Overview CMIP6-Endorsed MIPs

See updates http://www.wcrp-climate.org/index.php/wgcm-cmip/about-cmip

Date: 19 August 2015

Based on the assessment of the 10 CMIP6 endorsement criteria that were agreed at the WGCM 18th Session in October 2014, the CMIP Panel and WGCM co-chairs have now endorsed 21 Model Intercomparison Projects (MIPs).

The CMIP6 design will be described in a Geoscientific Model Development special issue with submissions of an overview paper and the CMIP6-Endorsed MIP contributions envisaged by end of March 2016. The description of the experiments and forcing data sets presented in this special issue will define CMIP6 in detail. Updated information on CMIP6 can also be found at the CMIP Panel website at http://www.wcrp-climate.org/index.php/wgcm-cmip/about-cmip.

Table 1. Overview of the 21 CMIP6-Endorsed MIPs endorsed by mid-August 2015. MIPs marked with * are “Diagnostic MIPS”.

	Short name of MIP	**Long name of MIP**
1	**AerChemMIP**	Aerosols and Chemistry Model Intercomparison Project
2	**C4MIP**	Coupled Climate Carbon Cycle Model Intercomparison Project
3	**CFMIP**	Cloud Feedback Model Intercomparison Project
4	**DAMIP**	Detection and Attribution Model Intercomparison Project
5	**DCPP**	Decadal Climate Prediction Project
6	**FAFMIP**	Flux-Anomaly-Forced Model Intercomparison Project
7	**GeoMIP**	Geoengineering Model Intercomparison Project
8	**GMMIP**	Global Monsoons Model Intercomparison Project
9	**HighResMIP**	High Resolution Model Intercomparison Project
10	**ISMIP6**	Ice Sheet Model Intercomparison Project for CMIP6
11	**LS3MIP**	Land Surface, Snow and Soil Moisture
12	**LUMIP**	Land-Use Model Intercomparison Project
13	**OMIP**	Ocean Model Intercomparison Project
14	**PMIP**	Palaeoclimate Modelling Intercomparison Project
15	**RFMIP**	Radiative Forcing Model Intercomparison Project
16	**ScenarioMIP**	Scenario Model Intercomparison Project
17	**VolMIP**	Volcanic Forcings Model Intercomparison Project
18	**CORDEX***	Coordinated Regional Climate Downscaling Experiment
19	**DynVar***	Dynamics and Variability of the Stratosphere-Troposphere System
20	**SIMIP***	Sea-Ice Model Intercomparison Project
21	**VIACS AB***	VIACS Advisory Board for CMIP6

Application for CMIP6-Endorsed MIPs

Please return to CMIP Panel Chair Veronika Eyring (email: Veronika.Eyring@dlr.de)

Date: 10 November 2014

The recently proposed, revised CMIP structure (see information on the CMIP Panel website at http://www.wcrp-climate.org/index.php/wgcm-cmip/about-cmip) provides for a small set of experiments to be routinely performed by modeling groups whenever they develop a new model version. The output from these so-called *ongoing CMIP Diagnostic, Evaluation and Characterization of Klima (DECK)* experiments and the *CMIP6 Historical Simulation* will be distributed for community use via the ESGF infrastructure. Other Model Intercomparison Projects (MIPs) will build on the CMIP DECK experiments and the CMIP6 Historical Simulation and augment them to address a broad range of scientific questions. Additionally proposed MIP experiments together with the CMIP DECK experiments and the CMIP6 Historical Simulation will constitute the suite of simulations for the next phase of CMIP.

MIPs are invited to request endorsement for the next phase of CMIP (i.e., CMIP6). Applications from MIPs requesting status as a CMIP6-Endorsed MIP should be sent to the CMIP Panel Chair. The current set of MIP proposals is now complete and will be revised on the agreed timeline. We will review any additional proposals in a year from now at the next WGCM meeting in October 2015. A MIP may propose that a subset or even all of their experiments be included as part of the suite of simulations constituting CMIP6. The CMIP Panel will, together with the WGCM co-chairs, decide whether a MIP and its experiments meet the criteria for endorsement for CMIP6. Note that it is expected that all additional experiments proposed for CMIP6 will be scientifically analyzed and exploited by the MIP.

CMIP6-Endsored MIPs can make full use of the ESGF infrastructure. In order to minimize the burden imposed on modeling groups wishing to participate, the MIPs seeking to be part of CMIP Phase X must agree to comply with the CMIP standards in terms of experimental design, data format and documentation. In general the WGCM encourages adhering to the standards in place for CMIP.

The main criteria for MIPs to be endorsed for CMIP6 are

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.
3. The MIP adopts the CMIP modeling infrastructure standards and conventions.
4. All experiments are tiered, well-defined, and useful in a multi-model context and don’t 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 modeling group.
6. A sufficient number of modelling centers (~8) are committed to performing all of the 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 CMIP6 Special Issue.
10. The MIP considers reporting on the results by co-authoring a paper with the modelling groups.

AerChemMIP (Aerosols and Chemistry MIP)
Application for CMIP6-Endorsed MIPs

Date: 10 June 2015

- Co-chairs of MIP

William Collins (UK) (W.Collins@reading.ac.uk)
Jean-François Lamarque (US) (lamar@ucar.edu)
Michael Schulz (Norway) (michael.schulz@met.no)

- Members of the Scientific Steering Committee

Olivier Boucher (France) (olivier.boucher@lmd.jussieu.fr)
Veronika Eyring (Germany) (veronika.eyring@dlr.de)
Arlene Fiore (US) (amfiore@ldeo.columbia.edu)
Michaela Hegglin (UK) (m.i.hegglin@reading.ac.uk)
Gunnar Myhre (Norway) (gunnar.myhre@cicero.oslo.no)
Michael Prather (US) (mprather@uci.edu)
Drew Shindell (US) (drew.shindell@duke.edu)
Steve Smith (US) (ssmith@pnnl.gov)
Darryn Waugh (US) (waugh@jhu.edu)

Goal of the MIP

Past climate change has been forced by a wide range of chemically reactive gases, aerosols, and well mixed greenhouse gases (WMGHGs), in addition to CO_2. Scientific questions and uncertainties regarding chemistry-climate interactions range from regional scales (e.g., tropospheric ozone and aerosols interacting with regional meteorology), to long-range connections (e.g., hemispheric transport of air pollution, the impacts of lower stratospheric ozone and temperatures on surface climate), to global integration (e.g., the lifetimes of CH_4 and N_2O).

AerChemMIP proposes to contribute to CMIP6 through the following: 1) diagnose forcings and feedbacks involving NTCF[1]s, (namely tropospheric aerosols, tropospheric O_3 precursors, and CH_4) and the chemically reactive WMGHGs (e.g., N_2O, also CH_4, and some halocarbons** including impacts on stratospheric O_3), 2) document and understand past and future changes in the chemical composition of the atmosphere, and 3) estimate the global-to-regional climate response from these changes.

---

[1] Near Term Climate Forcers

To participate in the CMIP6/AerChemMIP project models will need to be run for the CMIP6 DECK experiments with the same setup, i.e. with the same levels of sophistication activated in the chemistry and aerosol schemes. In particular it will be essential to have PI control and Historical simulations with the full chemistry (where used) and aerosols. It is also realised that valuable contributions to answering the AerChemMIP scientific questions can be made by groups unable to participate in CMIP6. Participation from these groups is welcomed and encouraged in the wider AerChemMIP project but the data will not form part of the official CMIP6 submission.

Recently, WCRP endorsed “Biogeochemical forcings and feedbacks” as a Theme of Collaboration, similar in scope to the Grand Challenges. AerChemMIP is ideally suited to provide significant contributions to this theme through simulations in all Tiers. In particular, simulations Tier 1.3 look at the role of methane changes (which have an important biogeochemical component) on the historical climate. In addition, air quality is a theme central to AerChemMIP (see Tier 1.2 for example), and therefore will fulfill some of the goals of “Biogeochemical forcings and feedbacks” highlighted in http://www.wcrp-climate.org/images/documents/jsc/JSC36/WCRP-GC-biogeochemistry.pdf.

The AerChemMIP Tier 1 simulations focus primarily on understanding atmospheric composition changes (from NTCFs and other chemically-active anthropogenic gases) and their impact on climate. We have devised a series of experiments that contrast the forcing of various NTCFs with that of WMGHGs in historical and future climate change. In addition, the proposed chemistry-climate simulations will enable diagnosis of changes in regional air quality (AQ) through its coupling to large-scale changes in O_3-CH_4-$PM_{2.5}$. We will work in collaboration with RFMIP and DAMIP to provide a comprehensive analysis of ERF and the regionally-resolved climate forcing signature from tropospheric NTCFs. For some of the specifically attributable climate forcings, in particular those at the 10s of mW m^{-2} level, the actual climate change will be difficult to detect in a transient simulation or even a time slice of several decades. AerChemMIP is a joint, consolidated effort for CMIP6 from two international communities -- Aerosol Comparisons between Observations and Models (AeroCom, http://aerocom.met.no/Welcome.html) and the IGAC/SPARC Chemistry-Climate Model Initiative (CCMI, http://www.met.reading.ac.uk/ccmi/). Experiments suggested for CCMI Phase 2 [*Eyring et al.*, 2013b], which are traditionally run using chemistry-climate models (CCMs) with mostly prescribed sea surface temperatures and sea ice concentrations, complement this set of AerChemMIP/CMIP6 experiments. Further experiments in AeroCom phase III aim to understand sensitivity of aerosol forcing to aerosol formation and loss processes.

We do not specifically consider the very long-lived F-gases (SF_6, PFCs, and some HFCs) as their abundance is not affected by chemistry on a century time scale.

Overview

Aerosols and ozone were identified in IPCC AR5 (Myhre et al., 2013) as the main sources of uncertainty in the radiative forcing since pre-industrial times. Uncertainties in projecting the chemically reactive WMGHGs as well as future air quality from global changes were also identified in AR5 [*Kirtman et al.*, 2013]. In addition to changing anthropogenic emissions evaluated in AR5, natural aerosols originating from biogenic sources, dust or sea-salt are a primary contributor to the uncertainty in current forcing (Carslaw et al. 2013). Due to the nonlinear response of clouds to the background level of aerosols, the response of the climate system to human perturbations will depend critically on the natural aerosol background (Carlton et al., 2010).

Beyond aerosols, the biogeochemistry of ecosystems provides large sources of the WMGHGs CH_4 and N_2O, as well as O_3 precursors (lightning and soil nitrogen oxides, volatile organic compounds, wildfire emissions). These sources are likely to be affected by climate change, leading to a variety of feedbacks that to date have only been quantified from a limited number of studies (and models) and thus demand for a coordinated set of simulations that allows for a consistent and clean comparison between models.

Anthropogenic emissions of NTCFs have been responsible for a climate forcing that is presently nearly equal in magnitude to CO_2-forcing. These emissions have led to a variety of global climate impacts such as regional patterns of temperature and precipitation, with a magnitude similar to the global-mean equivalent ERF of WMGHG. In addition, NTCF ERF is inherently inhomogeneous, and there is some evidence that where NTCF on a regional scale is large, the climate response differs from the globally equivalent ERF – i.e., there is some regional response to regional ERF.

NTCF emissions are also responsible for driving regional and local air quality (AQ). This has led to the recognition that a combined strategy of mitigating climate change and air pollution together has clear economic benefits compared to separate mitigation (IPCC, 2014 – WG3 SPM). In our future world, most, if not all scenarios lead to changes in the emissions and meteorology that determine air quality and create pollution episodes. The knowledge base used to manage air pollution to date must be updated based on more comprehensive information that CMIP6 will provide on future air chemistry climatologies. The exposure risks of human health and assets (agriculture, built environment, ecosystems) will be driven by daily variations in surface ozone and particulate matter in addition to deposition of nitrate and sulfate and any land-use change interacting with atmospheric changes.

The forcing of climate by ozone changes results from tropospheric increases and lower stratospheric decreases, with interaction between those. They are the result of combined impacts from climate change and multiple emission changes. For example, one of the largest components of CH_4 emissions' ERF is that from the increase in tropospheric O_3. In addition, stratospheric O_3 losses due to ozone depleting substances (ODS) since the 1970s has led to significant cooling of the lower stratosphere, and through the Antarctic ozone hole is linked to changes in tropospheric circulation and rainfall patterns in the southern hemisphere, especially during summer (WMO, 2014). In the Southern Hemisphere, future summertime circulation changes are controlled by both the ozone recovery rate and the rate of GHG increases [*Eyring et al.*, 2013a], indicating the need to account for ozone changes in future climate projections.

Since some models participating in CMIP6 do not have interactive chemistry schemes, AerChemMIP will also provide historical and future time-varying ozone, and stratospheric water vapour concentration fields for CMIP6. The ozone database will be an update of the database provided for CMIP5 by [Cionni et al., 2011]. These datasets will be generated from a mixture of CCMs and CTMs simulations which are not themselves part of CMIP6.

Because of the central role of aerosols in many of the AerChemMIP simulations and analysis, we suggest that climate models without prognostic aerosol schemes refrain from participating in AerChemMIP. It is important to note that the models participating in AerChemMIP must use for the corresponding, qualifying DECK and historical simulation emissions of aerosols and aerosol precursors provided by Smith et al. in the end of 2015. The CMIP6 aerosol forcing dataset shall not be used for DECK and historical simulations. However modeling groups are encouraged to participate with their AerChemMIP model version in the prescribed aerosol subset of RFMIP simulations, where the CMIP6 aerosol forcing dataset is required. In that case the aerosol radiative effects will have to be decoupled from the interactive aerosol scheme and prescribed optical properties from the RFMIP aerosol forcing dataset shall be used.

Overview of the Proposed Tier 1 Experiments

The AerChemMIP Tier 1 simulations focus on two science questions

1. How have NTCF and ODS emissions contributed to global ERF and affected regional climate over the historical period?
2. How will future policies (on climate/AQ/land use) affect the NTCFs and their climate impacts?
3. How have WMGHG concentrations forced climate (including through their chemical impacts) over the historical period?

In the following sections, we discuss each question separately and provide for each science question the description of the simulations necessary to answer the stated question. Note that we emphasize the use of the Effective Radiative Forcing (ERF) to measure climate forcing. We have provided at the end of this document a description of the methodology associated with this calculation.

1. How have NTCF and ODS emissions contributed to global ERF and affected regional climate over the historical period?

Anthropogenic non-CO_2 emissions (e.g., NTCFs, GHGs like halocarbons and N_2O,…) have led to a climate forcing that is commensurate to CO_2-forcing on regional scales, especially over the last few decades.

By way of their associated large uncertainty in radiative forcing since pre-industrial times, ozone and aerosols in particular are a key factor behind the large uncertainty in constraining climate sensitivity over the record of observed data. These NTCFs have an inhomogeneous spatial distribution and the degree of regional temperature and precipitation responses to such heterogeneous forcing remains an open question within the scientific community. It is further unclear whether NTCFs, which are primarily located at Northern Hemisphere mid latitude land areas have led to a larger climate response there, relative to forcing from WMGHGs.

One unambiguous regional response to inhomogeneous climate forcing concerns the Southern hemisphere summertime surface circulation changes induced by the Antarctic ozone hole as an indirect response to ozone-depleting halocarbons. These changes have been argued to lead to changes in rainfall patterns, ocean circulation, and sea-ice cover. The relative role of these ozone-induced changes compared to other anthropogenic forcings and natural variability is not fully resolved by the scientific community (with some studies reaching contradictory conclusions). Hence there is a need for multi-model ensemble of simulations, especially with models resolving stratospheric chemistry that isolate the role of stratospheric ozone depletion.

Experiment 1.1: Transient historical coupled ocean climate impacts of NTCFs and of ozone depleting halocarbons (note: this builds on CMIP6-historical-simulation, which is used as the reference simulation, and requires AerChemMIP diagnostics therein)

- 1.1.1 Perturbation: Historical WMGHG (including halocarbon) concentrations, 1850 NTCF emissions. 165 years, 1-3 ensemble members
- 1.1.2 Perturbation: Historical WMGHG concentrations and NTCF emissions, 1950 halocarbons. 65 years (branched from CMIP6 historical in 1950), 1 up to the number of ensemble members performed for the CMIP6 historical

Experiment 1.2: Estimating ERFs through specified transient historical SST simulations (see note on ERFs below).

Note that these simulations uses transient SST from AOGCM simulations in Experiment 1.1 and not constant SST over the historical period. Perform 1850-2014 (1 ensemble member only) simulation with all forcings as in CMIP6 historical **but** with

- 1.2.1 1850 all NTCF emissions (including biomass burning). 165 years
- 1.2.2 1950 ODSs. 65 years (1950-2014)

Experiment 1.3. Time-slice simulations based on the 1850 control SSTs to compute the ERF for 1850 and 2014 for all NTCF (e.g. AR5 fig 8.15). This requires two simulations

- 1.3.1 Control: 1850 WMGHG concentrations and 1850 NTCF emissions. 30 years
- 1.3.2 Perturbation: 1850 WMGHG concentrations, 2014 NTCF emissions. 30 years

2. How will future policies (on climate/AQ/land use) affect the NTCFs and their climate impact? What are the patterns of associated climate forcing, and how do these patterns translate into temperature and precipitation changes?

For the upcoming decades policy makers will be making choices in 3 broadly defined areas 1) climate change policies (targeting mostly WMGHGs), 2) air quality policies (targeting mostly NTCF emissions including CH4 that are precursors of tropospheric aerosols and tropospheric ozone) and 3) land-use policies. AerChemMIP aims to identify the patterns of chemical change at the global and regional levels as well as the ERF associated with NTCF mitigation efforts (focusing on policy choices in areas 1 and 2 above), and their climate (surface temperature and precipitation) and environmental (health, ecosystem, visibility, …) impact between 2015 and 2055 (this is the time frame over which aerosol and precursor emissions are expected to be significant). The impact analysis will be performed by contrasting the following simulations: a) SSP3-7 from ScenarioMIP (note that additional diagnostics will have to be included for those simulations to be useful for AerChemMIP) since it has high aerosol emissions and b) a perturbation experiment where air quality policies (or maximum feasible reductions) are applied to the SSP3-7 NTCF emissions, and therefore lead to much reduced NTCF emissions. These perturbations will be designed in collaboration with ScenarioMIP to ensure that perturbations are consistent with the underlying story line of the scenario in consideration.

Experiment 2.1: Transient coupled ocean climate impacts

- 2.1.1 Reference: SSP3-7 (to be performed under ScenarioMIP)
- 2.1.2 Perturbation: SSP3-7 with reduced NTCF (aerosol and tropospheric ozone precursors, including methane) 41 years, 1-3 ensemble members (3 recommended)

Experiment 2.2: Estimating ERFs through fixed-SST simulations (SSTs from 2.1.1)

- 2.2.1 Control: as Experiment 2.1.1 using archived SSTs from 2.1.1
  41 years, one ensemble
- 2.2.2 Perturbation: Only black carbon emissions as in Experiment 2.1.2 (this is to isolate the specific role of black carbon in near-term policy decisions)
  41 years, one ensemble
- 2.2.3 Perturbation: All aerosol precursor emissions (but not NOx) as in 2.1.2,
  41 years, one ensemble
- 2.2.4 Perturbation: All ozone precursors except methane kept the same as in 2.1.2,
  41 years, one ensemble
- 2.2.5 Perturbation: Methane kept the same as in 2.1.2,
  40 years, one ensemble

3. How have chemically reactive WMGHGs affected the forcing over the historical period?

Under this question, we focus in Tier one on estimating the forcing from changes in methane on ozone (tropospheric and stratospheric), aerosol oxidation, and emissions of natural aerosols, including the climate impacts associated with those changes. Note that only ERF estimates are calculated, while the associated transient coupled simulations are in Tier 2.

Experiment 3.1: Estimating ERFs through specified SST simulations (SSTs taken from CMIP6 historical simulation)

Perform 1850-2015 (1 ensemble member only) simulation with all forcings (and including chemistry feedbacks on tropospheric and stratospheric ozone) as in transient historical **but** with

3.1.1 1850 CH_4. 165 years

Total amount of Tier 1 simulation years

Experiments 1.x.x: 540y - 940y (overlap w DAMIP ca ???y-???y) (overlap w RFMIP ca ??y)
Experiments 2.x.x: 246y - 328y (excluding 2.1.1, run under ScenarioMIP)
Experiments 3.x.x: 165y

Synergy with other MIPs – Model diagnostics

Experiment 1.1.1/1.1.2 parallels similar forcing attribution simulations in DAMIP but include chemistry responses and diagnostics.
Experiments 1.2.4/1.2.5/3.2.1/3.2.2: These parallel similar ERF calculations in RFMIP, but start from emission changes rather than concentration changes
Experiments 2.1.1/2.1.2 extend the ScenarioMIP simulations to separate out the impact of AQ policies and NTCFs

Model diagnostics specific to AerChemMIP Tier 1 experiments need to be implemented also in the DECK and CMIP6-historical-simulation. The diagnostics will be contributed to the CMIP6 data request by AerChemMIP. If models have not all components to compute dynamic aerosols, tropospheric or stratospheric chemistry, models are requested to consider using the forcing fields of chemical compounds provided by AerChemMIP when performing AerChemMIP Tier 1 experiments.

Overview of the Proposed Tier 2 and 3 Experiments

AerChemMIP also includes additional experiments to document with an eventually more limited set of models complementing science questions, which are based on tier 1 experiments, and make efficient use of the general set-up of CMIP6. Tier 2 and 3 add detail to the Tier 1 experiments 1.1, 1.2, 1.3 and 3.1 by addressing extra combinations of NTCFs and reactive WMGHGs. They also add two additional science objectives.

4. Quantifying the climate feedbacks through changes in natural emissions (FDBCK)
5. Quantifying the uncertainties associated with anthropogenic emissions (ChemDOC)

Note that all except one Tier 2 and 3 simulations rely on AGCM simulations, i.e. sea-surface temperatures (and sea-ice distribution) are specified from an existing fully-coupled simulation in Tier 1.

Experiment 1.1: Transient historical coupled ocean climate impacts of NTCFs and of ozone depleting halocarbons

1.1.3 Perturbation: Historical WMGHG (including halocarbon) concentrations, 1850 aerosol (but not NOx) emissions, 165 years, 1-3 ensemble members . Tier 2

Experiment 1.2: Estimating ERFs through specified transient historical SST

1.2.3 1850 all tropospheric ozone precursor emissions. 165 years. Tier 2
1.2.4 1850 all aerosol emissions (except NOx). 165 years. Tier 2

Experiment 1.3. Time-slice simulations based on the 1850 control SSTs to compute the ERF for 1850 and 2014 (e.g. AR5 fig 8.15).

1.3.3 Perturbation: 1850 WMGHG concentrations, 2014 aerosol (not NOx) emissions. 30 years. Tier 2
1.3.4 Perturbation: 1850 WMGHG concentrations, 2014 BC emissions. 30 years. Tier 2
1.3.5 Perturbation: 1850 WMGHG concentrations, 2014 tropospheric ozone precursor emissions. 30 years. Tier 2
1.3.6 Perturbation: 1850 WMGHG (except CH4) concentrations, 1850 NTCF emissions, 2014 CH4 concentrations. 30 years. Tier2
1.3.7 Perturbation: 1850 WMGHG (except N2O) concentrations, 1850 NTCF emissions, 2014 N2O concentrations. 30 years. Tier 2
1.3.8 Perturbation: 1850 WMGHG (except ODS) concentrations, 1850 NTCF emissions, 2014 ODS concentrations. 30 years. Tier 2
1.3.9 Perturbation: 1850 WMGHG concentrations, 2014 NOx emissions. 30 years. Tier 3
1.3.10 Perturbation: 1850 WMGHG concentrations, 2014 CO/VOC emissions. 30 years. Tier 3

Experiment 3.1: Estimating ERFs through specified SST simulations

3.1.2 1850 N2O. 165 years. Tier 2

4. Quantifying the climate feedbacks through changes in natural emissions

Climate change will have affected (and will affect) the natural emissions of natural NTCFs and reactive WMGHGs. These natural emissions will have an ERF and so feedback onto climate change. Ideally experiments would analyse the effects of constant vs varying emissions, however this is too complex to implement. The experiments proposed here simply double the natural emissions. The ERFs of natural WMGHGs (e.g. wetland methane) are not calculated as these can be obtained from experiment 1.3.

Experiment 4.1: Quantifying the ERFs of double natural emissions (based on 1850) control.

4.1.1 1850 doubled dust emissions. 30 years. Tier 2
4.1.2 1850 doubled sea salt emissions. 30 years. Tier 2
4.1.3 1850 doubled emissions of oceanic DMS. 30 years. Tier 3
4.1.4 1850 doubled fire emissions. 30 years. Tier 3
4.1.5 1850 doubled biogenic VOCs. 30 years. Tier 3
4.1.6 1850 doubled lightning NOx. 30 years. Tier 3
4.1.7. 1850 doubled wetland emissions of methane. 30 years. Tier 3

In addition, we will study the chemistry impact of land-use changes through biogenic emissions (LU-NTCF). This will be performed with AerChemMIP diagnostics, but using the same protocol as the land-use change experiment in LUMIP, i.e. based on SSP3-7 but with land use from SSP1-2.6.

Total amount of simulation years. Tier 2 = 900-1230, Tier 3 =140

Model Diagnostics and Performance Metrics for Model Evaluation

AerChemMIP will contribute to the CMIP6 data request by suggesting aerosol and chemistry related output that is required for model evaluation (including the characterization of air quality extremes) and for diagnosing radiative forcings from NTCFs. In addition, AerChemMIP will contribute to the development of the Earth System Model Evaluation Tools (ESMValTool, [*Righi et al.*, 2014]), the documentation of aerosol parameters via the AeroCom tools and will include important chemistry-related diagnostics and performance metrics for CMIP6 model evaluation.

Design of Effective Radiative Forcing simulations.

The proposed simulations combine analysis of the effective radiative forcing (ERF) and the consequent climate impacts of NTCFs. The RF from WMGHGs will be provided by RFMIP.

The ERFs are calculated by comparing the net TOA radiation fluxes between two runs with the same SSTs but with perturbed NTCF emissions (see below). Internal variability (mainly clouds) generates considerable noise therefore 30 years of simulation are needed to characterize the present day ERF from NTCFs. Alternatively, models that can nudge their simulated model winds (only, towards meteorological analyses or previously generated meteorological fields) should be able to identify a statistically-significant signal with as little as 3 years of simulation. In a similar way a pair of runs driven by evolving SSTs but with and without evolving NTCF emissions will provide the time evolution of the NTCF ERF.
For the temperature and precipitation impacts, simulations with a coupled ocean are needed. Again, this requires a pair with and without evolving NTCF emissions in order to compute the impacts. The internal variability in the coupled ocean models is larger than with fixed SSTs, so at least 3 ensemble members will be needed.

The effective radiative forcing (ERF) was introduced in IPCC AR5 [*Boucher et al.*, 2013; *Myhre et al.*, 2013]. The definition is given as follows: '*ERF is the change in net TOA downward radiative flux after allowing for atmospheric temperatures, water vapour and clouds to adjust, but with surface temperature or a portion of surface conditions unchanged*'. This is different from the traditional radiative forcing (RF) concept where surface and tropospheric temperature and other variables such as water vapour and clouds must be kept fixed. Quantification of a climate driver by ERF and RF provides different results for some aerosol effects where the latter concept allows quantification of semi-direct effect and second indirect aerosol effect (ERF of aerosol-radiation interaction and aerosol-cloud interaction, respectively). For greenhouse gases RF and ERF are more similar in magnitude, but the latter has larger uncertainty.

Two ways to simulate ERF is currently used, namely; i) net TOA fluxes from fixed-sea surface temperature (SST) simulations and ii) regression of transient temperature response with the initial radiative perturbation [*Gregory et al.*, 2004]. The two methods for simulating ERF are illustrated in [*Boucher et al.*, 2013; *Sherwood et al.*, 2014]. Both ERF methods have their advantages and disadvantages [*Boucher et al.*, 2013; *Myhre et al.*, 2013]. The regression method can be applied to many of the typical CMIP runs, but require long runs (at least 30 years) with a significant radiative perturbation. The fixed-SST method can by applied to relatively small radiative perturbations, but not all modelling groups have access to fixed-SST type simulations.

The fixed-SSTs approach can further be applied with additional radiation calls to diagnose the various aerosol effects [*Ghan et al.*, 2012]. Separate diagnostics for shortwave and longwave changes are applied. To diagnose the indirect aerosol effect and semi-direct effect the scattering and absorption by aerosols are neglected by setting refractive indexes of anthropogenic aerosol to zero, see [*Ghan et al.*, 2012] for further details.

References

Boucher, O., et al. (2013), Clouds and Aerosols, 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*, edited by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J.

Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley, pp. 571-657, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Cionni, I., V. Eyring, J. F. Lamarque, W. J. Randel, D. S. Stevenson, F. Wu, G. E. Bodeker, T. G. Shepherd, D. T. Shindell, and D. W. Waugh (2011), Ozone database in support of CMIP5 simulations: results and corresponding radiative forcing, *Atmos. Chem. Phys. Discuss.*, *11*(4), 10875-10933.

Eyring, V., et al. (2013a), Long-term ozone changes and associated climate impacts in CMIP5 simulations, *J Geophys Res-Atmos*, *118*(10), 5029-5060.

Eyring, V., et al. (2013b), Overview of IGAC/SPARC Chemistry-Climate Model Initiative (CCMI) Community Simulations in Support of Upcoming Ozone and Climate Assessments, *SPARC Newsletter*, *40*, 48-66.

Ghan, S. J., X. Liu, R. C. Easter, R. Zaveri, P. J. Rasch, J.-H. Yoon, and B. Eaton (2012), Toward a minimal representation of aerosols in climate models: Comparative decomposition of aerosol direct, semi-direct and indirect radiative forcing, *J. Climate*, doi: 10.1175/JCLI-D-1111-00650.00651, in press.

Gregory, J. M., W. J. Ingram, M. A. Palmer, G. S. Jones, P. A. Stott, R. B. Thorpe, J. A. Lowe, T. C. Johns, and K. D. Williams (2004), A new method for diagnosing radiative forcing and climate sensitivity, *Geophysical Research Letters*, *31*(3), L03205.

Kirtman, B., et al. (2013), *Chapter 11. Near-term Climate Change: Projections and Predictability 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*, Cambridge University Press Cambridge, United Kingdom.

Myhre, G., et al. (2013), 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*, edited by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley, pp. 659-740, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Righi, M., V. Eyring, K.-D. Gottschaldt, C. Klinger, F. Frank, P. Jöckel, and I. Cionni (2014), Quantitative evaluation of ozone and selected climate parameters in a set of EMAC simulations, *Geosci. Model Dev. Discuss.*, *7*, 6549-6627.

Sherwood, S. C., S. Bony, and J. L. Dufresne (2014), Spread in model climate sensitivity traced to atmospheric convective mixing, *Nature*, *505*(7481), 37-+.

C^{4}MIP Application for CMIP endorsement

June 2015

The Coupled Climate Carbon Cycle MIP, C^{4}MIP, requests formal endorsement by WGCM for the next phase of CMIP (CMIP6).

Background and motivation

The carbon cycle is the key addition to physical climate models that makes them “Earth System Models” (ESMs). CMIP5 was the first CMIP phase to include ESMs as the standard climate change modeling tool and carbon cycle results featured strongly in the IPCC 5th Assessment Report (see for example WG1 SPM, TS, chapters 6, 9, 12 and WG2 chapter 4). WG1 SPM highlighted the direct link from anthropogenic emissions to global climate change through the policy relevant Transient Response to Cumulative Emissions (TCRE). This is a key advance over AR4 – an advance only possible due to the inclusion of the carbon cycle in physical climate models.

C^{4}MIP has been central in this development work, and the first C^{4}MIP intercomparison paper (Friedlingstein et al., J. Clim., 2006) now has more than 1000 citations. Extensive use was made of the carbon cycle simulations in the CMIP5 database which led to a *Journal of Climate* Special Issue titled “Climate–Carbon Interactions in the CMIP5 Earth System Models” (http://journals.ametsoc.org/page/C4MIP)

Further development, evaluation and assessment of carbon cycle processes are one of the key focus areas for global climate modelling centres and we fully expect carbon cycle processes to be of fundamental importance in CMIP6 (and beyond).

Proposals from MIPs should include the following information:

**   Preliminary information used to determine whether a MIP should be endorsed for CMIP6 or not.*
*** Information that must be provided later (and before the panel can determine which experiments, if any, will be incorporated in the official CMIP6 suite).*

➢ Name of MIP

Coupled Climate Carbon Cycle MIP (C^{4}MIP)

➢ Co-chairs of MIP (including email-addresses) *(alphabetical order)*

Vivek Arora, Canadian Centre for Climate Modelling and Analysis, Canada, vivek.arora@ec.gc.ca

Pierre Friedlingstein, University of Exeter, UK, p.friedlingstein@exeter.ac.uk

Chris Jones, Met Office Hadley Centre, UK, chris.d.jones@metoffice.gov.uk

➢ Members of the Scientific Steering Committee

Co-chairs plus a steering committee providing additional expertise:

  - Ocean biogeochemistry: Laurent Bopp (IPSL, France), John Dunne (GFDL, USA) and Tatiana Ilyina (MPI, Germany)
  - Nitrogen cycle: Sönke Zaehle (MIP, Germany)
  - Permafrost: Charles Koven (LBL, USA)
  - Observations: Heather Graven (Imperial College, UK), Martin Jung (MPI, Germany)
  - Evaluation/iLAMB: Forrest Hoffman (ORNL, USA), Jim Randerson (UC Irvine, USA)
  - Land use change/LUMIP: Julia Pongratz (MPI, Germany), Victor Brovkin (MPI, Germany)

Additional experts on related activities might be invited to attend SSC meetings (eg. land use forcing (LUMIP), methane emissions, offline analysis (OCMIP, TRENDY), Remote sensing data, TCRE, …)

Several of the SSC members are also members of other MIPs : ScenarioMIP (Friedlingstein), LUMIP (Jones, Brovkin, Pongratz), OCMIP (Bopp)

- **Link to website (if available)**

emps.exeter.ac.uk/mathematics/research/xcs/c4mip/

- **Goal of the MIP and a brief overview**
  - The primary focus of C4MIP is to understand and quantify future (century-scale) changes in land and ocean carbon storage and fluxes
  - Idealized experiments will be used to separate and quantify the sensitivity of land and ocean carbon cycle to changes in climate and changes in atmospheric CO_2 concentration
  - Historical experiments will be used to evaluate model performance and investigate potential for future constraints
  - Future scenario experiments will be used to quantify future changes in carbon storage and hence quantify the atmospheric CO_2 concentration and related climate change for given CO_2 emissions, or diagnose the emissions compatible with a prescribed atmospheric CO_2 concentration pathway

- **References (if available)**
  - Friedlingstein, P. and coauthors (2006): Climate–carbon cycle feedback analysis: Results from the C4MIP model intercomparison, Journal of Climate, 19(14), 3337-3353
  - Arora, V.K. and coauthors (2013) Carbon–Concentration and Carbon–Climate Feedbacks in CMIP5 Earth System Models. Journal of Climate, Vol. 26, Iss. 15, pp. 5289-5314.
  - Friedlingstein, P. and coauthors (2013) Uncertainties in CMIP5 Climate Projections due to Carbon Cycle Feedbacks. Journal of Climate, Vol. 27, Iss. 2, pp. 511-526.
  - Jones, C. D. and coathors (2013) Twenty-First-Century Compatible CO_2 Emissions and Airborne Fraction Simulated by CMIP5 Earth System Models under Four Representative Concentration Pathways. Journal of Climate, Vol. 26, Iss. 13, pp. 4398-4413.
  - Gillett, N. P., et al. (2013) Constraining the Ratio of Global Warming to Cumulative CO2 Emissions Using CMIP5 Simulations. Journal of Climate, Vol. 26, Iss. 18, pp. 6844-6858.
  - Schwinger J., et al. (2014) Non-linearity of ocean carbon cycle feedbacks in CMIP5 earth system models. Journal of Climate, Vol. 27, Iss. 11, pp, 3869–3888.
  - IPCC AR5 WG1 SPM, TS and Chapters 6, 9 and 12
  - Anav, A., et al. "Evaluating the land and ocean components of the global carbon cycle in the CMIP5 Earth System Models." Journal of Climate 26.18 (2013): 6801-6843.

- Todd-Brown, K. E. O., et al. "Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations." Biogeosciences 10.3 (2013): 1717-1736.
- 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.

- **An overview of the proposed experiments**

The science background, proposed experimental design, proposed analysis and required diagnostics will be fully documented in a paper to be submitted to the CMIP6 GMD special issue.

C4MIP will build on the DECK and the CMIP6 Historical Simulation. The following simulations are pre-requisite for C4MIP participation:

DECK simulations

Control simulation

Requested for diagnosis of model drift (drift in land and ocean carbon pools)

- C4MIP expects a control run for each configuration of a model which contributes. Concentration-driven control run is required, with prescribed pre-industrial atmospheric CO2, to parallel concentration driven simulations. **An emissions-driven control run (emissions set to zero, but simulated evolution of atmospheric CO2 concentration) is *required* to parallel emissions driven simulations**.

CMIP 1% per year increasing CO_2 up to $4xCO_2$ simulation

Requested as baseline for the C4MIP climate-carbon cycle feedback analysis; also requested for assessment of TCRE. Nitrogen deposition in this simulation should be held fixed at the pre-industrial values used for the control run.

CMIP6 Historical Simulation

CMIP6 Emissions-driven Historical Simulation with interactive CO2 and forced by **prescribed CO2 emissions**

- As for the control run, C4MIP expects historical simulations with each configuration of the model. **Hence an emissions-driven historical simulation is *required***. Will be used for model evaluation and starting conditions for scenario simulations from Tier1

C4MIP SIMULATIONS

C4MIP- Tier 1

1.1: 1%BGC ("esm1pcbgc"): biogeochemically-coupled version of 1% per year increasing CO_2 up to $4xCO_2$ simulation. CO_2 increase only affects carbon cycle models, radiative code sees pre-industrial CO_2

Requested for the C4MIP climate-carbon cycle feedback analysis. Nitrogen deposition in this simulation should be held fixed at the pre-industrial values used for the control run.

1.2: Emission-driven future scenario (SSP-based RCP SSP5-8.5) up to 2100 ("esmssp5-85"): Requested for analysis of impact of carbon cycle feedbacks on climate projections over the 21st century. Also requested for assessment of cumulative emissions compatible with climate targets. Starting conditions taken from emissions-driven Historical simulation.

C4MIP- Tier 2

2.1: 1%RAD ("esm1pcrad"): radiatively-coupled version of the 1% per year increasing CO_2 up to $4xCO_2$ simulation. CO_2 increase only affects the radiative code, carbon cycle models see pre-industrial CO_2

Requested for further C4MIP climate-carbon cycle feedback analysis (non-linearities/synergies). Nitrogen deposition in this simulation should be held fixed at the pre-industrial values used for the control run.

2.2: 1% with N-deposition: additional feedback simulations for models with an interactive nitrogen cycle*.

2.2.1 ("esm1pccouNdep"): 1%COU-Ndep. Fully coupled, concentration driven 1% simulation with time varying anthropogenic nitrogen deposition**.

2.2.2 ("esm1pcbgcNdep"): 1%BGC-Ndep. BGC-coupled, concentration driven 1% simulation with time varying anthropogenic nitrogen deposition.

* note, "interactive nitrogen cycle" could mean either or both terrestrial or marine. Any model whose carbon stores/fluxes would be affected by the presence or absence of nitrogen deposition to perform these runs

** N-deposition fields to be provided by C4MIP in conjunction with AerchemMIP. Details TBC.

2.3: Concentration-driven CMIP6 Historical/SSP5-8.5 simulation, BGC mode

Requested for assessment of CO2-carbon cycle feedbacks over the 21st century; also for assessment of CO2 induced warming. Extension to 2300 recommended. Requires groups to have performed the C-driven SSP5-8.5 scenario (and extension) in ScenarioMIP.

2.3.1: HistBGC ("esmhistbgc"). Historical, concentration-driven, simulation parallel to standard Historical run, but with radiative effects of CO2 disabled – i.e. the radiation code is fed the time-invariant CO2 concentration from the control run.

2.3.2: SSP5-8.5-BGC ("esmssp5-85bgc"). SSP5-RCP8.5, concentration-driven, simulation to 2100 parallel to standard SSP5-8.5 simulation from ScenarioMIP, but with radiative effects of CO2 disabled – i.e. the radiation code is fed the time-invariant CO2 concentration from the control run.

2.3.3: SSP5-8.5-EXT -BGC ("esmssp5-85extbgc"). Extension of 2.3.2 to 2300 following the "high" extension of SSP5-8.5 decided on in ScenarioMIP.

SUMMARY OF C4MIP PROPOSAL

Category	Type of Scenario	Emission or concentration driven	Coupling mode	Simulation years	Short name	Use by other MIPs
Tier 1						
1%BGC	Idealised 1% per year CO2 only, BGC mode	C-driven	CO2 affects BGC	140	esm1pcbgc	OCMIP, LS3MIP
SSP5-8.5	SSP5-8.5 up to 2100	E-driven	Fully coupled	85	esmssp5-85	ScenarioMIP, LUMIP, OCMIP, LS3MIP
Tier 2						
1%RAD	Idealised 1% per year CO2 only, RAD mode	C-driven	CO2 affects RAD	140	esm1pcrad	OCMIP, LS3MIP
1%COU-Ndep	Idealised 1% per year CO2 only, fully coupled, increasing N-deposition	C-driven	Fully coupled	140	esm1pccou Ndep	OCMIP
1%BGC-Ndep	Idealised 1% per year CO2 only, BGC mode, increasing N-deposition	C-driven	CO2 affects BGC	140	esm1pcbgc Ndep	OCMIP
Hist/SSP 5-8.5-BGC	Historical+SSP5-8.5 up to 2300, BGC mode	C-driven	CO2 affects BGC	i. 155 ii. 85 iii. 200	esmhistbgc, esmssp5-85bgc and esmssp5-85extbgc	ScenarioMIP, OCMIP, LS3MIP, DAMIP

➢ An overview of the proposed evaluation/analysis of the CMIP DECK and CMIP6 experiments

For CMIP5 a Special Collection of C4MIP papers was assembled in Journal of Climate. Whilst we have not yet decided on specific papers for CMIP6 this was successful last time and may be pursued again. Key analysis papers include:

- an update on the climate carbon cycle feedback and it’s components (gain, beta, gamma on land and ocean)
- an analysis of the historical (emissions driven) scenario and its evaluation
- an analysis of the emissions driven scenario and quantification of the role of carbon cycle feedbacks under future climate change

We envisage these key papers will be co-authored with the modeling community.

C4MIP for CMIP6 will have a strengthened focus on model evaluation against observation-based estimates of carbon quantities.

Using the emission-driven historical simulation from DECK:

- Model evaluation: coordinated top-down and bottom-up metrics of performance for key land and ocean quantities (as in Anav et al., 2013, Hoffman et al., 2014), in emission driven historical simulations: analysis of simulated atmospheric CO_2 (concentration and isotopic composition) and evaluation against long-term observations (eg. Mauna Loa). Work in conjunction with European project, CRESCENDO.
- Use of emerging constraints based on carbon cycle interannual variability to constrain future projections (as in Cox et al., Nature 2013, Wenzel et al., JGR 2014) using the historical simulation and the CMIP1% simulations
- Quantify and explain changes since CMIP5 (show "demonstrable progress")
- Link with WGCM/WGNE metric panel for essential carbon cycle variables and with Obs4MIP for observation datasets.

Using the CMIP 1% DECK and the 1% BGC, RAD simulations from C4MIP:

- Quantification of the strength of carbon-concentration and carbon-climate feedbacks terms of Friedlingstein et al. (JClim. 2006) and their non-linearities (as in Gregory et al. JClim. 2009), assessment of magnitude, uncertainty, and changes since CMIP5
- Role of external forcing of the nitrogen cycle by anthropogenic N-deposition in influencing carbon cycle feedbacks and sensitiviteis.
- Assessment of the transient climate response to cumulative carbon emissions (TCRE), its magnitude, uncertainty and changes since CMIP5 as well as the impact of climate change on TCRE (using COU and BGC simulations).
- Quantification of response of natural CH_4 and N_2O emissions to climate and CO_2 changes

Using the SSP5-8.5 simulations

- In the emission driven case, these simulations will allow to quantify the effect of climate change (and land use change) on the global carbon cycle and atmospheric CO_2 and hence on the climate response (when compared to the SSP5-8.5 concentration driven simulation from ScenarioMIP)
- Quantification the uncertainty in simulated atmospheric CO_2 concentration (emissions-driven simulations).
- Analysis of TCRE based on SSP scenarios and its comparison with TCRE estimated from the idealized CMIP 1% simulations. Characterization of cumulative emissions allowed to likely stay below a given climate target (as in IPCC AR5).
- Analysis of changes in land and ocean carbon pools for future scenarios as a result CO_2 increase, climate change and of anthropogenic LUC (in coordination with LUMIP).
  - C4MIP E-driven SSP5-8.5 to form the control for LUMIP E-driven simulation with alternative land-use scenario (from SSP1-2.6).
- Assessment of risk of longer term carbon release from permafrost, vegetation dieback, change in oceanic circulation and impact on ocean carbon sink for the extension up to 2300.
- For the SSP5-8.5, BGC mode: analysis of CO2-carbon cycle feedbacks over the 21st century in a scenario world (as opposed to the idealized 1% world); assessment of CO2 induced warming (by comparison with the fully coupled scenario run); assessment of processes relating carbon uptake to ocean heat uptake and how feedback metrics vary on long timescales (especially in the ocean).

- Proposed timing*

We propose to ask modeling groups to provide results by end of 2016 from their COU and BGC 1% per year CO2 simulations. The analyses of historical and scenarios will depend on when their results become available (to be coordinated with Historical CMIP6 and ScenarioMIP).

- **Synergies with other MIPS**

ScenarioMIP

ScenarioMIP will coordinate the concentration-driven scenario simulations. C4MIP will coordinate the emission-driven scenario simulations. This will allow investigating the impact of carbon cycle feedbacks on climate projections. It will hence confirm (or infirm) the CO_2 concentration pathways used in ScenarioMIP and provided by the IAM models.

LUMIP

Scenario simulations will include land-use change as a forcing. The analysis of its impact on land carbon cycle and climate system

OCMIP

Analysis of oceanic response in 1% and SSP scenarios will be done in collaboration with OCMIP.

LS3MIP

Analysis of land response in 1% and SSP scenarios will be done in collaboration with LS3MIP.

DAMIP

Analysis of the emission-driven historical run in fully coupled mode and BGC only mode will help detection and attribution of historical role of CO2 emissions on climate and carbon cycle.

- **Data request**

Sent separately

Output request will require further harmonization with LUMIP, LS3MIP and OCMIP requests.

- **Model diagnostics and performance metrics for model evaluation**
    - We propose to use iLAMB/ESMVal and other software evaluation packages for evaluation of the carbon cycle in the historical emissions forced simulations

- For each proposed experiment to be included in CMIP6**
  - the experimental design;
  - the science question and/or gap being addressed with this experiment;
  - possible synergies with other MIPs;
  - potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, and (D) policy makers.
- If possible, a prioritization of the suggested experiments, including any rationale**
- All model output archived by CMIP6-Endorsed MIPs is expected to be made available under the same terms as CMIP output. Most modeling groups currently release their CMIP data for unrestricted use. If you object to open access to the output from your experiments, please explain the rationale.**
- List of output and process diagnostics for the CMIP DECK/CMIP6 data request**

  - whether the variable should be collected for all CMIP6 experiments, or only some specified subset and whether the output is needed from the entire length of each experiment or some shorter period or periods;
  - whether the output might only be relevant if certain components or diagnostic tools are used interactively (e.g. interactive carbon cycle or atmospheric chemistry, or only if the COSP simulator has been installed);
  - whether this variable is of interest to downstream users (such as impacts researchers, WG2 users) or whether its principal purpose is for understanding and analysis of the climate system itself. Be as specific as possible in identifying why the variable is needed.
  - whether the variables can be regridded to a common grid, or whether there is essential information that would be compromised by doing this;
  - the relative importance of the various variables requested (indicated by a tiered listing) is required if the data request is large.
- Any proposed contributions and recommendations for**
  - model diagnostics and performance metrics for model evaluation;
  - observations/reanalysis data products that could be used to evaluate the proposed experiments. Indicate whether these are available in the obs4MIPs/ana4MIPs database or if there are plans to include them;
  - tools, code or scripts for model benchmarking and evaluation in open source languages (e.g., python, NCL, R).
- Any proposed changes from CMIP5 in NetCDF metadata (controlled vocabularies), file names, and data archive (ESGF) search terms.**
- Explanation of any proposed changes (relative to CMIP5) that will be required in CF, CMOR, and/or ESGF.**

Cloud Feedback Model Intercomparison Project (CFMIP) Application for CMIP6-Endorsed MIPs

Mark Webb, Chris Bretherton, Sandrine Bony, Jen Kay, Steve Klein, Pier Siebesma,
Bjorn Stevens, George Tselioudis, Masahiro Watanabe,
Peter Good, Timothy Andrews, Roger Marchand, Robin Chadwick and Hervé Douville

Updated 31st March 2015

The primary goal of CFMIP is to inform improved assessments of climate change cloud feedbacks. However, the CFMIP approach is increasingly also being used to understand other aspects of climate response, such as regional-scale precipitation and non-linear changes.

CFMIP started in 2003 and its first phase (CFMIP-1) organised an intercomparison based on perpetual July SST forced Cess style +2K experiments and 2xCO_2 equilibrium mixed-layer model experiments containing ISCCP simulator in parallel with CMIP3 (McAvaney and Le Treut, 2003). Results from CFMIP-1 had a substantial impact on the evaluation of clouds in models and in the identification of low level cloud feedbacks as the primary cause of inter-model spread in cloud feedback, and featured prominently in the fourth and fifth IPCC assessments.

The subsequent objective of CFMIP-2 was to inform improved assessments of climate change cloud feedbacks by providing better tools to support evaluation of clouds simulated by climate models and to understand cloud-climate feedback processes. CFMIP-2 organized further experiments as part of CMIP5, introducing seasonally varying SST perturbation experiments for the first time, as well as fixed SST CO_2 forcing experiments to examine cloud adjustments, and idealized 'aquaplanet' experiments to establish the contributions of land and zonally asymmetric circulations to cloud feedback uncertainties (Bony et al., 2011). CFMIP-2 also introduced satellite simulators to CMIP via the CFMIP Observation Simulator Package (COSP), not only the ISCCP simulator, but additional simulators to facilitate the quantitative evaluation clouds using a new generation of active RADARs and LIDARs in space. Additionally CFMIP-2 introduced into CMIP5 process diagnostics such as temperature and humidity budget tendency terms and high frequency 'cfSites' outputs at 120 locations around the globe. CFMIP also organized a joint project with the GEWEX Global Atmospheric System Study (GASS) called CGILS (the CFMIP-GASS Intercomparison of LES and SCMs) to develop cloud feedback intercomparison cases to assess the physical credibility of cloud feedbacks in climate models by comparing Single Column Models (SCM) versions of GCMs with high resolution Large Eddy Simulations (LES) models. Additionally CFMIP-2 developed the CFMIP-OBS data portal and the CFMIP diagnostic codes repository (see http://www.cfmip.net for more details).

Early studies arising from CFMIP-2 include numerous model evaluation studies using COSP, studies attributing cloud feedbacks and cloud adjustments to different cloud types, and the finding that idealized 'aquaplanet' experiments without land or Walker circulations are able to capture the essential differences between models' global cloud feedbacks and cloud adjustments. Process outputs from CFMIP have also been used to develop and test physical mechanisms proposed to explain and constrain inter-model spread in cloud feedbacks in the CMIP5 models. CGILS has demonstrated a consensus in the responses of LES models to climate forcings and identified a number of shortcomings in the physical representations of cloud feedbacks in climate models. Additionally the CFMIP experiments have, due to their idealized nature, proven useful in a number of studies not directly related to clouds, but instead analyzing the responses of regional precipitation and circulation patterns to CO_2 forcing and climate change. Studies using CFMIP-2 outputs from CMIP5 remain ongoing and many further results are expected to feed into future assessments of the representation of clouds and cloud feedbacks in climate models. For a list of publications arising from CFMIP-2, please refer to the CFMIP publications page at http://www.cfmip.net.

Given the previous record of CFMIP activities and the case outlined below we would like to request that CFMIP be endorsed as a CMIP6 project to continue support for community activities in this important area of research. We provide information on our plans for CFMIP-3 structured according to the provided criteria below.

Name of MIP: The Cloud Feedback Model Intercomparison Project (CFMIP)

Co-chairs: Mark Webb mark.webb@metoffice.gov.uk, Chris Bretherton breth@washington.edu

Members of the Scientific Steering Committee: Mark Webb (Met Office), Chris Bretherton (U. Washington), Sandrine Bony (IPSL), Jen Kay (CIRES), Steve Klein (PCMDI), Pier Siebesma (KNMI), Bjorn Stevens (MPG), George Tselioudis (NASA GISS), Masahiro Watanabe (U. Tokyo)

Link to website: http://www.cfmip.net

Goal of the MIP and a brief overview: The primary goal of CFMIP is to inform improved assessments of climate change cloud feedbacks. However, the CFMIP approach is increasingly being used to understand other aspects of climate response, such as circulation, regional-scale precipitation and non-linear changes. This involves bringing climate modelling, observational and process modelling communities closer together and providing better tools and community support for evaluation of clouds and cloud feedbacks simulated by climate models and for understanding of the mechanisms underlying them. This is to be achieved by:

- Ongoing organized coordinated model inter-comparison activities which include experimental design as well as specification of model output diagnostics to support quantitative evaluation of modelled clouds with observations (e.g. COSP) and in-situ measurements (e.g. cfSites) as well as process-based investigation of cloud maintenance and feedback mechanisms (e.g. cfSites, budget tendency terms, etc.)
- Ongoing development and improvement of COSP and CFMIP-OBS infrastructure.
- Ongoing collaboration with the cloud process modelling community (via GASS collaboration) on CGILS and via new efforts to develop a hierarchy of experiments linking GCMs with cloud resolving models (CRMs) and Large Eddy Simulation (LES) models run on large domains (e.g. via the IMPULSE project consortium).
- Organising annual meetings to provide a focus for community activities relevant to CFMIP and also to the broader community working to understand changes in clouds, circulation and precipitation which impact regional projections of climate change. (These two communities are increasingly becoming connected because the experiments designed for CFMIP are also useful in addressing a broader range of questions not directly related to clouds.)

References:

- Andrews, T., (2014), Using an AGCM to diagnose historical effective radiative forcing and mechanisms of recent decadal climate change. *J. Climate*, 27, 1193–1209, doi:10.1175/JCLI-D-13-00336.1.
- Bony, S., Webb, M., Bretherton, C. S., Klein, S. A., Siebesma, P., Tselioudis, G., & Zhang, M. (2011). CFMIP: Towards a better evaluation and understanding of clouds and cloud feedbacks in CMIP5 models. Clivar Exchanges, 56(2), 20-22.
- Good, P., Andrews, T., Bouttes, N., Chadwick, R., Gregory, J. M., Lowe, J. A. (2014). The nonlinMIP intercomparison project: physical basis, experimental design and analysis principles. In preparation; (attached)
- McAvaney BJ, Le Treut H (2003) The cloud feedback intercomparison project: (CFMIP). In: CLIVAR Exchanges—supplementary contributions. 26: March 2003.
- Skinner, C.B., M. Ashfaq, and N.S. Diffenbaugh (2012). Influence of twenty-first-century atmospheric and sea surface temperature forcing on West African climate. J. Climate, 25, 527-542.

- Stevens B., Bony S., Frierson, D.M, Jakob, C., Kageyama, M., Pincus, R, Shepherd, T., Sherwood, S., Siebesma, A. P., Sobel, A., Watanabe, M., Webb, M.J. (2014). Clouds, Circulation and Climate Sensitivity: A Grand Science Challenge. World Climate Research Programme Report No. 8/2014

We argue below the CFMIP and its proposed experiments meet the requirements laid out by the CMIP panel, as outlined below.

1. CFMIP and its experiments directly address the key science questions of CMIP6. The question that CFMIP most directly addresses is `How does the Earth system respond to forcing?' The CFMIP emphasis on understanding cloud feedbacks makes CFMIP highly relevant to this question. The next most relevant question is `What are the origins and consequences of systematic model biases?' CFMIP has a strong model evaluation component via the use of satellite simulators, process diagnosis and comparisons with LES, and a proven track record in investigating the link between errors in cloud processes and cloud feedbacks. CFMIP is also relevant to the question `How can we assess future climate changes given climate variability, climate predictability, and uncertainties in scenarios?' CFMIP will continue to supplement fully coupled CMIP experiments with idealised experiments that focus on basic understanding of the dominant uncertainties associated with cloud feedbacks. This will continue to support work which relates variability on observable timescales (e. g. seasonal to decadal) to longer term climate change responses (e.g. via 'emergent constraints'). For example the amipPiForcing experiment proposed below will support studies relating cloud variability and feedbacks on observable timescales to long term cloud feedbacks (Andrews, 2014).

Note also that the WCRP Grand Challenge on Clouds, Circulation and Climate is led by two CFMIP committee members (Bony and Stevens), and has three additional CFMIP committee members on its steering committee (Webb, Siebesma, Watanabe), including one of the CFMIP co-chairs. This puts CFMIP in an excellent position to directly address the questions arising from the WCRP Grand Challenge.

2. CFMIP builds on and connects to the shared CMIP DECK and CMIP6 historical experiments. The AMIP experiment is the control simulation for the CFMIP amip4K, amip4xCO2 and amipFuture experiments which were proposed by CFMIP for CMIP5 and which we would like to see continued in CMIP6 as Tier I experiments. The proposed Tier II experiments also connect to the AMIP DECK experiment; the AMIP preindustrial forcing experiment and amip minus 4K experiments also use the DECK AMIP experiment as a control. The abrupt +/- 4% solar constant experiments build on and contrast with the DECK abrupt4xCO2 experiment, as do the abrupt4xCO2 and abrupt0.5CO2 experiments. Additionally the atmosphere-only timeslice experiments build on the abrupt4xCO2 experiment, decomposing the regional response of each model's abrupt4xCO2 run into separate responses to each aspect of forcing and warming. Additionally CFMIP will propose additional process diagnostics and simulator outputs for the CMIP6 historical experiment, which will allow process based comparisons with the AMIP experiments to assess the impact of coupled SST errors on the simulation of clouds and regional precipitation patterns in the CMIP6 models.

3. CFMIP will continue to follow the CMIP modeling infrastructure standards and conventions, in terms of experimental design, data format and documentation. CFMIP-2 experiments were organized as part of CMIP5 and the CFMIP co-chairs have demonstrated the ability to follow all of the relevant standards in experimental protocols, in specification of diagnostic output requests, data formats and documentation. We commit to continuing in this spirit for CFMIP experiments which are coordinated through CMIP6.

4. All experiments are tiered, well-defined, and useful in a multi-model context and don't overlap with other CMIP6 experiments.

These are outlined below, and detailed specifications are provided in the accompanying spreadsheet. They are tiered into Tiers I and II. Additionally we give guidance on other experiments currently under development which we may propose as additional Tier II experiments in the future. Alternatively these additional experiments may be coordinated outside of CMIP.

These experiments are we believe useful in the multi-model context because the common purpose that they share is a focus on understanding the inter-model uncertainty/spread in cloud adjustments and cloud feedbacks as well as that in regional precipitation and circulation change and non-linear change. Investigation of inter-model requires multi-model analysis and hence all of these experiments are useful (and in fact require) a multi-model context. The usefulness of the Tier I experiments to a number of climate researchers has already been demonstrated by the large number of publications produced using CFMIP-2 experiments.

We have checked for overlaps with other CMIP6 experiments and are confident that links with other MIPS (e.g. nonLinMIP, GeoMIP, SolarMIP, RFMIP and PMIP) are based on complementary but non-overlapping experiments.

Summary of proposed experiments

Tier I Science questions, activities and experiments

1.1 Continuation of CFMIP-2 experiments - Lead coordinator: Mark Webb (Met Office)

Science Question: What are the physical mechanisms underlying the range of cloud feedbacks and cloud adjustments predicted by climate models, and which models have the most credible cloud feedbacks?

The CMIP5/CFMIP-2 experiments and diagnostic outputs have enabled considerable progress on these questions but participation by a larger fraction of modelling groups is required in CMIP6 for a more comprehensive assessment of the uncertainties across the full multi-model ensemble. Our proposal is essentially to retain the CFMIP-2/CMIP5 experiments in Tier I for CMIP6. The experiments to be retained are amip4K, amip4xCO2, amipFuture, aquaControl, aqua4xCO2 and aqua4K. These build on the amip DECK experiment. As the output requirements for the DECK are not yet finalised, it is possible that the DECK AMIP experiment will not contain all of the output diagnostics required for CFMIP. For this reason we also request an additional CFMIP AMIP experiment including the full set of CFMIP diagnostics, both for model evaluation and for interpretation of feedbacks and adjustments in conjunction with other Tier I CFMIP experiments. If all of the proposed CFMIP diagnostics are included in the DECK experiment, this additional CFMIP AMIP experiment will not be required.

Tier II Science questions, activities and experiments

2.1 Abrupt +/-4% Solar Forced AOGCM experiments - Lead coodinators: Chris Bretherton (UW), Roger Marchand (UW), Bjorn Stevens (MPI)

Science Question: How do responses in the climate system due to changes in solar forcing differ from changes due to CO_2, and is the response sensitive to the sign of the solar forcing?

Rapid adjustments in clouds and precipitation are now recognized as significant components of models' responses to CO_2 forcing. While they can easily be separated from conventional feedbacks in SST forced experiments, such a separation in coupled models is complicated by various issues, including the response of the ocean on decadal timescales. A number of studies have examined cloud feedbacks in coupled models subject to a solar forcing, which is generally associated with much smaller cloud and precipitation adjustment, due to a smaller atmospheric absorption for a given top of atmosphere forcing. Solar forcing also has a weaker impact on the stratosphere than CO_2, potentially resulting in different upper tropospheric meridional temperature gradients and storm track responses.

A +4% solar experiment would be equivalent to the abrupt4xCO2 experiment but would increase the solar constant abruptly by 4 percent, resulting in a radiative forcing of a similar magnitude to that due to CO_2 quadrupling. This would provide a useful complement to the DECK abrupt4xCO2 experiment, and would support our understanding of regional responses of the coupled system with and without CO_2 adjustments. A complementary -4% abrupt solar forcing experiment would allow the examination of feedback asymmetry under climate cooling, and would also help with the interpretation of model responses to geo-engineering scenarios and volcanic forcing, and relate to past climates.

2.2 Abrupt2xCO2 and abrupt0.5xCO2 Experiments (nonLinMIP) - Lead Coordinator Peter Good (Met Office Hadley Centre)

Science Question: To what extent is regional-scale climate change per CO_2 doubling state-dependent (nonlinear), and why? How does the balance of mechanisms differ for high-forcing compared to low-forcing scenarios or paleoclimate simulations?

To address this question we propose two new experiments for Tier II, abrupt2xCO2 and abrupt0.5xCO2, to explore global and regional-scale nonlinear responses, highlighting different behavior under business-as-usual scenarios, mitigation scenarios and paleoclimate simulations. Additional experiments may be proposed for Tier II in the future, or coordinated via CFMIP outside of CMIP6. These include 100-year extensions to abrupt4xCO2 and abrupt2xCO2; a 1% ramp-down from the end of the 1pctCO2 experiment; an abrupt step-down to 1xCO2 from year 100 of the abrupt4xCO2. These would be used to explore longer-timescale responses, quantify nonlinear mechanisms more precisely and understand the reversibility of climate change.

2.3 amipMinus4K Experiment: Lead Coordinator: Mark Webb (Met Office)

Science Question: Are cloud feedbacks symmetric when subject to climate cooling rather than warming, and if not, why not?

An amipMinus4K experiment would take a similar form to the amip4K experiment, except that the sea surface temperatures would be uniformly reduced by 4K. This will be used to investigate asymmetric responses of clouds to a cooling climate in an idealized experiment, providing a link to PMIP. This experiment also complements the abrupt0.5xCO2 and the -4% solar experiments in that one can identify asymmetries in the warming/cooling response with and without interactions with the ocean. This experiment has been proposed for CFMIP following discussions with PMIP representatives (Pacale Braconnot, Masa Kageyama, and Masakazu Yoshimori).

2.4 Feedbacks in AMIP experiments: Lead Coordinator: Tim Andrews (Met Office)

Science question: Are climate feedbacks during the 20th century different to those acting on long term climate change and climate sensitivity?

Experiment and rationale: The previous CFMIP design was unable to diagnose time-dependent feedbacks that potentially undermine the simple linear forcing-feedback paradigm and which may be relevant to the gap between observed and modeled estimates of climate sensitivity. To address this we propose an additional experiment called 'amipPiForcing' (amip pre-industrial forcing), which is exactly the same as the standard amip run (i.e. SSTs and sea-ice) but run for the period 1870-present with constant pre-industrial forcings (i.e. all anthropogenic and natural forcing boundary conditions identical to the piControl run). Since the forcing constituents do not change in this experiment it readily allows a simple diagnosis of the simulated atmospheric feedbacks to observed SST changes, which can then be compared to feedbacks representative of long term change and climate sensitivity (e.g. from abrupt4xCO2 or amip4K). This has an advantage over the alternative approach of first estimating the forcing and adjustments (e.g. from RFMIP) and removing them from the amip experiment since the approach here only requires a single experiment (rather than pairs) which reduces the noise. The experiment has the additional benefit, by differencing with the standard amip run, of providing detailed information on the transient effective radiative forcing and adjustments in models relative to pre-industrial for the standard AMIP period. The inclusion of CFMIP process diagnostics not available in the RFMIP experiments will also enable a deeper understanding of the factors underlying forcing and feedback differences in the present and future climate.

2.5 Timeslice experiments for understanding regional climate responses to CO_2 forcing. Co-ordinators: Rob Chadwick (Met Office) and Hervé Douville (CNRM)

Science questions:

- How do regional climate responses (of e.g. precipitation) in a coupled model arise from the combination of responses to different aspects of CO_2 forcing and warming (uniform SST warming, pattern SST warming, direct CO_2 effect, plant physiological effect)?
- Which aspects of forcing/warming are most important for causing inter-model uncertainty in regional climate projections?
- Can inter-model differences in regional projections be related to underlying structural or resolution differences between models through improved process understanding, and could this help us to constrain the range of regional projections?
- What impact do coupled model SST biases have on regional climate projections?

We propose a set of 6 20-year atmosphere-only timeslice experiments to decompose the regional responses of each model's abrupt4xCO2 run into separate responses to each aspect of forcing and warming (uniform SST warming, pattern SST change, increased CO_2, plant physiological effect). As well as allowing regional responses in each individual model to be better understood, this set of experiments should prove especially useful for understanding the causes of model uncertainty in regional climate change.

The experiments are: 1) sstPi – the same as amip but with monthly-varying SSTs and sea-ice from years 101-120 of each model's own control run rather than observed fields; 2) sstPi4K – the same as sstPi but with SSTs uniformly increased by 4K; 3) sstPi4xCO2 – the same as sstPi but CO_2 as seen by the radiation scheme is quadrupled; 4) sstPi4xCO2Veg – the same as sstPi4xCO2 but with the plant physiological response also able to respond to the increased CO_2; 5) sstPiFuture – the same as sstPi but a seasonally varying monthly mean climatology of the SST pattern anomaly taken from years 91-140 of each model's own abrupt4xCO2 minus piControl is scaled to have a global mean increase of 4K and applied; 6) sstPiTot – the same as sstPiFuture but also with $4xCO_2$ including the plant effect. sstPiTot is used to establish whether a timeslice experiment can adequately recreate the coupled $abrupt4xCO_2$ response in each model, and then forms the basis for a decomposition using the other experiments.

We also propose an additional amip based experiment, amipTot: the same as sstPiTot but with the SST pattern anomaly climatology from sstPiFuture added instead to the observed background SSTs and sea-ice (as for other amip experiments). Comparison of amipTot and sstPiTot should help to illuminate the impact of SST biases on regional climate responses in each model, and how this contributes to inter-model uncertainty.

2.6 Atmosphere-only experiments for understanding the role of cloud-radiative effects in the large-scale atmospheric circulation in current and perturbed climates. Co-ordinators: Sandrine Bony (IPSL) and Bjorn Stevens (MPI).

Science questions:

- How do cloud-radiative effects impact the structure, the strength and the variability of the general atmospheric circulation in the present-day climate?
- How much do cloud-radiative feedbacks contribute to the spread of circulation and precipitation responses in climate change?
- Can we identify robust aspects of the climate response to global warming that do not depend on cloud-radiative feedbacks?

It is increasingly recognized that clouds, and cloud-radiative effects in particular, play a critical role in the general circulation of the atmosphere (ITCZ, MJO, storm tracks, hurricanes) and its response to global warming. A better assessment of this role would greatly help interpret model biases (how much do biases in cloud-radiative properties contribute to biases in the structure of the ITCZ, in the position and strength of the storm tracks, in the lack of intra-seasonal variability, etc) and to inter-model differences in simulations of the current climate and in climate change projections (especially changes in regional precipitation and extreme events). More generally, a better understanding of how clouds couple to circulation is expected to improve our ability to answer two of the four science questions raised by the WCRP Grand Challenge on Clouds, Circulation and Climate Sensitivity: what controls the position, the strength, and the variability of the storm tracks and of the tropical rainbelts?

These questions provided the scientific motivation for the Clouds On/Off Klima Intercomparison Experiment (COOKIE) project proposed by the european consortium EUCLIPSE and CFMIP in 2012. The COOKIE experiments, which have been run by 4 to 8 climate models (depending on the experiment), consisted in switching off the cloud-radiative effects (clouds seen by the radiation code - and the radiation code only- were artificially made transparent) in an atmospheric model forced by prescribed SSTs. By doing so, the atmospheric circulation could feel the lack of cloud-radiative heating within the atmosphere, but the land surface could also feel the lack of cloud shading, which led to changes in land-sea contrasts. The change in circulation between On and Off experiments was resulting from both effects, obscuring a bit the mechanisms through which the atmospheric cloud-radiative effects interact with the circulation for given surface boundary conditions. As the LW cloud-radiative effects are felt mostly within the troposphere (and represent most of the LW+SW cloud-radiative heating) while the SW effects are felt mostly at the surface, we could better isolate the role of tropospheric cloud-radiative effects on the circulation by running atmosphere-only experiments in which clouds are made transparent to radiation only in the LW.

We propose in Tier II a set of simple experiments similar to the amip, amip4K, aquaControl and aqua4K experiments of CMIP5/CFMIP2 (and Tier 1 of CMIP6) but in which cloud-radiative effects are switched off in the LW part of the radiation code. These experiments will be referred to as offlwamip, offlwamip4K, offlwaquaControl and offlwaqua4K. The analysis of idealized (aqua-planet) experiments will allow us to assess the robustness of the impacts found in more realistic (AMIP) configurations. It will also facilitate the interpretation of the results using simple dynamical models or theories, in collaboration with large-scale dynamicists (e.g. DynVar). The comparison of the inter-model spread of simulations between AMIP and offlwAMIP experiments for present-day and globally warmer climates will help identify which aspects of the spread depend on the representation of cloud-radiative effects, and which aspects do not, thus better highlighting other sources of spread.

Additional CFMIP experiments under consideration for the future

We also propose to use these CMIP6 experiments as the foundation for further experiments planned in the context of the Grand Challenge on Clouds, Circulation and Climate Sensitivity. These will include for example sensitivity experiments to assess the impacts of different physical processes on cloud feedbacks and regional circulation/precipitation responses, and others designed to test specifically proposed cloud feedback mechanisms. Additional experiments further idealizing the aquaplanet framework to a non-rotating rotationally symmetric case are also under development. These will be proposed as additional Tier II experiments at a future time, or coordinated by CFMIP outside of CMIP6.

5. Unless a Tier 1 experiment differs only slightly from another well-established experiment, it must already have been performed by more than one modeling group. All of the proposed Tier I experiments were previously included in CMIP5 and so are well established and already performed by multiple groups.

6. A sufficient number of modelling centers (~8) are committed to performing all of CFMIP's Tier 1 experiments and providing all the requested diagnostics needed to answer at least one of its science questions. Fourteen modeling groups have so far agreed to participate in CFMIP as part of CMIP6, implying that they are prepared to perform the Tier I experiments. These are ACCESS (Australia), BCC (China), CanESM (Canada), CESM (USA), CNRM (France), FGOALS (China), GFDL (USA), IPSL-ESM (France), MIROC6-GCM (Japan) NICAM (Japan), MPI-ESM (Germany), MRI (Japan) and UKESM (United Kingdom).

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. Our analysis plan is outlined below.

We commit to contributing to the creation of the CMIP6 data request and to analyzing the data, as we did for CMIP5. This will include making proposals for an updated COSP request in CMIP6 (see the proposal from the COSP PMC), and also additional improvements to the CFMIP diagnostic specifications relating to temperature and humidity budget increments, 3D radiative fluxes, inclusion of aerosol diagnostics across CFMIP experiments, and the introduction of additional locations in the cfSites specification.

We also commit to identifying observations needed for model evaluation and improved process understanding, and to contributing directly to making such datasets available as part of obs4MIPs. For example the CFMIP community has up to now played a central role in providing versions of CloudSat and CALIPSO datasets designed for direct comparison with CMIP5 data through the CFMIP-OBS website (see http://climserv.ipsl.polytechnique.fr/cfmip-obs/) and part of this work has recently involved publishing this data via the ESG and linking into obs4MIPS (see for example references to CFMIP-OBS on the obs4MIPS website at https://www.earthsystemcog.org/projects/obs4mips/aboutus). This work will continue.

CFMIP analysis activities are ongoing and the CFMIP community is ready to analyse CMIP6 data at any time. We would like modelling groups to perform the proposed CFMIP/CMIP6 experiments at the same time or shortly after their DECK experiments. Subsequent CFMIP experiments which are not included in CMIP6 will build on the proposed DECK and CMIP6/CFMIP experiments and some will start as soon as CMIP6 DECK experiments start to become available. We envisage a succession of CFMIP related intercomparisons addressing different questions arising from the GC spanning the duration of CMIP6.

We commit to scientifically analyze, evaluate and exploit the proposed experiments, and have identified leads within CFMIP for different aspects of this activity. An overview of the proposed evaluation/analysis of the CMIP DECK and CMIP6 experiments follows:

- CFMIP will continue to exploit the CMIP DECK and CMIP6 experiments to understand and evaluate cloud processes and cloud feedbacks in climate models. The wide range of analysis activities described above in the context of CFMIP-2 will be continued in CFMIP-3 using the CMIP DECK and CMIP6 experiments, allowing the techniques developed in CFMIP-2 to applied to an expanding number of models, including the new generation of models currently under development. These activities will include evaluation of clouds using additional simulators (see proposal regarding COSP below), investigation of cloud processes and cloud feedback/adjustment mechanisms using process outputs (cfSites, budget tendency terms, etc). The inclusion of COSP and budget tendency terms in additional DECK experiments (e.g. abupt4xCO2 and some scenario experiments, also see proposal for COSP below) will enable the CFMIP approach to be applied to a wider range of experimental configurations. (Lead coordinator Mark Webb).

- Analysis of the +/-4% solar model runs would include an evaluation of both rapid adjustments and longer-term responses on global and regional top-of-atmosphere radiative fluxes, cloud types (using ISCCP and other COSP simulators) and precipitation characteristics, as well as comparison of these responses with responses in DECK abrupt4xCO2 experiments. GeoMIP and SolarMIP have expressed a strong interest in these CFMIP experiments and joint analysis of these CFMIP experiments with GeoMIP and SolarMIP experiments is anticipated, specifically with the goal of determining to what degree results from abrupt solar forcing ONLY experiments and abrupt CO2 ONLY experiments can be used to predict what happens when both forcing are applied simultaneously, as done in the GeoMIP experiments (Lead coordinator Chris Bretherton).
- Analysis of nonlinear climate processes will primarily involve comparing the abrupt4xCO2, abrupt2xCO2 and abrupt0.5xCO2 experiments over the same timescale (Good et al., 2014). (Lead coordinator Peter Good).
- Analysis of amipPiForcing has already been done in detail for a single model in Andrews (2014). We propose to use this has a starting point for a multi-model analysis. (Lead coordinator Timothy Andrews).
- An overview analysis of regional responses and model uncertainty in the timeslice and amipTot experiments will be carried out by the co-ordinators, in collaboration with members of contributing modeling groups. We anticipate that further detailed analysis on the processes at work in different regions will be carried out by a variety of research groups with interest and expertise in a particular region: for example a set of similar experiments has previously been used to examine the climate response of the West African monsoon in CCSM3 (Skinner et al. 2012). The timeslice and amipTot experiments have already been successfully run with HadGEM2 (Met Office), and are currently in the planning stage for CNRM. (Lead coordinator Robin Chadwick).
- When analyzed together with the amip4K experiment, the amipMinus4K experiment allows one to exploit the CFMIP process diagnostics to understand for asymmetries in the climate response to warming and cooling which have been noted in PMIP experiments. These might arise from cloud phase responses in middle- and high-latitude clouds or from the adiabatic cloud liquid water path response feedback which is important over land regions and which would be expected to be weaker with cooling because of the non-linearity in the Clausius-Clapeyron relation. (Lead coordinator Mark Webb).

8. The MIP has completed the MIP template questionnaire. We have done this.

9. The MIP contributes a paper on its experimental design to the CMIP6 Special Issue. We agree to do this.

10. The MIP considers reporting on the results by co-authoring a paper with the modelling groups. We agree to do this. Separate papers will be prepared for each of the experiment groups proposed.

Answers to other questions in the MIP template questionnaire

All model output archived by CMIP6-Endorsed MIPs is expected to be made available under the same terms as CMIP output. Most modeling groups currently release their CMIP data for unrestricted use. If you object to open access to the output from your experiments, please explain the rationale. We have no objection to this.

List of output and process diagnostics for the CMIP DECK/CMIP6 data request. Please see the accompanying spreadsheet and outline below.

Any proposed contributions and recommendations for model diagnostics and performance metrics, observations/reanalysis data products, tools, code or scripts. We have provided a database of performance metrics and codes at the CFMIP Diagnostics Code Repository and a set of observational data for comparison with CFMIP outputs at the CFMIP-OBS site. Both are accessible via the CFMIP website http://www.cfmip.net. We welcome additional contributions to both of these databases.

Any proposed changes from CMIP5 in NetCDF metadata (controlled vocabularies), file names, and data archive (ESGF) search terms. None expected.

Explanation of any proposed changes (relative to CMIP5) that will be required in CF, CMOR, and/or ESGF. None expected.

CFMIP Recommended Outputs For CMIP6 DECK experiments and CFMIP experiments.

CFMIP recommends a set of diagnostic outputs for the CMIP6 DECK and CFMIP experiments which are based on those from CFMIP-2, with some modifications. These are detailed in the accompanying spreadsheet CMIP6DataRequestCompilationCFMIP_20150331.xls, and are summarized below. The recommendations are in two parts. The first part describes updates to the CFMIP process diagnostics compared to those which were requested in CMIP5, in terms of additional variables and the experiments in which they are requested. This set was drawn up by the CFMIP committee and ratified by the modeling groups following a presentation at the 2014 CFMIP meeting. The second part describes recommendations for COSP outputs in the DECK and CMIP6 Historical experiments which were drawn up by the COSP Project Management Committee (PMC). Please refer to the request scoping worksheet in the accompanying spreadsheet for a summary of which outputs are requested in which DECK experiments CFMIP experiments.

For participation in CFMIP it is required that modeling groups commit to performing all of the Tier I experiments, and sufficient diagnostic outputs to answer at least one scientific question. Since a number of the science questions of CFMIP (e.g. those pertaining to precipitation responses) require no diagnostic outputs beyond the standard 'Amon' outputs from CMIP5, a modeling group may qualify for participation in CFMIP even if they run the Tier I experiments without CFMIP simulators or process outputs. Such a submission would be useful, in the main for the precipitation analysis aspects of CFMIP. However we strongly recommend that participating groups additionally submit as many of the COSP and process outputs as they are practically able to, to support investigations of the full range of scientific questions of CFMIP in CMIP6.

Proposed updates to CFMIP process outputs for the CMIP DECK, CMIP6 Historical and CMIP6 CFMIP experiments.

CFMIP Committee

March, 2015

The diagnostic request for CMIP5/CFMIP2 is summarised and motivated in the CFMIP-2 proposal document [Bony et al., 2009], and documented in detail in the CMIP5 Standard Output documentation at http://cmip-pcmdi.llnl.gov/cmip5/output_req.html in excel spreadsheet format (Worksheet 'CFMIP output' indicates which tables appear in which experiments and for which periods, which other worksheets such as cfMon, cfDay etc indicate the variables in each table). Our view is that the CFMIP-2 diagnostics set is fundamentally sound and forms a suitable basis for the process diagnostics in the DECK, CMIP6 Historical and CMIP6 CFMIP experiments. Thus, we present this proposal as changes with respect to the CMIP5/CFMIP-2 protocol in the accompanying spreadsheet, which includes a request scoping worksheet indicating which outputs are requested in which experiments, including the CMIP6 DECK + CMIP6 Historical experiments and the CFMIP experiments proposed within CMIP6. In the sections below we present and motivate the specific requested changes.

In this section we cite a number of peer reviewed publications. Please refer to http://www.cfmip.net -> CFMIP Publications for full references.

cfSites Outputs: The CFMIP cfSites outputs were requested in CMIP5 for 120 locations in the amip, amip4K, amipFuture and amip4xCO2 experiments, and for 73 locations along the Greenwich meridian in the aquaplanet experiments. These outputs have so far been used to evaluate the models with in-situ measurements (e.g. Nuijens et al. (submitted), Guichard et al. (in prep), Neggers et al. (submitted) and to examine cloud feedbacks on short timescales such as over the diurnal cycle (Webb et al. 2015). For CMIP6 we have dispensed with the cfSites outputs in the aquaplanet experiments, and in amipFuture, retaining them in amip, amip4K and amip4xCO2 only. At the request of the US CLIVAR ETOS WG we have added Ascension Island and St. Helena to the list in light of upcoming field work/additional radiosondes from these islands, increasing the total number of locations to 122.

Temperature and humidity tendency terms: CFMIP-2 requested cloud, temperature and humidity tendency terms. In CMIP6 we have omitted the cloud condensate tendency terms because no publications have arisen from those saved in CMIP5. The temperature and humidity tendency terms from CMIP5 have been widely used however. Temperature and humidity tendency terms have been demonstrably useful for understanding the roles of different parts of the model physics in cloud feedbacks, adjustments, and present-day variability (Williams et al 2013, Webb and Lock 2013, Kamae and Watanabe 2012, Demoto et al 2013, Sherwood et al 2014, Ogura et al 2014, Brient et al. (submitted), Xavier et al. (submitted)). They have also been used to understand regional warming patterns such as polar amplification in coupled models (e.g. Yoshimori et al 2013,2014). For CMIP6 we have improved the definitions of the temperature and humidity tendency terms, and added some additional terms such as clear-sky radiative heating rates to more precisely quantify the contributions of different processes to the temperature and humidity budget changes underlying cloud feedbacks and adjustments. A shortcoming of the CMIP5 protocol was that we were unable to interpret the physical feedback mechanisms in coupled model experiments due to lack of process diagnostics. For this reason we are additionally requesting these budget terms in the DECK abrupt4xCO2 experiment and the pre-industrial control.

Additional daily diagnostics: So-called ‘clustering’ approaches are now commonly used for assessing the contributions of different cloud regimes (e.g. stratocumulus, trade cumulus, frontal clouds, etc) to present day biases in cloud simulations and to inter-model differences in cloud feedbacks (e.g. Williams and Webb 2009, Tsushima et al., 2013, Tsushima et al., submitted). We have added some additional daily 2D fields to the standard package of CFMIP daily outputs to allow further investigation of feedbacks between clouds and aerosols associated with the changing hydrological cycle (aerosol loadings and cloud top effective radii/number concentrations) and a clearer diagnosis of the roles of convective and stratiform clouds (convective vs stratiform ice and condensed water paths and cloud top effective radii/number concentrations).

Proposal of request of COSP diagnostics for CMIP/DECK, CMIP6 Historical and CMIP6 CFMIP experiments

COSP Project Management Committee

March, 2015

1 Introduction

The initial design for CMIP6 has recently been published [Meehl et al., 2014]. It includes a set of 'DECK + CMIP6 Historical' experiments to be run by modelling groups whenever they develop a new model version:

- AMIP (1979-end)
- Pre-industrial control
- 1% yr^{-1} CO_2 increase up to $4xCO_2$
- Abrupt $4xCO_2$
- CMIP6 Historical run

These experiments are called the CMIP Diagnostic, Evaluation and Characterization of Klima (DECK) plus CMIP6 Historical experiments. In this document, we present the proposal of the list of COSP diagnostics to be requested to for the DECK + CMIP6 Historical experiments and additionally the CMIP6 CFMIP experiments. This proposal is the outcome of discussions by the COSP Project Management Committee (PMC) and the CFMIP Committee. The COSP diagnostic request for CMIP5/CFMIP2 is summarised and motivated in the CFMIP-2 proposal document [Bony et al., 2009], and documented in detail in the CMIP5 Standard Output documentation at http://cmip-pcmdi.llnl.gov/cmip5/output_req.html in excel spreadsheet format (Worksheet 'CFMIP output' indicates which tables appear in which experiments and for which periods, which other worksheets such as cfMon, cfDay etc indicate the variables in each table). Our view is that the CFMIP-2 diagnostics set fundamentally sound and forms a suitable basis for the COSP request for the DECK, CMIP6 Historical and CMIP6 CFMIP experiments, subject to some modifications. Thus, we present this proposal as changes with respect to the CMIP5/CFMIP-2 protocol in the accompanying spreadsheet. The request scoping sheet also shows which outputs are requested in which experiments. We have tried to address the concerns raised in the CMIP5 survey by simplifying the technical difficulty of the requests (sometimes at the expense of extra data) and basing the requests upon a frozen well-tested and already-released version of COSP (v1.4). In the sections below we present and motivate the specific requested changes.

2 Description of proposed changes

2.1 Change #1: Replacement of curtain data by full 3D fields, and deletion of cfOff table (proposed by Alejandro Bodas-Salcedo)

In CFMIP-2, the production of data along the A-train track ("curtain" data, table cf3hr offline) involved a substantial amount of post-processing. A second post-processing step required the gridding and time-averaging of these data to produce the monthly means requested in the cfOff table. This proved quite difficult for many modelling centres. Although not from the ESG archive, this type of data has been used in several model evaluation papers [Bony et al., 2009; Bodas-Salcedo et al., 2008; Field et al., 2011; Williams et al., 2013] involving case-study comparison of models with along-track observations from CloudSat and Calipso. We believe that by simplifying the request, the modelling centres will find easier to contribute these data. Hence, we propose to drop the orbital sampling, i.e. to request globally-complete fields on a standard lat/lon grid. Given this change, the calculation of monthly-averages from gridded 3-hourly data is trivial, and therefore we propose to delete the cfOff table.

2.2 Change #2: New table cfMonExtra. Add CloudSat and CALIPSO CFADs to cfMonExtra (proposed by Alejandro Bodas-Salcedo and Mark Webb)

Optimisations to the code in COSP v1.4 mean that it is now practical to run the CloudSat simulator inline in models and so for longer periods. We propose the introduction of a new table cfMonExtra for the inclusion of monthly mean COSP diagnostics used for model evaluation in the AMIP DECK experiment, but which we don't consider appropriate for coupled or climate change experiments. In this new table we include Cloud Frequency/Altitude Diagram (CFAD) diagnostics for CloudSat and CALIPSO for the entire AMIP integration. CFADs for CloudSat and CALIPSO have appeared in a number of published studies [e.g. Nam et al., 2014; Franklin et al., 2013; Bodas-Salcedo et al., 2011; Bodas-Salcedo et al., 2012; Nam and Quaas, 2012; Nam and Quaas, 2013; Kay et al., 2012; Kodama et al., 2012; Marchand et al., 2009; Abel and Boutle, 2012] and their inclusion as monthly means in the AMIP DECK experiment will make them available for analysts in a more convenient form than the higher frequency outputs currently requested in CMIP5.

2.3 Change #3: Standard monthly COSP and daily COSP 2D outputs in all of the DECK, CMIP6 Historical and CMIP6 CFMIP Experiments (proposed by Mark Webb and Steve Klein)

Many of the standard monthly COSP and daily COSP 2D have been shown to be valuable in the CMIP5 experiments, not only for cloud evaluation [e.g. Franklin et al., 2013; Bodas-Salcedo et al., 2012; Nam and Quaas, 2013; Lacagnina and Selten, 2014; Bodas-Salcedo et al., 2014; Klein et al., 2013; Cesana and Chepfer, 2012; Tsushima et al., 2013] but also in quantifying the contributions of different cloud types to cloud feedbacks and forcing adjustments in climate change experiments [e.g. Tsushima et al., 2013; Zelinka et al., 2012a; Zelinka et al., 2012a; Zelinka et al., 2013; Zelinka et al., 2014]. We propose to include these in all of the DECK, CMIP6 Historical and CMIP6 CFMIP experiments as standard for the entire length of the runs, to support evaluation of cloud, cloud feedbacks and cloud adjustments and to investigate trends in the observational record.

2.4 Change #4: Move PARASOL reflectance to cfMonExtra (proposed by Robert Pincus)

Top-of-atmosphere reflectance measurements from PARASOL were part of the standard request for CMIP5. They have been used in some applications [e.g. Nam et al. 2012] but have not been widely exploited. The proposal is to move them from the cfMon to cfMonExtra tables to reduce the number of integrations for which they are requested and to focus on model evaluation applications.

2.5 Change #5: Add MISR CTH-COD to cfMonExtra. Add MISR CTH-COD and ISCCP CTP-OD histograms to cf3hr (proposed by Roger Marchand)

Histograms of cloud-top-height (or cloud-top-pressure) and optical-depth produced by ISCCP have been widely used in the evaluation of climate models, often in combination with the ISCCP-simulator now part of COSP. Because top-of-atmosphere outgoing longwave fluxes are related to cloud-top-height and outgoing shortwave fluxes are related to cloud-optical-depth this framework provides a way to evaluate the distribution of model clouds in a way that is closely related to their radiative impact. Similar histograms of cloud-top-height and optical-depth are being produced from observations by the Multiangle Imaging Spectro-Radiometer (MISR). While similar, the cloud-top-height in the MISR dataset is obtained using a stereo-imaging technique that his purely geometric and insensitivity to the calibration of the MISR cameras. This technique provides more accurate retrievals of cloud-top-height for low-level and mid-level clouds, and more reliable discrimination of mid-level clouds from other clouds, while ISCCP provides greater sensitivity to optically-thin high-level clouds. In addition, ISCCP and MISR histograms can be combined to separate optically-thin high-level clouds into multi-layer and single-layer categories [Marchand et al. 2010]. We therefore recommend using both ISCCP and MISR observations and instrument-simulators in the evaluation of climate model, and such an analysis is underway using a few CFMIP5 models that have run the MISR simulator [Hillman et al. 2014]. While monthly data are useful for the broad evaluation of models on monthly or longer time scales, the acquisition of high frequency (Three hourly) data will enable analysis of events that are not well resolved with monthly data, including the diurnal cycle, the Madden-Julian Oscillation (MJO) and various synoptic states or weather patterns, such as frontal passages. We recognize that this represents a large increase in data-volume compared with monthly averages and propose collection of this three hourly data only for a period of about 1 year.

2.6 Change #6: Add MODIS cloud fractions (total, liquid, ice) to cfMonExtra (proposed by Robert Pincus)

The partitioning between liquid and ice phase has significant impacts on the energy and hydrologic impacts of clouds. As models move towards predicting more details of the aerosol distributions, including the ice nucleation ability, evaluation of the phase partitioning on the global scale will become more important. Evaluation to date has been based primarily on polarization measurements from active and passive sensors [e.g. Doutriaux-Boucher and Quaas, 2004; Komurcu et al., 2014] and height-resolved partitioning estimates from the CALIPSO sensor are requested below. Cloud phase estimates from the MODIS simulator were not available in CFMIP2 but may prove a useful complement by virtue of greater geographic sampling and longer time records.

2.7 Change #7: MODIS COT-particle size histograms by phase in cfMonExtra, cfDayExtra, cf3hr (proposed by Robert Pincus)

The joint distribution of optical thickness and particle size provides a window on the microphysical processes within clouds [Nakajima et al., 1991] and is influenced by direct and some indirect effects of aerosols on cloud optical properties [Han et al. 2002]. As models move towards predicting more details of the aerosol properties and cloud-aerosol interactions the assessment of these processes becomes more pressing.

Estimate of particle size from MODIS have been difficult to use for model evaluation to date because of observational artefacts not treated by the MODIS simulator. These artefacts are reduced by the use of observations at wavelengths with greater absorption by condensed water (e.g. by exploiting reflectance at 3.7 μm instead of 2.1 μm). The MODIS simulator and accompanying data for CFMIP3 will use measurements at 3.7 μm to infer particle size. This will also act to make output from the MODIS simulator roughly consistent with the PATMOS-X observations in the same way that distributions of optical thickness from the MODIS, MISR, and ISCCP simulators are nearly equivalent.

2.8 Change #8: add CALIPSO ice and liquid 3D cloud fractions to cfMonExtra (proposed by Hélène Chepfer)

Changes in cloud optical depth associated with cloud phase feedbacks can dominate the changes in high-latitude clouds in future climate projections [e.g. Senior and Mitchell, 1993]. Cloud phase identification capabilities have been recently added to the CALIPSO simulator in COSP, and a compatible observational dataset has been produced [Cesana and Chepfer, 2013]. We propose to include these in the AMIP DECK experiment to support the evaluation of the simulation of cloud phase.

2.9 Change #9: CALIPSO total cloud fraction and PARASOL reflectance to cfDayExtra (proposed by Hélène Chepfer and Dimitra Konsta)

The multi-sensor A-train observations (CALIPSO-GOCCP and MODIS, PARASOL) allow to make the correlations between the different cloud variables at the instantaneous time scale, and at high resolution. The use of the high-frequency relationships between different variables allows for process-oriented model evaluation. These diagnostics will help test the realism of the co-variation of key cloud properties that control cloud feedbacks in models. Konsta at al. (2014) have used these diagnostics in a pilot analysis.

3 References Using Satellite Simulators for the Evaluation of Clouds in Models (partial)

Bodas-Salcedo, A. et al., 2008: Evaluating cloud systems in the Met Office global forecast model using simulated CloudSat radar reflectivities, *J. Geophys. Res.*, 113, D00A13. DOI: 10.1029/2007JD009620.

Bodas-Salcedo, A., et al., 2011: Satellite simulation software for model assessment. *Bull. Am. Meteorol. Soc.*, 92. DOI: 10.1175/2011BAMS2856.1.

Bodas-Salcedo, A., et al., 2012: The surface downwelling solar radiation surplus over the Southern Ocean in the Met Office model: the role of midlatitude cyclone clouds, *J. Climate*, 25. DOI: 10.1175/JCLI-D-11-00702.1.

Bodas-Salcedo, A., et al., 2014: Origins of the Solar Radiation Biases over the Southern Ocean in CFMIP2 Models, J. Climate, 27. DOI: 10.1175/JCLI-D-13-00169.1.

Cesana, G., and Chepfer, H., 2012: How well do climate models simulate cloud vertical structure? A comparison between CALIPSO-GOCCP satellite observations and CMIP5 models, *Geophys. Res. Let.* DOI: 10.1029/2012GL053153.

Cesana, G., and Chepfer, H., 2013: Evaluation of the cloud thermodynamic phase in a climate model using CALIPSO-GOCCP, *J. Geophys. Res.*, 118, 7922–7937. DOI: 10.1002/jgrd.50376.

Doutriaux-Boucher, M., and J. Quaas, 2004: Evaluation of cloud thermodynamic phase parametrizations in the LMDZ GCM by using POLDER satellite data, *Geophys. Res. Lett.*, 31, L06126. DOI: 10.1029/2003GL019095.

Field, P. R., et al., 2011: Using model analysis and satellite data to assess cloud and precipitation in midlatitude cyclones, *Q. J. R. Meteorol. Soc.*, 137, 1501-1515. DOI: 10.1002/qj.858.

Franklin, C. N., et al., 2013: Evaluation of clouds in ACCESS using the satellite simulator package COSP: Global, seasonal, and regional cloud properties, *J. Geophys. Res.*. DOI: 10.1029/2012JD018469.

Hillman, B., R. Marchand, T. P. Ackerman, A.Bodas-Salcedo, J. Cole, J.-C. Golaz, J. E. Kay, 2014: Comparing Cloud Biases in CMIP5: Insights Using MISR and ISCCP Observations and Satellite Simulators, *in preparation.*

Kay, J. E., et al., Exposing Global Cloud Biases in the Community Atmosphere Model (CAM) Using Satellite Observations and Their Corresponding Instrument Simulators, *J. Climate*, 25, 2012. DOI: 10.1175/JCLI-D-11-00469.1.

Klein, S. A. et al., 2013: Are climate model simulations of clouds improving? An evaluation using the ISCCP simulator, *J. Geophys. Res.*, 118. DOI: 10.1002/jgrd.50141.

Kodama, C., et al., 2012: An assessment of the cloud signals simulated by NICAM using ISCCP, CALIPSO, and CloudSat satellite simulators, *J. Geophys. Res,* 117. DOI: 10.1029/2011JD017317.

Komurcu, M., T. Storelvmo, I. Tan, U. Lohmann, Y. Yun, J. E. Penner, Y. Wang, X. Liu, and T. Takemura ,2014: Intercomparison of the cloud water phase among global climate models, J. Geophys. Res., 119, 3372–3400. DOI:10.1002/2013JD021119.

Konsta, D., J-L. Dufresne, H. Chepfer, A. Idelkadi and G. Cesana, 2014: Evaluation of clouds simulated by the LMDZ5 GCM using A-train satellite observations (CALIPSO, PARASOL, CERES). *Climate Dynamics*, under review.

Lacagnina, C., and Selten, F., 2014: Evaluation of clouds and radiative fluxes in the EC-Earth general circulation model, *Clim. Dyn.* DOI: 10.1007/s00382-014-2093-9.

Marchand, R., et al., 2009: A comparison of simulated cloud radar output from the multiscale modeling framework global climate model with CloudSat cloud radar observations, *J. Geophys. Res., 114,* D00A20. DOI: 10.1029/2008JD009790.

Marchand, R., T. Ackerman, M. Smyth, and W. B. Rossow ,2010: A review of cloud top height and optical depth histograms from MISR, ISCCP, and MODIS, *J. Geophys. Res.*, 115, D16206. DOI:10.1029/2009JD013422.

Nam, C., S. Bony, J.-L. Dufresne, and H. Chepfer, 2012: The "too few, too bright" tropical low-cloud problem in CMIP5 models, *Geophys. Res. Lett.*, 39. DOI:10.1029/2012GL053421.

Nam, C. C. W., and Quaas, J., 2012: Evaluation of Clouds and Precipitation in the ECHAM5 General Circulation Model Using CALIPSO and CloudSat Satellite Data. I, *J. Climate*, 25, 4975-4992. DOI:10.1175/JCLI-D-11-00347.1.

Nam, C. C. W. and Quaas, J., 2013: Geographically versus dynamically defined boundary layer cloud regimes and their use to evaluate general circulation model cloud parameterizations, *Geophys. Res. Let.* DOI: 10.1002/grl.50945.

Nam, C. W. W., et al., 2014: Evaluation of boundary layer cloud parametrizations in the ECHAM5 general circulation model using CALIPSO and CloudSat satellite data. *JAMES*. DOI: 10.1002/2013MS000277.

Tsushima, Y. et al., 2013: Quantitative evaluation of the seasonal variations in climate model cloud regimes, *Clim. Dyn.*, 41. DOI: 10.1007/s00382-012-1609-4.

Williams, K. D., et al., 2013: The Transpose-AMIP II experiment and its application to the understanding of Southern Ocean cloud biases in climate models, *J. Climate*, 26, 3258-3274. DOI: 10.1175/JCLI-D-12-00429.1.

Zelinka, M. D., et al., 2012a: Computing and Partitioning Cloud Feedbacks Using Cloud Property Histograms. Part I: Cloud Radiative Kernels. *J. Climate*. DOI: 10.1175/JCLI-D-11-00248.1.

Zelinka, M. D., et al., 2012b: Computing and Partitioning Cloud Feedbacks Using Cloud Property Histograms. Part II: Attribution to Changes in Cloud Amount, Altitude, and Optical Depth. *J. Climate*. DOI: 10.1175/JCLI-D-11-00249.

Zelinka, M. D., et al., 2013: Contributions of Different Cloud Types to Feedbacks and Rapid Adjustments in CMIP5, *J. Climate*. DOI: 10.1175/JCLI-D-12-00555.1.

Zelinka, M. D., et al., 2014: Quantifying Components of Aerosol-Cloud-Radiation Interactions in Climate Models. *J. Geophys. Res*. DOI: 10.1002/2014JD021710.

4 Other References

Abel, S. J., and Boutle, I. A., 2012: An improved representation of the raindrop size distribution for single-moment microphysics schemes, *Q. J. R. Meteorol. Soc.*, 138, 2151-2162. DOI: 10.1002/qj.1949.

Bony, S. et al., 2009: The Cloud Feedback Model Intercomparison Project : Summary of Activities and Recommendations for Advancing Assessments of Cloud-Climate Feedbacks, (http://www.cfmip.net -> CFMIP Strategy and Plans -> CFMIP2_experiments_March20th2009.pdf).

Han, Q, W B. Rossow, J Zeng, R Welch, 2002: Three Different Behaviors of Liquid Water Path of Water Clouds in Aerosol–Cloud Interactions. *J. Atmos. Sci.*, **59**, 726–735. DOI: 10.1175/1520-0469(2002)059<0726:TDBOLW>2.0.CO;2.

Meehl, G. et al., 2014: Climate Model Intercomparisons: Preparing for the Next Phase, *Eos Trans. AGU*, 95(9), 77. DOI: 10.1002/2014EO090001.

Nakajima, T, M D. King, J D. Spinhirne, L F. Radke, 1991: Determination of the Optical Thickness and Effective Particle Radius of Clouds from Reflected Solar Radiation Measurements. Part II: Marine Stratocumulus Observations, *J. Atmos. Sci.*, **48**, 728–751. DOI: 10.1175/1520-0469(1991)048<0728:DOTOTA>2.0.CO;2.

Senior, C. A., and Mitchell, J. F. B., 1993: Carbon Dioxide and Climate. The Impact of Cloud Parameterization. *J. Climate,* 6, 393–418. DOI: 10.1175/1520-0442(1993)006<0393:CDACTI>2.0.CO;2.

Acknowledgements

Thanks to D. Konsta for providing text for Section 2.9.

The nonlinMIP intercomparison project: physical basis, experimental design and analysis principles

P. Good[1], T. Andrews[1], N. Bouttes[2], R. Chadwick[1], J. M. Gregory[2,1], J. A. Lowe[1]

[1]{ Met Office Hadley Centre, Exeter, United Kingdom}
[2]{ NCAS-Climate, University of Reading, Reading, United Kingdom}
Correspondence to: P. Good (peter.good@metoffice.gov.uk)

Abstract

nonlinMIP aims to quantify and understand, at regional scales, climate responses that are non-linear under CO2 forcing (mechanisms for which doubling the CO2 forcing does not double the response). Non-linear responses can be large at regional scales, with important implications for understanding mechanisms and for GCM emulation techniques (e.g. energy balance models and pattern-scaling methods). However, these processes are hard to explore using traditional experiments, explaining why they have had little attention in previous studies. Some single model studies have established novel analysis principles and some physical mechanisms. There is now a need to explore robustness and uncertainty in such mechanisms across a range of models.

nonlinMIP addresses this using a simple, small set of CO2-forced experiments that are able to separate linear and non-linear mechanisms cleanly, with good signal/noise – while being demonstrably traceable to realistic transient scenarios. The design builds on the CMIP5 and CMIP6 DECK protocols, and is centred around a suite of abruptCO2 experiments, with a ramp-up-ramp-down experiment to test traceability to gradual forcing scenarios. The understanding gained will help interpret the spread in policy-relevant scenario projections.

Here we outline the basic physical principles behind nonlinMIP, and the method of establishing traceability from abruptCO2 to gradual forcing experiments, before detailing the experimental design and finally some analysis principles. The discussion on traceability of abruptCO2 to transient experiments is also relevant to the abrupt4xCO2 experiment in the CMIP5 and CMIP6 DECK protocols.

5 Introduction

Climate impacts assessments require, at regional scales, understanding of physical mechanisms of climate change in GCM projections. Also required is the ability to emulate (using fast simplified climate models) GCM behaviour for a much larger range of policy-relevant scenarios than may be evaluated using GCMs directly. These two requirements may be combined into a single question: what is the simplest conceptual framework that has quantitative predictive power and captures the key mechanisms behind GCM scenario projections?

Often, a pragmatic choice has been to assume some form of linearity. In studies of the global energy balance, linearity is often assumed in the form of a constant climate feedback parameter. This parameter may be used to quantify feedbacks in different models (e.g. Zelinka et al., 2013) or, in emulation methods, to parameterise global energy balance models (e.g. Huntingford and Cox, 2000). In understanding or emulating regional patterns of climate change, it is often assumed that regional climate change is roughly proportional to global mean warming. In emulation work, this is termed 'pattern scaling' (Mitchell, 2003;Santer et al., 1990;Tebaldi and Arblaster, 2014), but this assumption may also be applied either explicitly or implicitly in understanding mechanisms. Sometimes, patterns of change per K of global warming are quantified; often, physical mechanisms are studied for a single period of a single forcing scenario (implicitly assuming that the understanding is relevant for other periods or scenarios).

While these approximations appear to work well under some circumstances, significant limitations are increasingly being revealed in such assumptions. These are of two types: different timescales of response, and non-linear responses. In discussing this, a complication arises in that different linearity assumptions exist. Henceforth we define 'linear' as meaning 'consistent with linear systems theory' - i.e. responses that are linear in model forcing (i.e. where doubling the forcing doubles the response; this is different from assuming that pairs of responses are linearly related to each other – as in pattern scaling).

Even in a linear system (where responses are linear in forcing), the relationship between two system outputs (e.g. between global-mean temperature and regional sea surface temperature - SST) will in general be non-linear. This is due to different timescales of response in different locations and/or variables. Examples include lagged surface ocean warming due to a connection with the deeper ocean (Chadwick et al., 2013;Held et al., 2010;Williams et al., 2008;Manabe et al., 1990;Andrews and Ringer, 2014) or the direct response of precipitation to forcings (Andrews et al., 2010;Allen and Ingram, 2002;Mitchell et al., 1987). One (generally false) assumption of pattern scaling, then, is that regional climate responds over the same timescale as global-mean temperature. Different timescales of response are especially important in understanding and predicting behaviour under mitigation and geoengineering scenarios (or over very long timescales).

Non-linear system responses (e.g. Schaller et al., 2013) are more complex to quantify, understand and predict than those of linear systems. Some examples have been known for some time, such as changing feedbacks through retreating snow/sea-ice (Colman and McAvaney, 2009;Jonko et al., 2013), or the Atlantic Meridional Overturning Circulation. More recently, substantial non-linear precipitation responses have been demonstrated in spatial patterns of regional precipitation change in two Hadley Centre climate models with different atmospheric formulations (Good et al., 2012;Chadwick and Good, 2013). This is largely due to simultaneous changes in pairs of known robust pseudo-linear mechanisms (Chadwick and Good, 2013). Non-linearity has also been demonstrated in the response under idealised geoengineering scenarios, of ocean heat uptake, sea-level rise, and regional climate pattterns, with different behaviour found when forcings are decreasing than when they are increasing (Bouttes et al., 2013;Schaller et al., 2014).

Investigation of these mechanisms at regional scales has been constrained by the type of GCM experiment typically analysed. Most previous analyses (e.g. Solomon et al., 2007) have used results from transient forcing experiments, where forcing changes steadily through the experiment. There are three main problems with this approach. First, information about different timescales of response is masked. This is because the GCM response at any given time in a transient forcing experiment is a mixture of different timescales of response (Good et al., 2013;Held et al., 2010;Li and Jarvis, 2009), including short-timescale responses (e.g. ocean mixed layer response from forcing change over the previous few years) through long-timescale behaviour (including deeper ocean responses from forcing changes multiple decades to centuries earlier). Secondly, in transient forcing experiments, non-linear behaviour is hard to separate from linear mechanisms. For example, in an experiment where CO2 is increased by 1% per year for 140 years ('1pctCO2'), we might find different spatial patterns at year 70 (at 2xCO2) than at year 140 (at 4xCO2). This could be due to nonlinear mechanisms (due to the different forcing level and associated different climate state). However, it could also be due to linear mechanisms: year 140 follows 140 years of forcing increase, so includes responses over longer response timescales than at year 70 (only 70 years of forcing increase). Thirdly, signal/noise ratios of regional climate change can be relatively poor in such experiments.

These three issues may be addressed by the use of idealised abruptCO2 GCM experiments (Forster et al., 2012;Zelinka et al., 2013;Jonko et al., 2013;Good et al., 2013;Good et al., 2012;Chadwick and Good, 2013;Chadwick et al., 2013;Bouttes et al., 2013;Gregory et al., 2004): an experiment where CO2 forcing is abruptly changed, then held constant. In abrupt CO2 experiments, responses over different timescales are separated from each other. Further, responses at different forcing levels may be directly compared, e.g. by comparing the response in abrupt2xCO2 and abrupt4xCO2 experiments over the same timescale - both have identical forcing time histories, apart from the larger forcing magnitude in abrupt4xCO2. Thirdly, high signal/noise is possible: averages may be taken over periods of 100 years or more (after the initial ocean mixed layer adjustment, change is gradual in such experiments). Recent work (Good et al., 2011;Good et al., 2012;Good et al., 2013;Zelinka et al., 2013) has established that these experiments contain global and regional-scale information quantitatively traceable to more policy-relevant transient experiments - and equivalently, that they form the basis for fast simple climate model projections traceable to the GCMs.

The CMIP5 abrupt4xCO2 experiments have thus been used widely: including quantifying GCM forcing and feedback behaviour (Gregory et al., 2004;Zelinka et al., 2013), and for traceable emulation of GCM projections of global-mean temperature and heat uptake (Good et al., 2013;Stott et al., 2013). Abrupt4xCO2 is also part of the CMIP6 DECK protocol (Meehl et al., 2014).

NonlinMIP extends the CMIP5 and CMIP6 DECK designs to explore non-linear responses (via additional abruptCO2 experiments at different forcing levels. It also explores responses over slightly longer timescales (extending the CMIP5 abrupt4xCO2 experiment by 100 years).

6 Relating abruptCO2 to gradual forcing scenarios: the step-response model

In using the highly-idealised abruptCO2 experiments, it is essential that their physical relevance (traceability) to more realistic gradual forcing experiments is determined. Some GCMs could respond unrealistically to the abrupt forcing change. A key tool here is the step-response model (described below). This response-function method aims to predict the GCM response to any given transient-forcing experiment, using the GCM response to an abruptCO2 experiment. Such a prediction may be compared with the GCM transient-forcing simulation, as part of a traceability assessment (discussed in detail in section 5).

Once some confidence is established in traceability of the abruptCO2 experiments to transient-forcing scenarios, the step-response model has other roles: to explore the implications, for different forcing scenarios, of physical understanding gleaned from abruptCO2 experiments; to help separate linear and nonlinear mechanisms (section 5); and potentially as a basis for GCM emulation. The method description below also serves to illustrate the assumptions of linear system theory.

The step-response model represents the evolution of radiative forcing in a scenario experiment by a series of step changes in radiative forcing (with one step taken at the beginning of each year). The method makes two linear assumptions. First, the response to each annual forcing step is estimated by linearly scaling the response in a CO_2 step experiment according to the magnitude of radiative forcing change. Second, the response y_i at year i of a scenario experiment is estimated as a sum of responses to all previous annual forcing changes (see Figure 1 of Good et al., 2013 for an illustration):

$$y_i = \sum_{j=0}^{i} w_{i-j} x_j \qquad \text{(1a)}$$

where x_j is the response of the same variable in year j of the CO_2 step experiment. w_{i-j} scales down the response from the step experiment (x_j) to match the annual step change in radiative forcing from year i to year j of the scenario (denoted ΔF_{i-j}):

$$w_{i-j} = \frac{\Delta F_{i-j}}{\Delta F_s} \qquad (1b)$$

where ΔF_s is the radiative forcing change in the CO_2 step experiment. All quantities are expressed as anomalies with respect to a constant-forcing control experiment.

This approach can in principle be applied at any spatial scale for any variable for which the assumptions are plausible (e.g. Chadwick et al., 2013).

7 Linear and non-linear mechanisms, and the relevance of abruptCO2 experiments

Here we discuss further, with examples, the distinction between linear and nonlinear mechanisms, when they are important, and the relevance of abruptCO2 experiments.

7.1 Linear mechanisms: different timescales of response

Even in a linear system, regional climate change per K of global warming will evolve during a scenario simulation. This happens because different parts of the climate system have different timescales of response to forcing change.

This may be due to different effective heat capacities. For example, the ocean mixed layer responds much faster than the deeper ocean, simply due to a thinner column of water (Li and Jarvis, 2009). However, some areas of the ocean surface (e.g. the Southern Ocean and south-east subtropical Pacific) show lagged warming, due to a greater connection (via upwelling or mixing) with the deeper ocean (e.g. Manabe et al., 1990;Williams et al., 2008). The dynamics of the ocean circulation and vegetation may also have their own inherent timescales (e.g. vegetation change may lag global warming by years to hundreds of years, Jones et al., 2009). At the other extreme, some responses to CO2 forcing are much faster than global warming: such as the direct response of global mean precipitation to forcings (Allen and Ingram, 2002;e.g. Andrews et al., 2010;Mitchell et al., 1987) and the physiological response of vegetation to CO2 (Field et al., 1995).

In a linear system, patterns of change per K of global warming are sensitive to the forcing history. For example in Figure 1, a scenario is illustrated where forcing is ramped up, then stabilized. Three periods are highlighted, which may have different patterns of change per K of global warming, due to different forcing histories: at the leftmost point, faster responses will be relatively more important, whereas at the right, the slower responses have had some time to catch up. This is illustrated in Figure 2 for sea-level rise. The blue curves show that for RCP2.6, global-mean warming ceases after 2050, while sea-level rise continues at roughly the same rate throughout the century. This is largely because deep ocean heat uptake is much slower than ocean mixed-layer warming.

By design, abruptCO2 experiments separate different timescales of GCM response to forcing change. This is used, for example, (Gregory et al., 2004) to estimate radiative forcing and feedback parameters for GCMs: plotting radiative flux anomalies against global mean warming can separate 'fast' and 'slow' responses (see e.g. Figure 3).

7.2 Non-linear responses

Nonlinear mechanisms arise for a variety of reasons. Often, however, it is useful to describe them as state-dependent feedbacks. For example, the snow-albedo feedback becomes small at high or low snow depth. Sometimes, nonlinear mechanisms may be better viewed as simultaneous changes in pairs of properties. For example, convective precipitation is broadly a product of moisture content and dynamics (Chadwick and Good, 2013;Chadwick et al., 2012). Both moisture content and atmospheric dynamics respond to CO2 forcing, so in general we might expect convective precipitation to have a nonlinear response to CO2 forcing. Of course, more complex nonlinear responses exist, such as for the Atlantic Meridional Overturning Circulation.

In contrast to linear mechanisms, nonlinear mechanisms are sensitive to the magnitude of forcing. For example, the two points highlighted in Figure 4 may have different patterns of change per K of global warming, due to nonlinear mechanisms.

An example is given in Figure 5, which shows the albedo feedback declining with increased global temperature, due to declining snow and ice cover, and the remaining snow and ice being in areas of lower solar insolation (Colman and McAvaney, 2009).

AbruptCO2 experiments may be used to separate nonlinear from linear mechanisms. This can be done by comparing the responses at the same timescale in different different abruptCO2 experiments. Figure 6 compares abrupt2xCO2 and abrupt4xCO2 experiments over years 50-149. A 'doubling difference' is defined, measuring the difference in response to the first and second CO2 doublings. In most current simple climate models (e.g. Meinshausen et al., 2011), the radiative forcing from each successive CO2 doubling is assumed identical (because forcing is approximately linear in log[CO2], Myhre et al., 1998). With this assumption, a linear system would have zero doubling difference everywhere. Therefore, the doubling difference is used as a measure of nonlinearity. The question of which abruptCO2 experiments to compare, and over which timescale, is discussed in section 5.

In some GCMs, the forcing per CO2 doubling has been shown to vary with CO2 (Colman and McAvaney, 2009;Jonko et al., 2013). However, this variation depends on the specific definition of forcing used (Jonko et al., 2013). Currently this is folded into our definition of nonlinearity. If a robust definition of this forcing variation becomes available in future, it could be used to scale out any difference in forcing between pairs of abruptCO2 experiments, to calculate an 'adjusted doubling difference'.

As an example, Figure 7 maps the response to abrupt2xCO2 and abrupt4xCO2, and the doubling difference, for precipitation in HadGEM2-ES over the ocean (taken from Chadwick and Good). The nonlinearities are large - comparable in magnitude to the responses to abrupt2xCO2, albeit with a different spatial pattern.

8 Experimental design

nonlinMIP is composed of a set of abruptCO2 experiments (the primary tools), plus a CO2-forced transient experiment. These build on the CMIP5 and CMIP6 DECK protocols (the required runs from these are detailed in Table 1). The additional nonlinMIP runs (Table 2) are assigned three priority levels. Three options for participation are: 1) only the ‘essential’ simulation; 2) all ‘high priority’ plus the ‘essential’ simulations; or, preferably, 3) all simulations. The experiments in Table 1 are required in all cases. All experiments must be initialized from the same year of a pre-industrial control experiment, except for abrupt4xto1x (see Table 2). A typical analysis procedure is outlined in section 5.

The nonlinMIP design is presently limited to CO2 forcing, although the same principles could be applied to other forcings.

9 Basic analysis principles

This section outlines the general principles behind analysis of nonlinMIP results. The primary idea is to find where the step-response model (section 2) breaks: since the step-response model is based on a linear assumption, this amounts to detecting non-linear responses.

The aim is to focus subsequent analysis. If non-linearities in a quantity of interest are found to be small, then analysis may focus on understanding different timescales of response from a single abruptCO2 experiment: linearity means that the physical response (over a useful range of CO2 concentrations) is captured by a single abruptCO2 experiment. This represents a considerable simplification. If, on the other hand, non-linearities are found to be important, the focus shifts to understanding the different responses in different abruptCO2 experiments. The choice of which abruptCO2 experiments to focus on, and over which timescales, is discussed below.

9.1 First step: check basic traceability of abrupt4xCO2 to the transient-forced response near 4xCO2

This is to confirm that the abruptCO2 experiments contain realistic physical responses in the variables of interest (as previously done for global-mean temperature and heat uptake for a range of CMIP5 models (Good et al., 2013), and for other global-mean quantities for HadCM3 (Good et al., 2011). This also, rules out the most pathological non-linearities (e.g. if the response to an abrupt CO2 change in a given GCM was unrealistic).

The linear step-response model should first be used with the abrupt4xCO2 response, to predict the response near year 140 of the 1pctCO2 experiment (i.e. near 4xCO2). This prediction is then compared with the actual GCM 1pctCO2 result. This should first be done for global mean temperature: this assessment has been performed for a range of CMIP5 models (Good et al., 2013; see Figure 8), giving an idea of the level of accuracy expected. If the abruptCO2 response is fundamentally unrealistic, it is likely to show up in the global temperature change. This approach may then be repeated for spatial patterns of warming, and then for the quantities of interest. Abrupt4xCO2 is used here as it has larger signal/noise than abrupt2xCO2, yet is representative of forcing levels in a business-as-usual scenario by 2100. However, the tests may also be repeated using abrupt2xCO2 – but compared with year 70 of the 1pctCO2 experiment (i.e. at 2xCO2).

The step-response model emulation under these conditions should perform well for most cases: the state at year 140 of the 1pctCO2 experiment is very similar to that of abrupt4xCO2 (same forcing, similar global-mean temperature), so errors from non-linear mechanisms should be minimal. If large errors are found, this may imply caution about the use of abruptCO2 experiments for these variables, or perhaps point to novel non-linear mechanisms that may be understood by further analysis.

9.2 Second step: detecting nonlinear responses

Having established some level of confidence in the abruptCO2 physical response, the second step is to look for nonlinear responses. This first involves repeating the tests from step 1 above, but for different parts of the 1pctCO2 and 1pctCO2 ramp-down experiments, and using different abruptCO2 experiments for the step-response model.

An example is given in Figure 9 (but for different transient-forcing experiments). This shows results for global-mean precipitation in the HadCM3 GCM (Good et al., 2012). Here, the step-response model prediction using abrupt4xCO2 (red curves) only works where a transient-forced experiment is near to 4xCO2. Similarly, the prediction using abrupt2xCO2 (blue curves) works only near 2xCO2. Otherwise, quite large errors are seen, and the predictions with abrupt2xCO2 and abrupt4xCO2 are quite different from each other. This implies that there are large non-linearities in the precipitation response in this GCM, and that they may be studied by comparing the responses in the abrupt2xCO2 and abrupt4xCO2 experiments.

Having identified some non-linear response, and highlighted two or more abruptCO2 experiments to compare (in the previous example, abrupt2xCO2 and abrupt4xCO2), the non-linear mechanisms may be studied in detail by comparing the responses in the different abruptCO2 experiments over the same timescale (e.g. via the doubling difference, as in Figures 6,7). This allows (Good et al., 2012;Chadwick and Good, 2013) non-linear mechanisms to be separated from linear mechanisms (not possible in a transient-forcing experiment).

10 Conclusions

This paper outlines the basic physical principles behind the nonlinMIP design, and the method of establishing traceability from abruptCO2 to gradual forcing experiments, before detailing the experimental design and finally some general analysis principles that should apply to most studies based on this dataset.

Acknowledgements

This work was supported by the Joint UK DECC/Defra Met Office Hadley Centre Climate Programme (GA01101). Nathaelle Bouttes received funding from the European Research Council under the European Community’s Seventh Framework Programme (FP7/2007-2013), ERC Grant Agreement 247220, project “Seachange.”

References

Allen, M. R., and Ingram, W. J.: Constraints on future changes in climate and the hydrologic cycle, Nature, 419, 224-+, 10.1038/nature01092, 2002.

Andrews, T., Forster, P. M., Boucher, O., Bellouin, N., and Jones, A.: Precipitation, radiative forcing and global temperature change, Geophysical Research Letters, 37, Artn L14701
Doi 10.1029/2010gl043991, 2010.

Andrews, T., and Ringer, M. A.: Cloud feedbacks, rapid adjustments, and the forcing-response relationship in a transient co2 reversibility scenario, Journal of Climate, 27, 1799-1818, Doi 10.1175/Jcli-D-13-00421.1, 2014.

Bouttes, N., Gregory, J. M., and Lowe, J. A.: The reversibility of sea level rise, Journal of Climate, 26, 2502-2513, Doi 10.1175/Jcli-D-12-00285.1, 2013.

Chadwick, R., Boutle, I., and Martin, G.: Spatial patterns of precipitation change in cmip5: Why the rich don't get richer., Journal of Climate, accepted, 2012.

Chadwick, R., and Good, P.: Understanding non-linear tropical precipitation responses to co2 forcing, Geophysical Research Letters, 40, 10.1002/grl.50932, 2013.

Chadwick, R., Wu, P. L., Good, P., and Andrews, T.: Asymmetries in tropical rainfall and circulation patterns in idealised co2 removal experiments, Climate Dynamics, 40, 295-316, DOI 10.1007/s00382-012-1287-2, 2013.

Colman, R., and McAvaney, B.: Climate feedbacks under a very broad range of forcing, Geophysical Research Letters, 36, L01702
10.1029/2008gl036268, 2009.

Field, C. B., Jackson, R. B., and Mooney, H. A.: Stomatal responses to increased co2 - implications from the plant to the global-scale, Plant Cell Environ, 18, 1214-1225, DOI 10.1111/j.1365-3040.1995.tb00630.x, 1995.

Forster, P. M., Andrews, T., Good, P., Gregory, J. M., Jackson, L., and Zelinka, M. D.: Evaluating adjusted forcing and model spread for historical and future scenarios in the cmip5 generation of climate models, Journal of Geophysical Research-Atmospheres (accepted pending minor revisions), 2012.

Good, P., Gregory, J. M., and Lowe, J. A.: A step-response simple climate model to reconstruct and interpret aogcm projections, Geophysical Research Letters, 38, Artn L01703
Doi 10.1029/2010gl045208, 2011.

Good, P., Ingram, W., Lambert, F. H., Lowe, J. A., Gregory, J. M., Webb, M. J., Ringer, M. A., and Wu, P. L.: A step-response approach for predicting and understanding non-linear precipitation changes, Climate Dynamics, 39, 2789-2803, DOI 10.1007/s00382-012-1571-1, 2012.

Good, P., Gregory, J. M., Lowe, J. A., and Andrews, T.: Abrupt co2 experiments as tools for predicting and understanding cmip5 representative concentration pathway projections, Climate Dynamics, 40, 1041-1053, DOI 10.1007/s00382-012-1410-4, 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, Geophysical Research Letters, 31, L03205
10.1029/2003gl018747, 2004.

Held, I. M., Winton, M., Takahashi, K., Delworth, T., Zeng, F. R., and Vallis, G. K.: Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing, Journal of Climate, 23, 2418-2427, Doi 10.1175/2009jcli3466.1, 2010.

Huntingford, C., and Cox, P. M.: An analogue model to derive additional climate change scenarios from existing gcm simulations, Climate Dynamics, 16, 575-586, 2000.

IPCC: Summary for policymakers, 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, edited by: Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Jones, C., Lowe, J., Liddicoat, S., and Betts, R.: Committed terrestrial ecosystem changes due to climate change, Nat Geosci, 2, 484-487, Doi 10.1038/Ngeo555, 2009.

Jonko, A. K., Shell, K. M., Sanderson, B. M., and Danabasoglu, G.: Climate feedbacks in ccsm3 under changing co2 forcing. Part ii: Variation of climate feedbacks and sensitivity with forcing, Journal of Climate, 26, 2784-2795, Doi 10.1175/Jcli-D-12-00479.1, 2013.

Li, S., and Jarvis, A.: Long run surface temperature dynamics of an a-ogcm: The hadcm3 4xco(2) forcing experiment revisited, Climate Dynamics, 33, 817-825, 10.1007/s00382-009-0581-0, 2009.

Manabe, S., Bryan, K., and Spelman, M. J.: Transient-response of a global ocean atmosphere model to a doubling of atmospheric carbon-dioxide, J Phys Oceanogr, 20, 722-749, Doi 10.1175/1520-0485(1990)020<0722:Troago>2.0.Co;2, 1990.

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, 95, 77, 2014.

Meinshausen, M., Raper, S. C. B., and Wigley, T. M. L.: Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, magicc6-part 1: Model description and calibration, Atmos Chem Phys, 11, 1417-1456, DOI 10.5194/acp-11-1417-2011, 2011.

Mitchell, J. F. B., Wilson, C. A., and Cunnington, W. M.: On co2 climate sensitivity and model dependence of results, Q J Roy Meteor Soc, 113, 293-322, 1987.
Mitchell, T. D.: Pattern scaling - an examination of the accuracy of the technique for describing future climates, Climatic Change, 60, 217-242, 2003.
Myhre, G., Highwood, E. J., Shine, K. P., and Stordal, F.: New estimates of radiative forcing due to well mixed greenhouse gases, Geophysical Research Letters, 25, 2715-2718, 1998.
Santer, B., Wigley, T., Schlesinger, M., and Mitchell, J. F. B.: Developing climate scenarios from equilibrium gcm results, Report No. 47, Max Planck Institute for Meteorology, Hamburg, 1990.
Schaller, N., Cermak, J., Wild, M., and Knutti, R.: The sensitivity of the modeled energy budget and hydrological cycle to co2 and solar forcing, Earth Syst Dynam, 4, 253-266, DOI 10.5194/esd-4-253-2013, 2013.
Schaller, N., Sedláček, N. J., and Knutti, R.: The asymmetry of the climate system's response to solar forcing changes and its implications for geoengineering scenarios, Journal of Geophysical Research: Atmospheres, 10, 5171–5184, 2014.
Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L.: Contribution of working group i to the fourth assessment report of the intergovernmental panel on climate change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.
Stott, P., Good, P., Jones, G., Gillett, N., and Hawkins, E.: The upper end of climate model temperature projections is inconsistent with past warming, Environ Res Lett, 8, Artn 014024
Doi 10.1088/1748-9326/8/1/014024, 2013.
Tebaldi, C., and Arblaster, J. M.: Pattern scaling: Its strengths and limitations, and an update on the latest model simulations, Climatic Change, 122, 459-471, DOI 10.1007/s10584-013-1032-9, 2014.
Williams, K. D., Ingram, W. J., and Gregory, J. M.: Time variation of effective climate sensitivity in gcms, Journal of Climate, 21, 5076-5090, Doi 10.1175/2008jcli2371.1, 2008.
Zelinka, M. D., Klein, S. A., Taylor, K. E., Andrews, T., Webb, M. J., Gregory, J. M., and Forster, P. M.: Contributions of different cloud types to feedbacks and rapid adjustments in cmip5, Journal of Climate, 26, 5007-5027, Doi 10.1175/Jcli-D-12-00555.1, 2013.

Table 1. List of CMIP5/CMIP6 DECK experiments required by nonlinMIP.

Experiment	Description	Role
piControl	Pre-industrial control experiment	
Abrupt4xCO2	CO2 abruptly quadrupled, then held constant for 150 years.	Separate different timescales of response.
1pctCO2	CO2 increased at 1% per year for 140 years (i.e. as CMIP5 1pctCO2 experiment), then decreased by 1% per year for 140 years (i.e. returning to pre-industrial conditions).	To test traceability of the abruptCO2 experiments to more realistic transient-forcing conditions. Adding the ramp-down phase explores physics relevant to mitigation and geo-engineering scenarios.

Table 2. NonlinMIP experimental design. Three options are: only the 'essential' simulation; all 'high priority' plus the 'essential' simulations; or, preferably, all simulations. The experiments in Table 1 are required in all cases.

Experiment (priority)	Description	Role
Abrupt2xCO2 (essential)	As abrupt4xCO2 (see Table 1), but at double pre-industrial CO2 concentration.	To diagnose non-linear responses (in combination with abrupt4xCO2). Assess climate response and (if appropriate) make climate projections with the step-response model at forcing levels more relevant to mid- or low-forcing scenarios.
1pctCO2 ramp-down (high priority)	Initialised from the end of 1pctCO2. CO2 is decreased by 1% per year for 140 years (i.e. returning to pre-industrial conditions).	To test traceability of the abruptCO2 experiments to more realistic transient-forcing conditions. Adding the ramp-down phase explores a much wider range of physical responses, providing a sterner test of traceability. Relevant also to mitigation and geo-engineering scenarios, and offers a sterner test of.
Extend both abrupt2xCO2 and abrupt4xCO2 by 100 years (high priority)		Allow traceability tests (via the step-response model) against most of the 1pctCO2 ramp-up-ramp-down experiment. Explore longer timescale responses than in CMIP5 experiment. Permit improved signal/noise in diagnosing some regional-scale non-linear responses Provide a baseline control for the abrupt4xto1x experiment.
Abrupt4xto1x (medium priority)	Initialised from year 100 of abrupt4xCO2, CO2 is abruptly returned to pre-industrial levels, then held constant for 150 years.	Quantify non-linearities over a larger range of CO2 (quantifies responses at 1xCO2). Assess non-linearities that may be associated with the direction of forcing change.
Abrupt8xCO2 (medium priority)	As abrupt4xCO2, but at 8x pre-industrial CO2 concentration. Only 150 years required here.	Quantify non-linearities over a larger range of CO2.

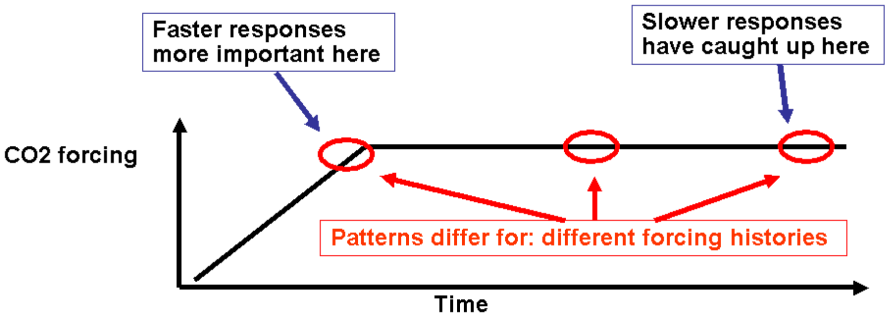


Figure 1. Schematic illustrating a situation where linear mechanisms can cause climate patterns to evolve. This represents a scenario where forcing (black line) is ramped up, then stabilised.

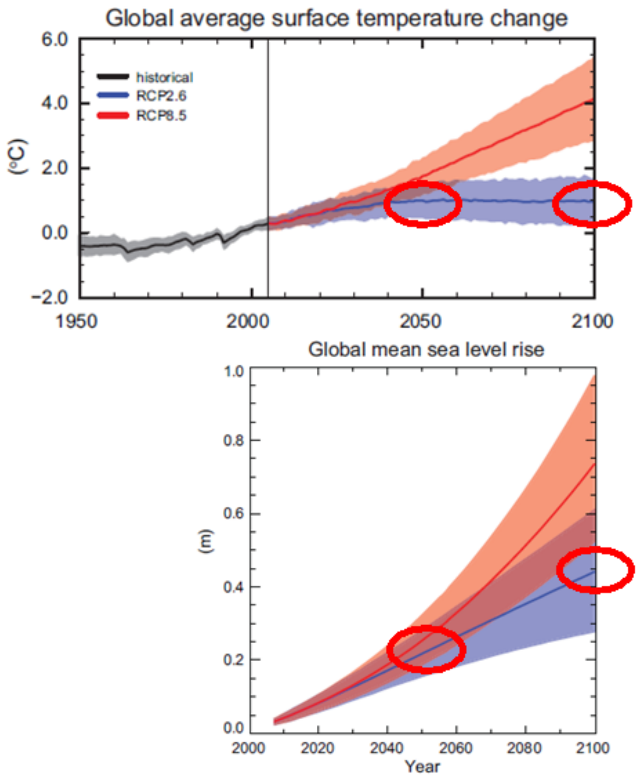


Figure 2. Adapted (red ovals overlaid) from the IPCC Fifth Assessment Report (IPCC, 2013), Figures SPM.7 and SPM.9. Global mean warming (top) and global mean sea level rise (bottom), relative to 1986-2005, for rcp8.5 (red) and rcp2.6 (blue).

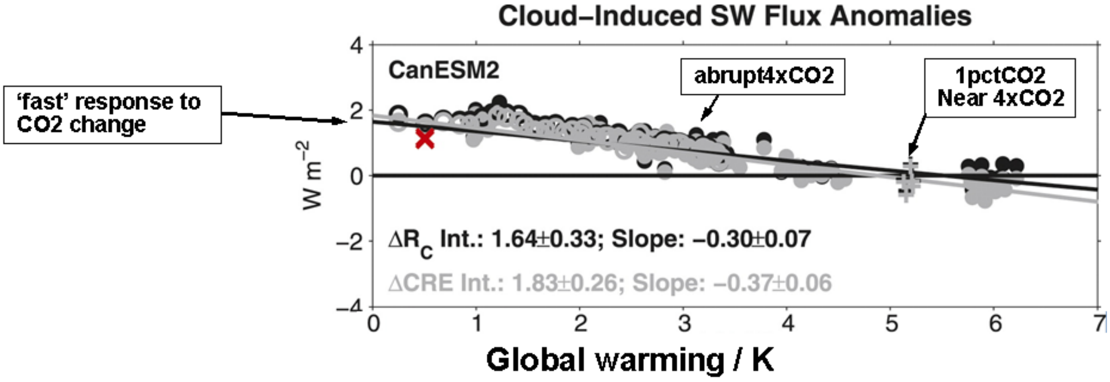


Figure 3. Illustrating a method (Gregory et al., 2004) for separating 'fast' and 'slow' responses to radiative forcing change. Figure adapted (labels in rectangles overlaid) from Zelinka et al. (2013). Global-mean cloud-induced SW flux anomalies against global warming, for the CanESM2 model (black & grey represent two methods of calculating cloud-induced fluxes). This also illustrates one test of traceability of abrupt4xCO2 to 1pctCO2 responses: the linear fit to the abrupt4xCO2 response (straight lines) passes through the 1pctCO2 response near 4xCO2 (i.e. near year 140 of that experiment).

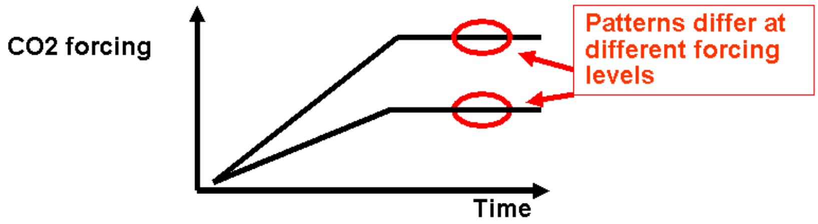


Figure 4. Schematic illustrating the point that nonlinear mechanisms can cause climate patterns to differ at different forcing (and hence global temperature) levels.

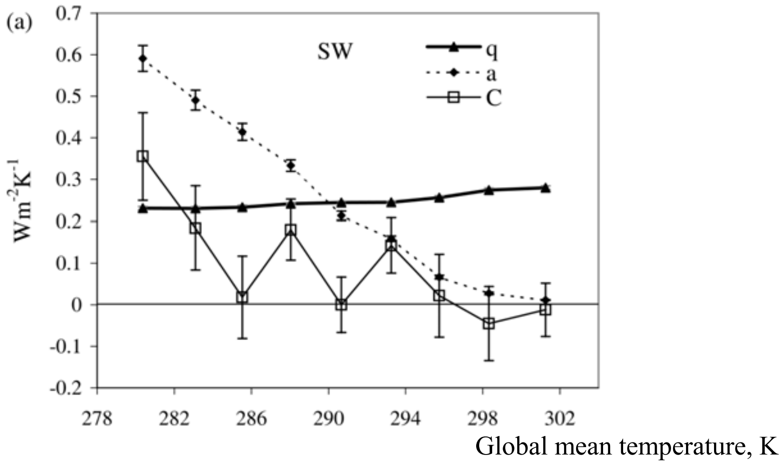


Figure 5. Albedo feedback (dotted line) strength (y-axis) decreasing with global mean temperature (x-axis, K) in a climate model (figure from Colman and McAvaney, 2009).

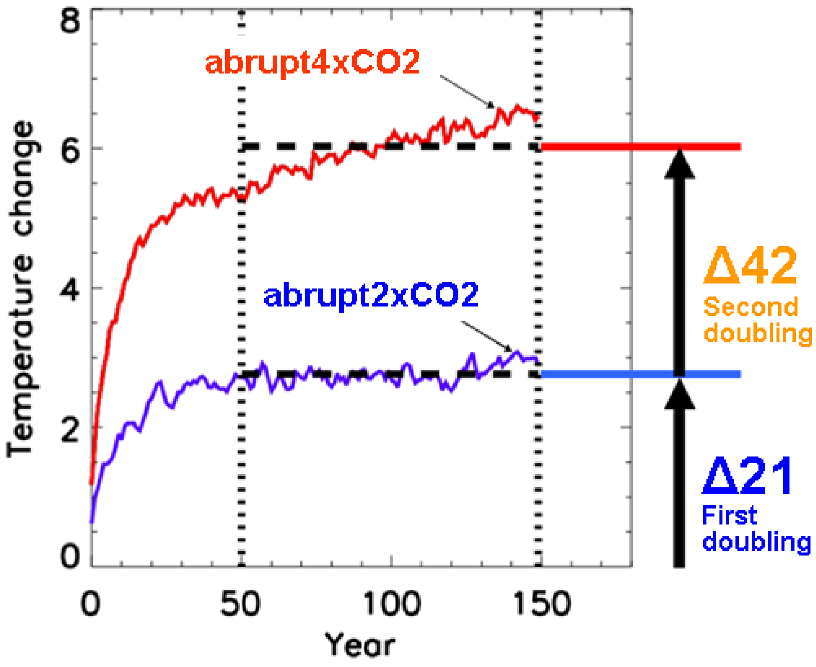


Figure 6. Defining the 'doubling difference'. Doubling difference = Δ42 – Δ21 (the difference in response between the first and second CO2 doublings. This is defined for a specific timescale after the abrupt CO2 change – in this example, it is the mean over years 50-149.

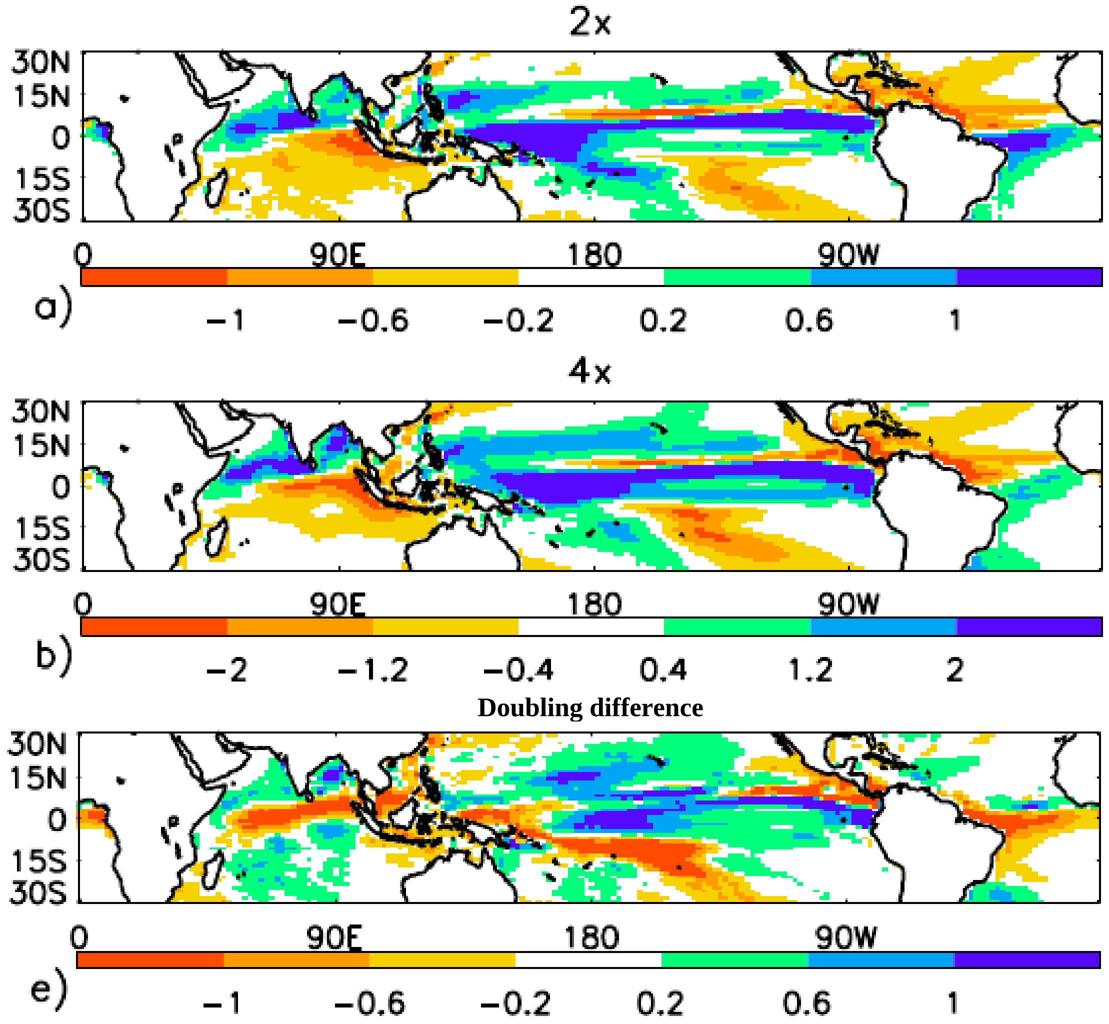


Figure 7. Non-linear regional precipitation responses over the ocean in HadGEM2-ES (figure from Chadwick and Good, 2013). Precipitation change (mm/day) averaged over years 50-149 for (top) abrupt2xCO2 and (middle) abrupt4xCO2, and the doubling difference (bottom). Note that the top and bottom panels have the same scale.

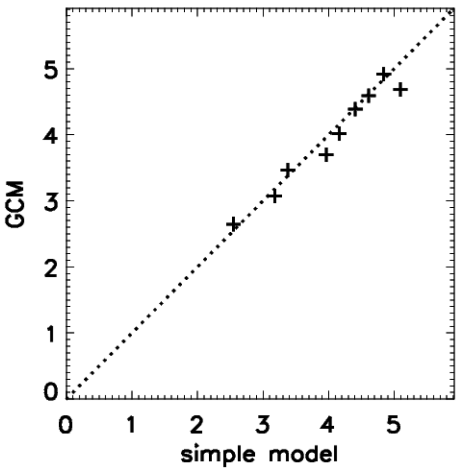


Figure 8. Checking basic traceability of abrupt4xCO2 to a transient forcing experiment (1pctCO2) (figure from Good et al., 2013). Global-mean warming (K) averaged over years 120-139 of 1pctCO2 for (y-axis) the GCM simulation and (x-axis) the reconstruction from abrupt4xCO2 using the step-response method.

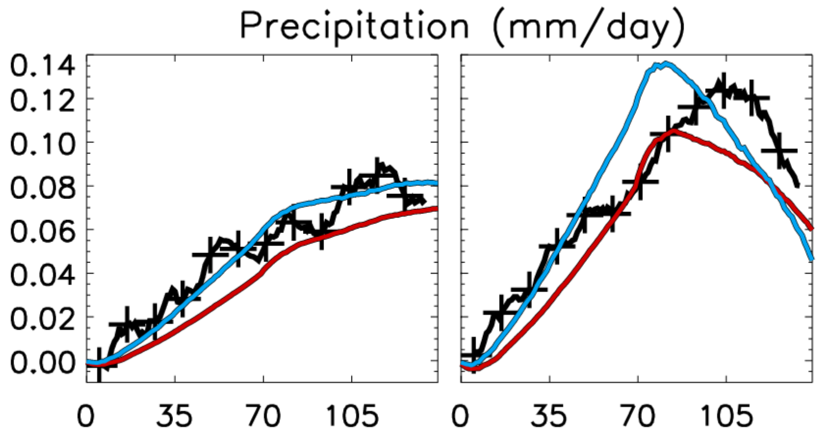


Figure 9. Finding nonlinear responses in transient forcing experiments. (figure from Good et al., 2012). Left: where CO2 is increased by 1% per year, then stabilised at 2x pre-industrial levels. Right: where CO2 is increased by 2% per year for 70 years, then decreased by 2% per year for 70 years. Black: GCM. Red: step-response model using the abrupt4xCO2 response. Blue: the abrupt2xCO2 response.

Proposal of Detection and Attribution Model Intercomparison Project (DAMIP)

Last updated: 10 June 2015

- Name of MIP*

  Detection and Attribution Model Intercomparison Project (DAMIP)

- Co-chairs of MIP as contacts*

Nathan Gillett, Canadian Centre for Climate Modelling and Analysis (nathan.gillett@ec.gc.ca)
Hideo Shiogama, National Institute for Environmental Studies (shiogama.hideo@nies.go.jp)

- Members of the Scientific Steering Committee*

Bernd Funke, Instituto de Astrofísica de Andalucía (bernd@iaa.es)
Gabriele Hegerl, University of Edinburgh (gabi.hegerl@ed.ac.uk)
Reto Knutti, ETH Zurich (reto.knutti@env.ethz.ch)
Katja Matthes, GEOMAR Helmholtz Centre for Ocean Research Kiel (kmatthes@geomar.de)
Benjamin D. Santer, Lawrence Livermore National Laboratory (santer1@llnl.gov)
Dáithí Stone, Lawrence Berkeley National Laboratory (dstone@lbl.gov)
Claudia Tebaldi, National Center for Atmospheric Research (claudia.tebaldi@gmail.com)

- Link to website (if available)*

  Not yet available.

- References (if available)*

Bindoff, N.L., et al., 2013: Detection and Attribution of Climate Change: from Global to Regional. 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 (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jones, G. S., P. A. Stott, and N. Christidis, 2013: Attribution of observed historical near surface temperature variations to anthropogenic and natural causes using CMIP5 simulations. J. Geophys. Res. Atmos., doi:10.1002/jgrd.50239.

Gillett, N. P., V. K. Arora, D. Matthews, P. A. Stott, and M. R. Allen, 2013: Constraining the ratio of global warming to cumulative CO2 emissions using CMIP5 simulations. J. Clim., doi:10.1175/JCLI-D-12–00476.1.

- Goal of the MIP and a brief overview*

  The primary goals of DAMIP are to facilitate improved estimation of the contribution of anthropogenic and natural forcing changes to observed global warming; to facilitate improved estimation of the contribution of those forcings to observed global and regional changes in other climate variables; to contribute to the estimation of how historical emissions have altered and are altering contemporary climate risk; and to facilitate and improve observationally-constrained projections of future climate change. Detection and attribution studies typically require unforced control simulations and historical simulations including all major anthropogenic and natural forcings. Such simulations will be carried out as part of the DECK and the *CMIP6 historical simulation* (hereafter we referred to the *CMIP6 historical simulation* as histALL). In addition such studies require additional simulations with individual forcings or subsets of forcings. We propose some such separated forcing experiments as Detection and Attribution MIP (DAMIP) for CMIP6. Combinations of histALL and separated forcing experiments from models participating in CMIP6 will be useful for model evaluation, better understanding of historical climate changes, and for deriving observational constraints on future climate change projections.

- An overview of the proposed experiments*

We propose some historical experiments using individual forcings or subsets of forcings. These experiments are CO_2-concentration driven for ESMs. These simulations should start at the same time as the histALL simulations and continue to at least 2020. Forcings identical to those in the histALL simulations should be used up to the end of those simulations, followed by forcings from the SSP2-4.5 simulation (Tier 1 of ScenarioMIP). Multi-member ensembles are vital for the separation of forced responses and internal variability. We require at least 3 ensemble members with different initial conditions for each experiment, and recommend that modeling groups which cannot afford to do this for all requested runs start by carrying out at least 3-member ensembles of the Tier 1 simulations. We also request three extension experiments with individual forcings up to 2100 under SSP2-4.5: well-mixed GHG changes only; stratospheric ozone changes only; and anthropogenic aerosol changes only. The minimum ensemble size of these is one. We also recommend modelling groups to perform a 500-year or longer piControl run to allow robust estimates of internal variability.

We propose four Tier 1 experiments for DAMIP/CMIP6. The first one is the enlargement of the ensemble size of **histALL** (the CMIP6 historical simulation) to at least three members. The other three Tier 1 experiments are Natural-only (**histNAT**), GHG-only (**histGHG**) and Aerosols-only (**histAER**) ensembles. Here, “XXX-only” mean that the agent XXX changes as in the histALL runs, but the other conditions are imposed and kept constant as in the piControl experiments. We require that forcing agents are perturbed exactly as in the histALL simulations: For example in the histGHG simulations the same well-mixed GHG concentrations are prescribed as in the histALL simulations. We request modelling groups to report what sets of emissions and boundary conditions are used in each run.

Three Tier 2 experiments are proposed: the extension of GHG-only up to 2100 (**ssp245GHG**), Stratospheric-Ozone-only (**histSOZ**) and the extension of Stratospheric-Ozone-only to 2100 (**ssp245SOZ**). Tier 3 experiments are Volcanic-only (**histVLC**), Solar-only (**histSOL**), the extension of Aerosol-only up to 2100 (**ssp245AER**), histALL with alternate estimates of aerosols (**histALL/estAER2**) and histALL with alternate estimates of NAT forcings (**histALL/estNAT2**).

Both DAMIP and DCPP propose an initial condition ensemble of histALL simulations. histALL, histNAT, histGHG, histAER, ssp245GHG and ssp245AER of DAMIP correspond closely to transient AGCM experiments for estimates of radiative forcing proposed in RFMIP-ERF. Combinations of DAMIP and RFMIP-Historical simulations will allow uncertainties in the aerosol response to be separated into those associated with simulating the climate response to a given distribution of aerosols, and the full uncertainty based on specified aerosol precursor emissions. DAMIP also co-sponsors "ALL minus land-use (LND_noLULCC_hist)" simulations of LUMIP. Two AerChemMIP experiments are closely related to histGHG of DAMIP: One is similar to histGHG except that ozone is allowed to vary, and a second includes historical changes in both GHGs and aerosols. histALL and histNAT runs from DAMIP will be used in diagnoses of GMMIP. histVLC runs are useful for GeoMIP and

VolMIP. Combinations of SSP2-4.5 runs of ScenarioMIP and ssp245GHG, ssp245SOZ and ssp245AER of DAMIP will allow the investigation of future climate responses to different forcing agents and observational constraints on future projections.

- An overview of the proposed evaluation/analysis of the CMIP DECK and CMIP6 experiments*

  A number of detection and attribution analyses of anthropogenic and natural forcing influences on historical climate changes are anticipated. These analyses are expected to address historical changes in temperature, the hydrological cycle, atmospheric circulation, ocean properties, cryospheric variables, extreme indices and other variables, from global to regional scales. The extension of DAMIP experiments from 2005 in CMIP5 to 2020 is essential to understand reasons for the recent *hiatus* of climate warming and improve signal-to-noise ratio for detection and attribution of changes in high-noise variables such as precipitation. The DAMIP experiments are also important for observational constraints on future climate change projections, climate sensitivity, TCR and TCRE. Using combinations of experiments from DAMIP and RFMIP, we can compare transient climate responses per unit radiative forcing across different forcing factors.

  It is anticipated that analyses of DAMIP simulations will form the basis of the assessment of the detection and attribution of climate change in the next IPCC assessment report. Observationally-constrained estimates based on DAMIP simulations may also provide a major contribution to projections of future climate in this report.

- Proposed timing *

  After modelling centers perform piControl (we recommend 500-year or longer simulations), histALL (the CMIP6 historical simulation) and SSP2-4.5 (ScenarioMIP Tier 1) experiments.

- For each proposed experiment for CMIP Phase 6**

Tier 1 experiments

(1.0) Enlarging ensemble size of the CMIP6 historical simulations to at least three members (histALL)

- the experimental design
  - All forcing historical simulations
  - Enlarging ensemble size of histALL to at least three members with different initial conditions. Please use forcings from SSP2-4.5 during 2015-2020.
  - Please provide outputs of experiments during 1850-2014 under the name of the CMIP6 historical simulation, and 2015-2020 as SSP2-4.5 of SceanrioMIP, not histALL.
  - DCPP proposes a 10-member ensemble of histALL up to 2030 also extended with SSP2-4.5.
- the science question and/or gap being addressed with this experiment

  Combinations of histALL, histNAT and histGHG will allow us to attribute observed climate changes to contributions from GHG, the other anthropogenic factors and natural forcing. Because better signal to noise ratio is vital to D&A analyses, we request at least 3 members for all historical experiments. Larger numbers of simulations also provide much larger samples of extreme events for climate risk analysis.

- possible synergies with other MIPs
  - This experiment will benefit all researchers who analyze historical climate changes, and the present climatology.
  - DCPP: DCPP proposes a 10 member ensemble of histALL up to 2030.
  - RFMIP-ERF: Combining radiative forcing estimated from RFMIP-ERF and transient climate responses from DAMIP (histALL, histNAT, histGHG, histAER, ssp245GHG and ssp245AER), we can investigate how feedbacks and adjustments vary with forcing factors.
  - RFMIP-Historical: Combinations of DAMIP (histALL, histNAT, histAER) and RFMIP-Historical will allow us to separate uncertainties in climate response based on specified aerosol evolution from the overall uncertainties in climate response to specified aerosol precursor emissions.
  - GMMIP: histALL and histNAT runs from DAMIP will be used in diagnoses of GMMIP.
  - ScenarioMIP: Combinations of histALL, histNAT, histGHG, histAER, ssp245GHG, ssp245AER of DAMIP and SSP2-4.5 of ScenarioMIP allow observational-constraints of uncertainties in future projections.
- potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, and (D) policy makers.

  Larger ensemble sizes of histALL should benefit (A) and (C) due to better signal to noise ratios of climate change signals and information about uncertainties associated with internal variability.

(1.1) Natural-only run (histNAT)

- the experimental design
  - Historical simulations forced by natural forcing agents only (i.e., solar irradiance change and volcanic activity), exactly as in histALL.
- the science question and/or gap being addressed with this experiment

  histALL and histNAT simulations will allow us to attribute observed changes to anthropogenic and natural influences. histALL, histNAT and histGHG simulations will allow us to attribute observed climate changes to contributions from GHG, the other anthropogenic factors and natural forcing. To better understand the role of natural forcings for climate is among the key aims of several SPARC activities (in particular Solaris/HEPPA and SSIRC). In this respect the experiments with natural forcing only (histNAT, histVLC and histSOL) are useful for SPARC.
- possible synergies with other MIPs
  - C20C+ Detection and Attribution Project: The event attribution project of C20C+ will make use of the histNAT and histALL simulations to estimate boundary SST conditions for their AGCM simulations of the hypothetical counterfactual world without human influences.
  - RFMIP-ERF: Combining radiative forcing estimated from RFMIP-ERF and transient climate responses from DAMIP (histALL, histNAT, histGHG, histAER,

    ssp245GHG and ssp245AER), we can investigate how feedbacks and adjustments vary with forcing factors.

    - RFMIP-Historical: Combinations of DAMIP (histALL, histNAT, histAER) and RFMIP-Historical will allow us to separate uncertainties in climate response based on specified aerosol evolution from the overall uncertainties in climate response to specified aerosol precursor emissions.
    - VolMIP: VolMIP proposes both historical and mechanism based simulations with a focus on volcanic eruptions.
    - GMMIP: histALL and histNAT runs from DAMIP will be used in diagnoses of GMMIP
    - ScenarioMIP: Combinations of histALL, histNAT, histGHG, histAER, ssp245GHG, ssp245AER of DAMIP and SSP2-4.5 of ScenarioMIP allow observational-constraints of uncertainties in future projections.
- potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, and (D) policy makers.

  histALL, histGHG and histNAT in CMIP5 were vital for IPCC AR5 to conclude “more than half of the observed increase in global mean surface temperature from 1951 to 2010 is very likely due to the observed anthropogenic increase in greenhouse gas concentrations”. The updated forcings, longer simulations, and larger ensemble sizes of these experiments using new models in DAMIP/CMIP6 will facilitate even more robust attribution assessments and a better understanding of observed climate changes. histALL and histNAT will be used for event attribution analyses of recent extreme weather and climate events, and can be used for D&A analyses of impact assessments. Those attribution studies will provide essential information for discussion of mitigation and adaptation policies.

(1.2) well-mixed GHG-only run (histGHG)

- the experimental design
    - Historical simulations forced by well mixed greenhouse gas changes only, as in the histALL simulations. Models with interactive chemistry schemes should either turn off the chemistry or use a preindustrial climatology of stratospheric and tropospheric ozone in their radiation schemes. This will ensure that ozone is fixed in all these simulations, and simulated responses in models with and without coupled chemistry are comparable.
- the science question and/or gap being addressed with this experiment
    - Combinations of histALL, histNAT and histGHG will allow the quantification of climate change attributable to GHG changes, other anthropogenic forcings and natural forcings.
    - Allows observationally-constrained TCR and TCRE to be estimated together the 1PCTCO2 simulation in the DECK.
- possible synergies with other MIPs
    - C4MIP: Carbon flux changes related to the GHG concentration changes.
    - RFMIP-ERF: Combining radiative forcing estimated from RFMIP-ERF and transient climate responses from DAMIP (histALL, histNAT, histGHG, histAER, ssp245GHG and ssp245AER), we can investigate how feedbacks and adjustments vary with forcing factors.

    - AerChemMIP: Two experiments of AerChemMIP (Experiment 1.1.1/1.1.2) are closely related to the histGHG simulation of DAMIP.
    - ScenarioMIP: Combinations of histALL, histNAT, histGHG, histAER, ssp245GHG, ssp245AER and SSP2-4.5 allow observational-constraints of uncertainties in future projections.
- potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, and (D) policy makers.

  histALL, histGHG and histNAT in DAMIP/CMIP6 will facilitate more robust attribution assessments and a better understanding of climate changes than those based on CMIP5. Furthermore, histGHG and ssp245GHG can be used to derive observationally constrained future climate projections, climate sensitivity, TCR and TCRE. Observationally-constrained projections will provide useful information to inform discussion of mitigation and adaptation policies.

(1.3) Anthropogenic-Aerosols-only runs (histAER/histAERchem)

Two experimental designs are proposed for histAER - Please select one of them.

If you like to perform both this experiment and the corresponding simulation in RFMIP-ERF, please apply the same setup.

(1.3a) Anthropogenic-Aerosols-only runs (histAER)

- the experimental design
    - Historical simulations forced by anthropogenic aerosol concentrations only or aerosol and aerosol precursor emissions only as in the histALL simulation (sulfate, black carbon, organic carbon, ammonia, NOx and VOCs).
    - (1.3a) is only for models in which changes in GHG concentrations do not affect aerosols and changes in aerosol precursors do not affect ozone. In addition models in which these interactions do occur, but for which (1.3b) cannot be implemented, should carry out this experiment.

(1.3b) Anthropogenic-Aerosols-only runs (histAERchem)

- the experimental design
    - Historical simulations forced by aerosol and aerosol precursor emissions only as in the histALL simulation (sulfate, black carbon, organic carbon, ammonia, NOx and VOCs).
    - Changes in well-mixed-GHGs, aerosol precursors and ozone precursors are prescribed as in histALL runs. However, in the radiation scheme, the concentrations of well-mixed-GHGs and the ozone climatology from the piControl runs are used. This procedure will allow the simulation of aerosol burdens consistent with histALL runs, and the simulation of their influences on climate.
    - (1.3b) is only for models in which changes in GHG concentrations affect aerosols or changes in aerosol precursors affect ozone.
- the science question and/or gap being addressed with this experiment

  Aerosols are a key source of uncertainty in historical and future climate simulations and the prime reason for high uncertainty in TCR and ECS constraints. Together with the histNAT and histALL simulations, these simulations will allow us to attribute observed climate changes to contributions from natural forcings, aerosols and “GHG+ozone+land use change". This approach will likely result in more tightly constrained estimates of attributable warming since the aerosol response, which is more uncertain, will be directly simulated.

- possible synergies with other MIPs
    - RFMIP-ERF: Combining radiative forcing estimated from RFMIP-ERF and transient climate responses from DAMIP (histALL, histNAT, histGHG, histAER, ssp245GHG and ssp245AER), we can investigate how feedbacks and adjustments vary with forcing factors.
    - RFMIP-Historical: Combinations of DAMIP (histALL, histNAT, histAER) and RFMIP-Historical will allow us to separate uncertainties in climate response based on specified aerosol evolution from the overall uncertainties in climate response to specified aerosol precursor emissions.
    - ScenarioMIP: Combinations of histALL, histNAT, histGHG, histAER, ssp245GHG, ssp245AER of DAMIP and SSP2-4.5 of ScenarioMIP allow observational-constraints of uncertainties in future projections.
    - AerChemMIP: histAER and the experiments of AerChemMIP are useful to understand climate responses to NTCF.
- potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, and (D) policy makers.

    Aerosols are large uncertainty sources for historical and future climate simulations. The histAER and ssp245AER experiments will allow the climate modelling community to further understand aerosol impacts on climate. Because the hydrological cycle and shortwave radiation are sensitive to aerosols, histAER and ssp245AER may be useful for impact studies regarding water availability and shortwave radiation inputs. A better understating of aerosol influence on climate is also important for policies controlling aerosol emissions.

Tier 2 experiments

(2.1) Extension of well-mixed GHG-only run (ssp245GHG)

- the experimental design
    - extensions of histGHG runs up to 2100 using the SSP2-4.5 concentrations. As in histGHG, models with interactive chemistry schemes should either run with the chemistry scheme turned off or use a preindustrial climatology of ozone in the radiation scheme.
- the science question and/or gap being addressed with this experiment

    Combinations of histALL, histGHG, histNAT, ssp245GHG and SSP2-4.5 (ScenarioMIP) will allow us to make estimates of future temperature changes that are constrained by observed historical changes. Simulated future responses to aerosols and GHG are scaled based on the scaling factors by which the historical simulated responses to aerosols and GHG must be multiplied to best fit observations.
- possible synergies with other MIPs
    - RFMIP-ERF: Combining radiative forcing estimated from RFMIP-ERF and transient climate responses from DAMIP (histALL, histNAT, histGHG, histAER,

    ssp245GHG and ssp245AER), we can investigate how feedbacks and adjustments vary with forcing factors.

  - ScenarioMIP: Allows the separation of future climate change signals of GHG and the other anthropogenic forcing factors. Combinations of histALL, histNAT, histGHG, histAER, ssp245GHG, ssp245AER of DAMIP and SSP2-4.5 of ScenarioMIP allow observational-constraints of uncertainties in future projections.

- potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, and (D) policy makers.

  The histGHG and ssp245GHG simulations can be used to obtain observationally constrained future climate projections, climate sensitivity, TCR and TCRE that wide communities of CM, IAM, IAV and policy makers are interested in.

(2.2) Stratospheric-Ozone-only (histSOZ)

- the experimental design
  - Historical simulations forced by changes in stratospheric ozone concentrations.
  - In models with coupled chemistry, the chemistry scheme should be turned off, and the simulated ensemble mean monthly mean 3D stratospheric ozone concentrations from the histALL simulations should be prescribed. Tropospheric ozone should be fixed at 3D long-term monthly mean piControl values, with a value of 100 ppbv ozone concentration in this piControl climatology used to separate the troposphere from the stratosphere.
  - In models without coupled chemistry the same stratospheric ozone prescribed in histALL should be prescribed. Stratospheric ozone concentrations will be provided by CCMI.
- the science question and/or gap being addressed with this experiment

  Stratospheric ozone changes have driven large changes in stratospheric temperature, and are also responsible for driving circulation changes in the Southern Hemisphere, with associated climate impacts. Only a few CMIP5 models carried out historical ozone simulations. A larger multi-model ensemble will allow us to more robustly identify in models and perhaps also in observations the influence of ozone on the stratosphere, the tropospheric circulation and climate, and the Southern Ocean and associated carbon cycle aspects.
- possible synergies with other MIPs
  - AerChemMIP: Comparison with the ODS-only simulation of AerChemMIP will allow the net climate effects of ODSs to compared with the net climate effects of stratospheric ozone changes, a key issue of concern for the WMO/UNEP Ozone Assessment.
- potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, (D) policy makers, and (E) Ozone research and policy community.

  These simulations will be of most relevance to (E) since they will allow the climate impacts of past ozone changes to be directly assessed.

(2.3) Extension of Stratospheric-Ozone-only run (ssp245SOZ/ssp245SOZchem)

- the experimental design
    - Extensions of histSOZ up to 2100 using the SSP2-4.5 scenario.
- the science question and/or gap being addressed with this experiment

  These simulations will allow the contribution of future stratospheric ozone changes to future climate change to be evaluated, including for example contributions to future Southern Hemisphere atmospheric circulation change, oceanic circulation changes, and carbon cycle impacts. These simulations will be relevant to future WMO/UNEP Ozone Assessments.
- possible synergies with other MIPs
    - WGCM Cryosphere CG: Ozone-only experiments are particularly relevant for the Cryosphere CG (Antarctic).
- potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, (D) policy makers, and (E) Ozone research and policy community.

  These simulations will be of most relevance to (E) since they will allow the climate impacts of future stratospheric ozone changes to be directly assessed.

Tier 3 experiments

(3.1) Volcanic-only run (histVLC)

- the experimental design

  Historical simulations forced by volcanic forcing as in histALL
- the science question and/or gap being addressed with this experiment

  histNAT, histVLC and histSOL allow us to investigate volcanic and solar influences on climate and to check additivity. The CMIP5 ensemble tended to overestimate the historical volcanic response. histVLC will be used for better understanding errors in the volcanic forcing and responses. To better understand the role of natural forcings for climate is among the key aims of several SPARC activities (in particular Solaris/HEPPA and SSIRC). In this respect the experiments with natural forcing only (histNAT, histVLC, and histSOL) are useful for SPARC.
- possible synergies with other MIPs
    - GeoMIP & VolMIP: The volcanic response of models can be validated against observations using histVLC, whereas GeoMIP experiments cannot. Thus histVLC experiments will provide useful context for interpreting simulated responses to stratospheric aerosol across models in the GeoMIP experiment. While VolMIP includes simulations of individual eruptions it does not include simulations of the transient response to historical eruptions, allowing better validation of long-term transient effects against observations, and its focus is on 19th century eruptions.
- potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, and (D) policy makers.

  Solar radiation management has recently attracted interest from the wide communities of climate model, IAM, IAV and policy makers. The histVLC experiment can be used for the validation of model responses to stratospheric aerosol injections against observations.

(3.2) Solar-only run (histSOL)

- the experimental design

  Historical simulations forced by solar forcing as in histALL

- the science question and/or gap being addressed with this experiment
  - Assess solar-only effects, separate solar and volcanic effects, additional output to clarify importance of radiation, ozone (interactive or prescribed) and stratospheric dynamics for solar signals.
- possible synergies with other MIPs
  - *To assess solar cycle contribution to prediction, SolarMIP runs could be used to estimate solar forcing spread and compare with Component A of DCPP (for which additional output is desirable, see below)*
- potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, and (D) policy makers.
  - *Primarily A, to a lesser extent B and D.*

(3.3) Extension of anthropogenic Aerosol-only run (ssp245AER/ ssp245AERchem)

- the experimental design
  - Extensions of histAER/histAERchem runs up to 2100 using the SSP2-4.5 scenario. Please use the same setup as the histAER/histAERchem runs but the SSP2-4.5 forcing.
- the science question and/or gap being addressed with this experiment
  - Combinations of histALL, histAER, histNAT, ssp245AER and SSP2-4.5 (ScenarioMIP) will allow us to make estimates of future temperature changes that are constrained by observed historical changes. Combining radiative forcing estimated from RFMIP-ERF and transient climate responses from DAMIP (histALL, histNAT, histGHG, histAER, ssp245GHG and ssp245AER), we can investigate how feedbacks and adjustments vary with forcing factors.
- possible synergies with other MIPs
  - RFMIP-ERF: Combining radiative forcing estimated from RFMIP-ERF and transient climate responses from DAMIP (histALL, histNAT, histGHG, histAER, ssp245GHG and ssp245AER), we can investigate how feedbacks and adjustments vary with forcing factors.
  - ScenarioMIP and AerChemMIP: Allows the separation of future climate change signals of aerosols and the other anthropogenic forcing factors. Combinations of histALL, histNAT, histGHG, histAER, ssp245GHG, ssp245AER of DAMIP and SSP2-4.5 of ScenarioMIP allow observational-constraints of uncertainties in future projections.
- potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, and (D) policy makers.

  Aerosols are large uncertainty sources for historical and future climate simulations. The histAER and ssp245AER experiments will allow the climate modelling community to further understand aerosol impacts on climate. Because the hydrological cycle and shortwave radiation are sensitive to aerosols, histAER and ssp245AER may be useful for impact studies regarding water availability and shortwave radiation inputs. A better understating of aerosol influence on climate is also important for policies controlling aerosol emissions.

(3.4) histALL with alternate estimates of anthropogenic aerosol emissions/concentrations (histALL/estAER2; Note distinction from original histALL submission is planned through use of the subexperiment field in the ESGF, pending standardisation of this field across CMIP6)

- the experimental design

  Same as histALL but forced with a different estimate of variations in anthropogenic aerosol emissions or concentrations. We are discussing with the aerosol emissions/concentrations groups for CMIP6 about the possibility that the production of these additional historical emission/concentration scenarios can be produced in parallel with the standard historical CMIP6 estimate.

- the science question and/or gap being addressed with this experiment

  This experiment will allow us to sample over uncertainties in aerosol forcing, and hence account for this source of uncertainty in estimates of attributable climate changes.

- possible synergies with other MIPs
  - C20C+ Detection and Attribution Project: The event attribution project of C20C+ will make use of the histALL/estAER2 with the histNAT and histNAT/estNAT2 simulations to estimate boundary SST conditions for their AGCM simulations of the hypothetical counterfactual world without human influences.

- potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, and (D) policy makers.

This subexperiment will help us to investigate the sensitivity of past climate change to uncertainties in our understanding of what has been driving these changes. In particular, testing the sensitivity to uncertainty in historical aerosol forcing will provide better understanding of the observational constraints of future greenhouse gas warming, with relevance to A, C, and D in terms of more confident estimates of uncertainty in predictions of future climate change under specified emissions scenarios.

(3.5) histALL with alternate estimates of solar and volcanic forcing (histALL/estNAT2; Note distinction from original histALL submission is planned through use of the subexperiment field in the ESGF, pending standardisation of this field across CMIP6)

- the experimental design

  Same as histALL but with second estimates of solar and volcanic forcing. We are discussing with the solar and volcanic forcing groups for CMIP6 about the possibility that the production of these additional historical forcing scenarios can be produced in parallel with the standard historical CMIP6 estimates.

- the science question and/or gap being addressed with this experiment

  This experiment will allow us to sample over uncertainties in natural forcings, and hence account for this source of uncertainty in estimates of attributable temperature changes. This experiment will also allow us to investigate the contribution of natural forcing uncertainties to simulated decadal temperature trends, such as for example the contribution of uncertainties in volcanic forcing to uncertainites in simulated temperature trends over the past two decades.

- possible synergies with other MIPs
  - C20C+ Detection and Attribution Project: The event attribution project of C20C+ will make use of the histALL/estAER2 with the histNAT and histNAT/estNAT2 simulations to estimate boundary SST conditions for their AGCM simulations of the hypothetical counterfactual world without human influences.
- potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, and (D) policy makers.

  A primary goal of attribution studies is to distinguish the magnitude of the role of anthropogenic forcing against natural forcing in historical climate change. With this estimate we will be able to understand the sensitivity of these estimates to limits in our understanding of past natural forcings, with relevance for D in terms of understanding the confidence in the anthropogenic role in past climate change.

➢ If possible, a prioritization of the suggested experiments, including any rationale **

Tier 1

(1.0) Enlarging ensemble size of the CMIP6 historical runs and the SSP2-4.5 runs of ScenarioMIP (2015-2020) to at least three members (histALL)

(1.1) Natural-only run (histNAT)

(1.2) well-mixed GHG-only run (histGHG)

(1.3) Anthropogenic-Aerosols-only run (histAER)

Tier 2

(2.1) Extension of well-mixed GHG-only run (ssp245GHG)

(2.2) Stratospheric-Ozone-only (histSOZ)

(2.3) Extension of Stratospheric-Ozone-only run (ssp245SOZ)

Tier 3

(3.1) Volcanic-only run (histVLC)

(3.2) Solar-only run (histSOL)

(3.3) Extension of Anthropogenic-Aerosols-only run (ssp245AER)

(3.4) histALL with alternate estimates of anthropogenic aerosol emissions or concentrations (histALL/estAER2)

(3.5) histALL with alternate estimates of solar and volcanic forcing (histALL/estNAT2)

To keep consistency between CMIP6 and the previous MIPs, we have histALL, histGHG and histNAT in Tier 1. The analysis of the CMIP5 ensemble has highlighted that aerosols remain the largest source of uncertainty in D&A analyses. Therefore we also propose histAER in the high priority, which will allow the aerosol response to be directly estimated and reduce uncertainties in regression coefficients. histALL, histNAT, histGHG and histAER correspond closely to some experiments of RFMIP and AerChemMIP.

The ssp245GHG simulations will be used for observational constraints of future projections, which will attract interest from wide communities. CCMI and AerChemMIP requested that histSOZ and ssp245SOZ be higher priority. Therefore we also place ssp245GHG, histSOZ and ssp245SOZ in Tier 2.

- All model output archived by CMIP6-Endorsed MIPs is expected to be made available under the same terms as CMIP output. Most modeling groups currently release their CMIP data for unrestricted use. If you object to open access to the output from your experiments, please explain the rationale. **

  No objection to open access.

- List of output and process diagnostics for the CMIP DECK/CMIP6 data request**
  - whether the variable should be collected for all CMIP6 experiments, or only some specified subset and whether the output is needed from the entire length of each experiment or some shorter period or periods;
  - whether the output might only be relevant if certain components or diagnostic tools are used interactively (e.g. interactive carbon cycle or atmospheric chemistry, or only if the COSP simulator has been installed);
  - whether this variable is of interest to downstream users (such as impacts researchers, WG2 users) or whether its principal purpose is for understanding and analysis of the climate system itself. Be as specific as possible in identifying why the variable is needed.
  - whether the variables can be regridded to a common grid, or whether there is essential information that would be compromised by doing this;
  - the relative importance of the various variables requested (indicated by a tiered listing) is required if the data request is large.

Standard outputs as in CMIP5, but subject to change based on further consultation with climate modeling communities, other MIPs (e.g., RFMIP, AerChemMIP, ScenarioMIP, CFMIP, DCPP, GMMIP and ISI-MIP) and WGCM GCs after April 2015.

Additional output for solar signal analysis:

1. *Zonal mean shortwave (SW) and longwave (LW) heating rates (as requested by DynVar)*
2. *2D or 3D ozone fields (prescribed or interactively calculated).*
3. *O_2 and O_3 photolysis rates from climate models with interactive chemistry (may be already requested by AerChemMIP?).*
4. *O_x as well as O_x total production and loss rates (may be already requested by AerChemMIP?)*
5. *TEM diagnostics (monthly mean v*, w*, EPflux divergence) indices (already requested by DynVar)*
6. *daily zonal mean temperatures and zonal wind*
7. *3D geopotential height at least at 10hPa level for the detection of sudden stratospheric warmings (SSWs) and the calculation of NAM and SAM*

Request for collecting the above mentioned additional output for DECK-AMIP & CMIP6 historical simulations. This output might be also requested for some runs in CFMIP (abrupt solar forcing change experiments), and DCPP (Component A experiments) in order to add the analysis to the experiments requested in SolarMIP.

- Any proposed contributions and recommendations for**

  - model diagnostics and performance metrics for model evaluation;
  - observations/reanalysis data products that could be used to evaluate the proposed experiments. Indicate whether these are available in the obs4MIPs/ana4MIPs database or if there are plans to include them;
  - tools, code or scripts for model benchmarking and evaluation in open source languages (e.g., python, NCL, R).

  Specification of the exact estimate of external forcing data used.

- Any proposed changes from CMIP5 in NetCDF metadata (controlled vocabularies), file names, and data archive (ESGF) search terms.**

  None.

- Explanation of any proposed changes (relative to CMIP5) that will be required in CF, CMOR, and/or ESGF.**

  NA

The Decadal Climate Prediction Project (DCPP)

Name of MIP: The Decadal Climate Prediction Project (DCPP)

Co-chairs of MIP: George Boer (george.boer@ec.gc.ca), Doug Smith (doug.smith@metoffice.gov.uk)

Members of the Scientific Steering Committee: The DCPP Panel consists of:

George Boer <George.Boer@ec.gc.ca>,
Christophe Cassou <cassou@cerfacs.fr>,
Francisco Doblas-Reyes <francisco.doblas-reyes@ic3.cat>,
Gokhan Danabasoglu <gokhan@ucar.edu>,
Ben Kirtman <bkirtman@rsmas.miami.edu>,
Yochanan Kushnir <kushnir@ldeo.columbia.edu>
Kimoto Masahide <kimoto@aori.u-tokyo.ac.jp>,
Jerry Meehl <meehl@ucar.edu>,
Rym Msadek <rym.msadek@noaa.gov>,
Wolfgang Mueller <wolfgang.mueller@mpimet.mpg.de >,
Doug Smith <doug.smith@metoffice.gov.uk>,
Karl Taylor <taylor13@llnl.gov>,
Francis Zwiers <fwzwiers@uvic.ca>

Link to website. DCPP material is available at (http://www.wcrp-climate.org/dcp-overview) and at (http://dcpp.pacificclimate.org/).

Goals and overview:

The decadal hindcast component of CMIP will follow the example of other coordinated experiments as a protocol-driven multi-model multi-national project with data production and data sharing as integral components.

The Goals of the decadal prediction component of CMIP include:

- the promotion of the science and practice of decadal prediction (forecasts on timescales up to and including 10 years)
- the provision of information potentially useful for the IPCC WG1 AR6 assessment report and other studies and reports on climate prediction and evolution
- the production and retention of a multi-year multi-model collection of decadal hindcast and forecast data in support of climate science and of use to the Global Framework for Climate Services

Scientific aspects of the DCPP include:

- a system view (data; analyses; initial conditions; ensemble generation; models and forecast production; post processing and assessment) of decadal prediction
- investigation of broad questions (e.g. sources and limits of predictability, current abilities with respect to decadal prediction, potential applications, …)
- provision of benchmarks against which to compare improvements in models and prediction quality

- information on processes and mechanisms of interest, e.g., the hiatus, climate shifts, AMOC etc., in a collection of hindcasts and forecasts

Practical aspects include:

- the coordination of efforts based on agreed experimental structures and timelines in order to promote research, intercomparison, multimodel approaches, applications, and to provide justification for research directions
- a contribution to the development of infrastructure, in particular a multi-purpose data archive of decadal hindcasts and forecasts useful for a broad range of scientific and application questions and of benefit to national and international climate prediction and climate services organizations

References and Publications:

Many decadal prediction papers have been published referring to the decadal component of CMIP5 and to other decadal prediction results. Chapter 11 of the AR5 also gives information and pertinent references.

The DCPP intends to contribute a paper on the DCPP experimental design to the GMD CMIP6 special issue and will also consider a restricted set of co-authored papers with modeling groups giving basic multi-model results.

Overview of the proposed experiments: The DCPP is organized into three Components:

- *Hindcasts*: the design and organization of a coordinated decadal prediction (hindcast) component of CMIP6 in conjunction with the seasonal prediction and climate modelling communities
- *Forecasts*: the ongoing production of experimental quasi-operational decadal climate predictions in support of multi-model annual to decadal forecasting and the application of the forecasts
- *Predictability, mechanisms, and case studies*: the organization and coordination of decadal climate predictability studies and of case studies of particular climate shifts and variations including the study of the mechanisms that determine these behaviours

Each component is explained in detail in the accompanying material

Overview of the proposed evaluation/analysis of CMIP DECK and CMIP6 experiments:

The DCPP is unique in bringing together researchers from communities with expertise in seasonal to interannual prediction (as represented by WGSIP), climate simulation (as represented by WGCM), and decadal variability and predictability in general (as represented by CLIVAR). The models used, and approaches taken, represent to varying degrees the special interests and abilities of these communities.

For climate models, control and sensitivity experiments are a necessary backdrop to climate change simulations. Most models used in the DCPP will also participate in other aspects of CMIP6 and will have performed DECK and 20th century climate change integrations as well as other simulations and MIP integrations. The data retained for these studies provides a background for DCPP-related studies of many different aspects of decadal prediction.

The DCPP therefore strongly encourages participation in the DECK experiments. However, it is recognized that for initialized climate prediction, initialization and verification using observation-based information is key and provides an alternative vantage point for understanding model and system behavior. DCPP participants from the forecasting community may place less emphasis on idealized experiments compared to participants from the climate simulation community. These forecasting models are, nevertheless, usually closely related to model versions which are used for climate simulation and for which the results of DECK simulations are available. The forecasting aspect encourages emphasis on methods of initializing models, generating ensembles of forecasts and, especially, on assessing results against observations. The two approaches represent complementary views for the understanding and prediction of forced and internally generated climate variations. A tiered set of retained data for the DCPP has been developed to assist in evaluation and analysis. It is intended to encompass a minimum set of forecast and analysis variables for verification and investigation but is not intended in any way to restrict the amount of data that groups retain for their DCPP integrations. In particular, these data are intended to permit the calculation of:

- means and variability statistics pertaining to the forecasts
- individual and multi-model predictability and skill measures
- information on atmospheric and oceanic structures and their evolution
- basic information on system budgets and balances

Additional special purpose data include:

- higher frequency data (for analysis of extremes, storminess etc.)
- data in support of other MIPs (see the “MIP connection table”)

We believe that the DCPP represents an important evolutionary advance from the CMIP5 decadal prediction component. The DCPP addresses a broad but integrated range of scientific issues broadly characterized as the ability of the system to be predicted on decadal timescales, the currently available skill, the mechanisms that control long timescale variability, and the ongoing production of forecasts of potential benefit for both science and applications.

Proposed timing:

The proposed timing for the DCPP generally follows that outlined for CMIP6 (slide 11). In particular, the availability of historical forcing and future scenario information are key to DCPP timing.

Experimental design:

- See the Goals and Overview section above for the scientific and practical questions addressed by the DCPP
- See the attached files which include some of the above material together with
  - A tiered list of proposed experiments (DCPP_Components_24Jun2015.doc).
  - Experiment and data request information in the files
    - DCPP_CMIP6DataRequestCompilationTemplate.xlsx

- DCPP_standard_output.xls

Changes from CMIP5 in documentation and data treatment:

The DCPP Panel’s Data Subgroup consisting of K. Taylor (a member of the WIP), F. Doblas-Reyes, R. Msadek and W. Mueller are our contacts with the WIP and the CMIP Panel. As noted also in the Experimental Design material, the hope is that, in conjunction with the WIP, a coordinated set of “basic” or “common” tiered data tables can be developed across MIPs together with “MIP specific” tables associated with individual MIPs.

The Decadal Climate Prediction Project (DCPP)

The term "decadal prediction" encompasses predictions on annual, multi-annual to decadal timescales. The possibility of making skilful forecasts on these timescales and the ability to do so is investigated by means of predictability studies and retrospective predictions (hindcasts) made using the current generation of climate models. Skilful decadal prediction of relevant climate parameters is a Key Deliverable of the WCRP's Grand Challenge of providing Regional Climate Information (http://www.wcrp-climate.org/index.php/gc-regionalclimate).

The DCPP consists of three Components. Groups are invited to participate in any or all of the Components, each of which are separately "tiered". It is not expected that all groups will participate in all experiments or tiers.

- *Component A, Hindcasts*: the design and organization of a coordinated decadal prediction (hindcast) component of CMIP6 in conjunction with the seasonal prediction and climate modelling communities
- *Component B, Forecasts*: the ongoing production of experimental quasi-operational decadal climate predictions in support of multi-model annual to decadal forecasting and the application of the forecasts
- *Component C, Predictability, mechanisms, and case studies*: the organization and coordination of decadal climate predictability studies and of case studies of particular climate shifts and variations including the study of the mechanisms that determine these behaviours

Many scientific and practical questions are involved. The understanding of the physical processes that govern the long timescale predictability of the climate system is vital to improving decadal predictions and these are explored using observations, climate model studies and the results of decadal hindcasts. The analysis of available observations for initializing forecasts, the improvement of the models used in the production of the forecasts, post processing of forecasts including bias adjustment, calibration and multi-model combination, together with the production and application of probabilistic decadal forecasts, are all involved in the research and development efforts contributing to the DCPP. As has been the case for weather forecasting, continued improvement in each of the components of a decadal forecasting system is expected to yield improvement in decadal prediction skill.

The Decadal Climate Prediction Panel in conjunction with the Working Group on Seasonal to Interannual Prediction (WGSIP), the Working Group on Coupled Modelling (WGCM), and the CLIVAR focus on decadal variability and predictability are involved in the coordination of the scientific and practical aspects of the DCPP.

DCPP Component A
A multi-year multi-model decadal hindcast experiment

The decadal hindcast component of CMIP follows the example of other coordinated experiments as a protocol-driven multi-model multi-national project with data production and data sharing as integral components.

The Goals of the decadal prediction component of CMIP include:
- the promotion of the science and practice of decadal prediction (forecasts on timescales up to and including 10 years)
- the provision of information potentially useful for the IPCC WG1 AR6 assessment report and other studies and reports on climate prediction and evolution
- the production and retention of a multi-year multi-model collection of decadal hindcast data in support of climate science and of use to the Global Framework for Climate Services (GFCS)

Scientific aspects of the DCPP to which Component A can contribute include:
- a system view (data; analyses; initial conditions; ensemble generation; models and forecast production; post processing and assessment) of decadal prediction
- investigation of broad questions (e.g. sources and limits of predictability, current abilities with respect to decadal prediction, potential applications, …)
- provision of benchmarks against which to compare improvements in models and prediction quality
- information on processes and mechanisms of interest (e.g., the hiatus, climate shifts, AMOC etc.) in a collection of hindcasts

Practical aspects of Component A include:
- the coordination of efforts based on agreed experimental structures and timelines in order to promote research, intercomparison, multimodel approaches, applications, and to provide justification for research directions
- a contribution to the development of infrastructure, in particular a multi-purpose data archive of decadal hindcasts useful for a broad range of scientific and application questions and of benefit to national and international climate prediction and climate services organizations

The basic elements of Component A are:
- a coordinated set of multi-model multi-member ensembles of retrospective forecasts initialized each year from 1960 to the present
- an associated hierarchy of data sets of results generally and readily available to the scientific and applications communities

Consultation and timing for Component A:
- The proposed timing for Component A generally follows that outlined for CMIP6 (slide 11). In particular, the availability of historical forcing and future scenario information are key to DCPP timing.

Details of the proposed Component A decadal prediction component are listed below.

DCPP Component A hindcast protocols

The approach parallels that of the "Near-term Decadal" component of CMIP5 (Taylor et al., 2009, dated 22 January, 2011, together with the Experiment Design Addendum at (http://cmippcmdi.llnl.gov/cmip5/experiment_design.html) although with important differences. Note the call for hindcasts to be produced every year, rather than every 5 years, over the hindcast period

Table 1. Basic Component A: Hindcast/forecast experiments

#	Experiment	Notes	# of years
TIER 1: Hindcast/forecast information			
A1	Ensembles of at least 5-year, but much preferably 10-year, *hindcasts* and *forecasts*	Coupled models with initialization based on observations Start date *every year* from 1960 to the present if at all possible; otherwise every second year at minimum. Start date on or before 31 Dec of the year preceding the forecast period (start dates on or before Nov 15 allow for DJF seasonal forecast results and are recommended) 10 ensemble members (more if possible) Prescribed CMIP6 historical values of atmospheric composition and/or emissions (and other conditions including volcanic aerosols). Future forcing as the SSP2-4.5 scenario.	(30-60)x10x(5-10) = 1500-6000 years of integration
TIER 2: To quantify the effects of initialization (encompasses CMIP6/historical simulations)			
A2	Ensembles of historical and near-future climate *simulations*	Made with the same model as used for hindcasts 1850 to 2030, with initial conditions from a preindustrial control simulation 10 ensemble members (more if possible) Prescribed historical and future forcing as for the A1 Experiment	170x10=1700 yrs of integration

Table 2. Other hindcast experiments (if resources permit)

#	Experiment	Notes	# of years
TIER 3: Effects of increased ensemble size			
A3	Increase ensemble size for the Tier 1 Experiment	*m* additional ensemble members to improve skill and examine dependence of skill on ensemble size	60x(5-10)x*m*=(300-600)*m* years of integration
TIER 4: Improved estimates of hindcast skill			
A4	Ensembles of at least 5-year, but much preferably 10-year, hindcasts and forecasts	As Experiment 1 but with no information from the future with respect to the forecast Radiative and other forcing information (e.g. greenhouse gas concentrations, aerosols etc.) maintained at initial state value or projected in a simple way. No inclusion of volcano or other short term forcing unless available at initial time.	1500-6000 years of integration
TIER 4: Improved estimates of the effects of initialization			
A5	Ensembles of at least 5-year, but much preferably 10-year, hindcasts and forecasts	Historical climate simulations up to the start dates of corresponding forecast with prescribed forcing Simulations continued from forecast start date but with the same forcing as in the Tier 1 Experiment, i.e. with NO information from the future with respect to the start date. These are uninitialized versions of Experiment 4 hindcasts.	1500-6000 yrs of integration

Explanatory comments

Table 1 lists the main DCPP Component A experiments. The Tier 1 hindcast experiment parallels the corresponding CMIP5 decadal prediction experiment in using the same specified forcing during the forecasts as is used for the Tier 2 historical climate simulations (which correspond to the CMIP6/historical simulations). The specification of historical and scenario forcing introduces some information from the future with respect to the forecast and may lead to slightly overestimated historical forecast skill measures. The main effect is expected to be due to the specification of short term radiative forcings such as volcanoes which occur during a forecast. Other forcings, such as those associated with greenhouse gas and aerosol emissions and/or concentrations, vary comparatively slowly over the five or ten year period of a forecast so affect the results very little. The benefits of using specified forcings include the use of common values across models, the ease of treatment within models, the possibility of documenting improvements with respect to CMIP5 hindcasts, the ability to estimate the effects of initialization by comparing forecasts and simulations which use the same forcings, and the estimation of drift corrections from hindcasts which include the forcings and so are more suitable for the purpose of future decadal forecasts.

Component A benefits from and builds on the experience gained from the decadal component of CMIP5. The DCPP Component A calls for hindcasts every year, rather than every 5 years which will improve the statistical stability of results, allow more sophisticated drift treatments, more clearly delineate skill levels, and foster improved assessment, combination, and calibration of the forecasts. Broad participation in Component A will potentially allow stratification of results according to the inclusion and initialization of climate components in the models, of model resolutions including model tops, and of methods of initialization and ensemble generation.

DCPP component A also provides an opportunity to study solar effects on climate. In order to take advantage of this, however, groups should use the correct ozone forcing timeseries which is important for the impact of solar variations.

Table 2 lists additional experiments which are of interest if resources permit. The Tier 3 experiment increases the ensemble size in order to quantify the expected benefits and as a guide to future forecast applications. It is not expected that many, if any, groups will undertake the Tier 4 experiments which require an additional large commitment of resources. These experiments are of interest in order to quantify the effects of specifying forcing during the forecast period and are included for completeness and in case the needed resources become available.

DECK and DECK/historical simulations. The DCPP is unique in bringing together researchers from communities with expertise in seasonal to interannual prediction as well as climate simulation. For climate models, control and sensitivity experiments are a backdrop to climate change simulations and most models used in the DCPP will also participate in other aspects of CMIP6 and will have performed DECK and 20th century climate change integrations. The DCPP strongly encourages participants to perform the DECK simulations but recognizes that this may not be feasible for all groups (those proposing to use high resolution models for prediction for instance). It is not intended that the DECK requirements should bar DCPP participation in these special cases.

Data retention. See the file (DCPP_CMIP6DataRequestCompilationTemplate.xlsx) for DCPP input to the CMIP6 Data Request. These requests apply to all Tiers. Data are to be served via the Earth System Grid (ESG) and to parallel CMIP5 although with changes to protocols as specified by the WGCM Infrastructure Panel (WIP). At this time, 6-hourly decadal prediction data for dynamical downscaling are not considered a priority. The hope is that, in conjunction with the WIP, a coordinated set of “basic” or “common” tiered data tables can be developed across MIPs together with “MIP specific” tables associated with individual MIPs.

DCPP Component B
Experimental real-time multi-model decadal predictions

The real-time decadal prediction component of CMIP will follow the example of other coordinated experiments as a protocol-driven multi-model multi-national project with data production and data sharing as integral components. It will build on the WMO structure already in place for seasonal forecasts. Forecasts and verification statistics will be made available via the web at WMO designated "Lead Centres" and mirrored via the ESGF. Lead Centres will collect forecast and verification data from designated "Contributing Centres". Lead Centres will produce a multi-model forecast together with uncertainties, and maintain an archive of previous real-time forecasts from Contributing Centres along with an assessment of performance as verifying observations become available.

Goals

- the promotion of the science and practice of decadal prediction by generating *real-time* multi-model decadal predictions
- the production and retention of ongoing multi-year multi-model decadal forecast data in support of the Global Framework for Climate Services (GFCS)
- the provision of information potentially useful for the IPCC WG1 AR6 assessment report and other studies and reports on climate prediction and evolution

Scientific aspects

- assess decadal predictions of key variables including temperature, precipitation, mean sea level pressure, the AMO, PDO, Arctic sea ice, the NAO, and tropical storms
- assess uncertainties and generate a consensus forecast
- permit the assessment decadal predictions of associated climate impacts of societal relevance

Practical aspects

- the coordination of efforts based on agreed experimental structures and timelines as specified in the protocol below
- the contribution to the development of infrastructure, in particular a multi-purpose data archive of ongoing decadal forecasts useful for a broad range of scientific and application questions and of benefit to national and international climate prediction and climate services organizations

The basic elements

- an ongoing coordinated set of multi-model multi-member ensembles of real-time forecasts updated each year.
- an associated hierarchy of data sets of results generally and readily available to the scientific and applications communities

Details of the DCPP Component B real-time decadal prediction component are listed below**.**

DCPP/WMO/CMIP Real-time decadal forecast protocols

Table 1. **Basic Component B: Real-time decadal forecasts**

#	Experiment	Notes	# of years
TIER 1: Real-time forecasts			
B1	Ensembles of ongoing real-time 5-year forecasts	Coupled models with initialization based on observations Start date *every year* ongoing Start date on or before 31 Dec (start dates on or before Nov 15 allow for DJF seasonal forecast results and are recommended) 10 ensemble members (more if possible) Atmospheric composition and/or emissions (and other conditions including volcanic aerosols) to follow a prescribed forcing scenario as in A1.	10x5=50 years of integration for 5-year forecasts
TIER 2: Increased ensemble size and duration			
B2.1	Increase ensemble size	*m* additional ensemble members to reduce noise and improve skill	5*m* yrs of integration
B2.2	Extend forecast duration to 10 years	To provide forecast information for the period 5 to10 years ahead	10x5=50 yrs of integration

Table 2. Component B Data

Because of its “quasi-real time” aspect, the data aspects of Component B differ somewhat from those of Components A and C.

Data to be served via WMO Lead Centres and mirrored on the Earth System Grid (ESG) with protocols paralleling CMIP5 although with modifications as specified by the WGCM Infrastructure Panel (WIP). Data to be archived by March 31st each year.

Priority	Description	Notes
Priority 1 - monthly means - basic variables - single level files	- surface air temperature, precipitation, mean sea level pressure, sea-ice, snow, 500hPa geopotential height, 850hPa temperature - vertically integrated amounts of energy, salt in the ocean - Atlantic MOC - fluxes of energy and moisture at the TOA and surface	Basic data sets for many investigations. Applies to quasi-real time decadal predictions currently being made.
Priority 2 **-** hindcast data for skill assessment and forecast calibration	- Same variables as for Priority 1	Hindcast data for models which have contributed to the multi-model prediction exercise since CMIP5
Priority as in the DCPP Data Retention Table	- Variables as in the DCPP Data Retention Table	More extensive data for forecast production, research and applications. Ongoing upon the completion of Component A hindcasts.

Explanatory comment

Component B real-time decadal forecasts are currently being produced based on CMIP5 and other models and hindcast data sets. The intent is that the forecasts produced by these models will be augmented and/or replaced by Component A results as they become available**.**

DCPP Component C: Predictability, Mechanisms and Case Studies

The climate system varies on multiple timescales which may be studied using physically based and statistical models. Diagnostic studies investigate climate system behaviour inferred indirectly from a long series of observations and/or model simulations. Prognostic studies investigate the behaviour of models when initial conditions or model features such as physical parameterizations, numerics or forcings are perturbed. The mechanisms involved are of great interest as they underpin the inherent predictability of the system and as they govern forecast skill.

Predictability studies based on models are often referred to as “perfect model” studies in the sense that one has perfect knowledge of the modelled climate system in terms of the computer code. However, they represent “attainable predictability” only to the extent that the model is sufficiently similar to the real system and it is important also to study their applicability. Predictability studies are intended to give an indication of the regions and timescales for which skilful forecasts may be possible and may also be used to study aspects of the physical mechanisms and processes involved. Predictability information can be inferred from the output of ensembles of decadal predictions as well as from independent predictability experiments.

Case studies are hindcasts which focus on a particular climatic event and the mechanisms and impacts involved. These are typically hindcast studies of an observed event although they can include particular kinds of events in model integrations (variations of AMOC and the associated variation of N Atlantic SSTs in models are an example). Studies of the skill with which a particular event (e.g. the hiatus, climate shift, an extreme year, etc.) can be forecast and the mechanisms which support (or perhaps make difficult) a skilful prediction are all of interest.

The DCPP and CLIVAR are proposing coordinated multi-model investigations of a restricted number of mechanism/predictability/case studies believed to be of broad interest to the community. Two research areas are the current foci of Component C. They are:

- Hiatus+: an investigation of the origin, mechanisms and predictability of long timescale variations in global mean temperature (and other variables) including periods of both enhanced warming and cooling with a focus on the current “hiatus”
- Volcanoes and prediction: an investigation of the influence and consequences of volcanic eruptions on decadal prediction and predictability

A description of the proposed experiments is found below.

The proposed experiments in Table 1 are intended to discover how models respond to simple imposed forcings in the tropical Pacific and in the North Atlantic. The questions at issue are the consistency of models’ responses to the forcings and the pathways through which the imposed forcings are expressed throughout the ocean and atmosphere, especially as this illuminates model behaviour and possible mechanistic links to retarded and accelerated global temperature variations and regional climate anomalies. In other words, to what extent can we attribute modulations of global mean temperatures to ocean variations, what are the respective roles of the Eastern Pacific and North Atlantic, and to what extent can we attribute decadal climate anomalies at regional scale to the AMV and IPV. Are AMV events a precursor of IPV shifts and vice versa and what are the mechanisms at play?

The proposed experiments in Table 2 investigate the predictability of the mid-1990s warming of the subpolar gyre, and its impact on climate variability.

The proposed experiments in Table 3 are directed toward an understanding of the effects of volcanoes on past and potentially on future decadal predictions. Radiative effects arising from the aerosol loading in the stratosphere, together with subsequent dynamical effects and/or coupled dynamical modes, are of interest.

Participants are invited to undertake as many of the Component C experiments as are of interest to them. **Please see** the Notes at the end of the section for additional details of the Component C experimental protocol.

Table 1.

Component C1: Haitus+: Accelerated and retarded rates of global temperature change and associated regional climate variations

Objectives: To investigate the role of eastern Pacific and North Atlantic sea surface temperatures in the modulation of global surface temperature trends and in driving regional climate variations.

#	TIER	Experiment	Notes	# of years
Pacemaker experiments				
C1.1	1	Coupled model restored to observed anomalies of sea surface temperature in the tropical eastern Pacific	-Follow the experimental design of Kosaka and Xie (2013). -Time period: 1950 to 2014 (from 1920 if possible) -Ensemble size: 10 members or more. -Restoring timescales: 40 $Wm^{-2}K^{-1}$ for a 50m deep mixed layer (or equivalent) -Monthly SST anomalies (base period 1950-2014) will be provided	65x10=650 years
C1.2	1	As above but for the North Atlantic	As C1.1 but restored to 12-month running mean SST anomalies (to be provided) in the North Atlantic, 10^oN to 60^oN -Time period: 1950 to 2014 -Ensemble size: 10 members (25 preferable) -Restoring timescales and ensemble generation: as for C1.1 -Minimization of drift if necessary (see Notes)	65x10=650 years

C1.3	2	As C1.2	As C1.2 but with restoring only in the extratropics 45-60^{o}N	65x10=650 years
C1.4	2	As C1.2	As C1.2 but with restoring in the tropical band 10-35^{o}N	65x10=650 years
C1.5	2	Control experiment	Restore N Atlantic SST to model control run climatology -Time period: 10 years -Ensemble size: 25 members -Restoring timescales and ensemble generation: as for C1.1 -No interannual changes in external forcings	25x10=250
C1.6	2	Climate impacts of AMV+	As C1.5 but restore N Atlantic SST to positive AMV anomaly (provided) superimposed on model climatology	25x10=250
C1.7	2	Climate impacts of AMV-	As C1.6 but for negative AMV anomaly pattern	25x10=250

Table 2.

Component C2: Case study of mid-1990s Atlantic subpolar gyre warming

Objectives: To investigate the predictability of the mid-1990s warming of the subpolar gyre, and its impact on climate variability.

#	TIER	Experiment	Notes	# of years
			Prediction experiments	
C2.1	3	Repeat hindcasts with altered initial conditions	Initialize with climatology (the average over 1960 to 2009) in N Atlantic “sub-polar ocean”[95^o W to 30^o E, 45^o N-90^o N] -Linear transition between climatology and actual observations over the 10^o buffer zone 35^o N-45^o N - 10 member ensembles - 5, but much preferably 10 years - Start dates end of 1993, 1994, 1995, 1996	4x(5,10)x10=200-400 years
C2.2	3	ditto	As above with start dates 1992, 1997, 1998, 1999	200-400 years

Table 3.

Component C3: Volcano effects on decadal prediction

Objectives:

- Assess the impact of volcanoes on decadal prediction skill
- Investigate the potential effects of a volcanic eruption on forecasts of the coming decade
- Investigate the sensitivity of volcanic response to the state of the climate system

#	TIER	Experiment	Notes	# of years
Prediction experiments with and without volcano forcing				
C3.1	1	Pinatubo	Repeat 1991 forecasts without Pinatubo forcing -5, preferably 10 years -10 ensemble members -Specify the "background" volcanic aerosol to be the same as that used in the 2015 forecast	(5,10)x10=50-100 years
C3.2	2	El Chichon	1982 hindcasts as above but without El Chichon forcing	50-100 years
C3.3	2	Agung	1963 hindcasts as above but without Agung forcing	50-100 years
#	**TIER**	**Experiment**	**Notes**	**# of years**
Prediction experiments for 2015 with added forcing				
C3.4	1	Added forcing	Repeat 2015-2019/24 forecast with Pinatubo forcing	50-100 years
C3.5	3	Added forcing	Repeat 2015-2019/24 forecast with El Chichon forcing	50-100 years
C3.6	3	Added forcing	Repeat 2015-2019/24 forecast with Agung forcing	50-100 years

Notes

Experiment C1.1 follows the design of Kosaka and Xie (2013) in which observed SST anomalies are imposed in the tropical Pacific region in coupled model simulations. The results will be compared to the standard historical simulations to infer the impact of the tropical Pacific SSTs. In particular:

- Observed monthly SST anomalies (base period 1950-2014) (to be supplied) are superimposed on the model climatology over the same period computed from historical simulations in order to avoid model drift..

- The SST signal is imposed either by altering surface fluxes or by restoring the SST directly. A restoring coefficient equivalent to 40 $Wm^{-2}K^{-1}$ for a 50m mixed layer is recommended. To minimise shocks, the restoring coefficient should decrease to zero over an 8^o buffer zone bounding the restored region. No restoring if sea ice present. Outside of the restoring region the model evolves freely allowing full climate system response.

- In order to sample uncertainties in the ocean initial state it is recommended that, if possible, ensemble members are generated by taking initial conditions from different historical simulations. Otherwise, ensembles may be generated by perturbing atmospheric conditions.
- Experiments should cover the period from 1950 to 2014, but starting from 1920 is desirable if possible. External forcings as for historical simulations.

Experiment C1.2 follows experiment C1.1 but uses observed SST anomalies in the North Atlantic rather than the tropical Pacific. Tests have shown that SST restoring in the Atlantic may lead to an undesirable response of the Atlantic Meridional Overturning Circulation (AMOC) and hence of SSTs in other regions (including the south Atlantic) which can obscure the results. It is recommended that groups monitor this potential response and take steps to minimise it if necessary. One way to minimise the AMOC response is to apply additional salinity restoring such that the upper ocean density remains unchanged. Another approach is to perform 3D restoring of temperature and salinity below the mixed layer. There is evidence that the signal to noise of the atmospheric response to North Atlantic SST is comparatively weak in models and additional ensemble members (up to 25) are requested if possible.

Experiments C1.3 and C1.4 follow experiment C1.2 but apply the SST anomalies in the subpolar and subtropical regions separately in order to assess their relative importance.

Experiments C1.5-1.7 are similar to experiment C1.2 but use an idealised SST pattern (which will be supplied). This pattern is derived from the difference between observations and the ensemble mean of coupled model historical simulations (Ting et al 2009) and is an estimate of unforced internal variability. Although this estimate is not perfect, since the modelled response to external factors such as anthropogenic aerosols may not be entirely correct, the experiments nevertheless provide information on the climate response to North Atlantic SST variations. Results are based on model control integrations rather than historical simulations and therefore may be performed before the updated CMIP6 forcings become available.

Kosaka, K. and S-P. Xie (2013) Recent global-warming hiatus tied to equatorial Pacific surface cooling, Nature, 501, 403–407

Ting M, Kushnir Y, Seager R, Li C (2009) Forced and internal twentieth-century SST in the North Atlantic, J Clim 22: 1469-1881, doi: 10.1175/2008JCLI2561.1

The Decadal Climate Prediction Project (DCPP)

We respond below to the various questions and comments in the email. The Aspen Workshop was very successful in reviewing activities and in improving the details of Component C. We believe that the DCPP is an important contribution to CMIP6 and attempt to respond to the points you raise below:

1. "It is probably possible to run Component A Tier 1 completely disconnected of the DECK and the CMIP6 Historical Simulation. Would DCPP consider switching round the Tiers in Component A?"

It would not be appropriate to exchange Component A Tier 1 and Tier 2 in our view. Tier 1 is the basic "prediction/predictability" experiment of the DCPP without which we would learn very little about our capability to make decadal predictions. We agree that it is not directly dependent on the DECK and CMIP6/historical, although we believe that substantially all groups will perform the DECK experiments. However, we also argue that decadal prediction provides an additional and alternative way of viewing model and system behaviour that compliments the historical/future simulation approach that has traditionally been part of CMIP. The Component A Tier 1 "prediction" (rather than simulation) experiment brings new possibilities for investigating and understanding climate variation which are more direct and time-dependent than is possible in simulations. The strong and direct connection between observations and predictions is particularly pertinent.

2. "As a side note, you might have seen, we have sent out a related email to the model groups ... in the special case where the burden of the entry card simulations are prohibitive...an exception to this policy can be granted on a model to model basis..."

As far as we are aware, virtually all Component A participants will perform both DECK and CMIP6/historical simulations. It may be the case that one or two potentially participating groups, with high resolution models for instance, will not have the resources to undertake all of the DECK simulations. It would, in our view, be a mistake to exclude such groups from participating in Component A since they may prove to be better able to represent the physical processes governing short term climate variability and thereby to attain higher skill, than lower resolution models. It would certainly be of interest to see what difference resolution would make for instance.

We are concerned that if the DECK requirements are too rigid it will have a detrimental impact on the DCPP by preventing participation in Component A (you will recall the discussion in Offenbach for example). Our response to this is to urge DECK compliance but indicate that it is not a rigid requirement under all circumstances..

3. "For Tier 2 of Component A it is unclear why there is no mentioning of additional ensemble members of the CMIP6 Historical Simulation ....."

Component A Tier 2 calls for exactly this, namely an ensemble of 10 historical simulations (which are extended to 2030 to overlap with decadal predictions initialized in 2020). The connection with the CMIP6/historical simulations is mentioned in the Explanatory comments and has been added to the Table 1, Tier 2 heading.

4. "Experiment Spreadsheet"

Please see the updated Spreadsheet, attached.

5. "Experiment Design. The science rationale for each experiment is clear, but need to be further specified in the GMD paper....."

Following the Aspen Workshop, where details were extensively discussed and agreed upon, we have prepared expanded "Notes" as part of the Component C description. We will, of course, provide rationales and details also in the GMD paper.

6. "Component B only makes sense if Component A has been run (i.e., Component A should be an entry card for Component B)."

As noted in the project description, Component B real-time decadal forecasts are currently being produced based on CMIP5 models and hindcast data sets. The intent is that the forecasts produced by these models will be augmented, and subsequently replaced, by Component A results as they become available. In this way the real-time decadal forecasting activity will evolve and improve in a natural way.

As indicated in the description of Component B, there are exploratory efforts to align Component B with the WMO structure already in place for seasonal forecasts. The intent is ultimately to contribute to applications and the GFCS. Component B is notable in that it promotes both science and applications.

7. "... it is not always clear in the proposal how this new effort will build on achievements from CMIP5 and what is different or better....."

All of the points mentioned in this comment are pertinent and, it seemed to us, to be implicit in the preamble to the DCPP and Component A. We have expanded on this in the Component A Notes.

8. "Data request: DCPP should recommend that their list of data should also be collected from the CMIP6 Historical Simulation..."

We always intended this to be the case and now state it explicitly under "Data retention" that the data request applies to “all Tiers”.

The treatment of data is a major CMIP6 infrastructure feature and we are anxiously awaiting more information on how this will be handled.

9. "Scientific analysis plan: The scientific analysis plan is rather general ....".

The DCPP is a comparatively new area of research and is evolving rapidly. The broad questions at issue are the extent to which skillful decadal predictions are possible, the global patterns of currently available skill, and the identification of the mechanisms involved. Each of the Components is directed toward one or more of these broad questions.

One of the main outputs of the DCPP will be a data set which will encourage analysis, especially by the broader diagnostics community, of the array of questions that are of interest to decadal prediction (e.g. methods of initialization, ensemble generation, model drift and bias, etc. etc.). If DCPP data is readily available we know it will encourage many diverse and novel analyses of interest. We do not feel it would be profitable, at this stage, to restrict or direct these analyses, and we would rather count on scientific (free) enterprise to exploit DCPP results.

10. "It was not clear if DCPP is willing to contribute an article to the GMD CMIP6 special issue...."

We were not entirely clear as to where in the document this should appear. This is now mentioned explicitly under the “References and Publications” section of the application.

11. "... please also respond to any requests from Martin Juckes concerning information needed to prepare the data request...."

We have responded to Juckes' requests and stand ready to respond to follow up. The CMIP6 endorsement "Criterion 3" represents a major responsibility for modelling groups but also for CMIP. Our recent interactions with the ESGF and WIP websites have not been promising and we await further clarification of the critical technical aspects around data retention and distribution.

12. We thank the Panel for providing such detailed comments and feedback. We are pleased that the Panel finds the DCPP proposal pertinent and scientifically interesting and hope that these responses to your comments and queries are helpful.

Best regards

GJB and DS

Flux-anomaly-forced model intercomparison experiment (FAFMIP)

Steering committee

Jonathan Gregory (chair, j.m.gregory@reading.ac.uk), Detlef Stammer (detlef.stammer@zmaw.de), Stephen Griffies (stephen.griffies@noaa.gov)

Goals and overview of experiments

FAFMIP is proposed in support of the WCRP Grand Challenge on sea level rise and regional impacts. Projections of regional sea level change by CMIP5 AOGCMs, like earlier AOGCM generations, show a substantial spread due to the different models' differing simulations of regional ocean density and circulation changes, especially in high latitudes and the North Atlantic (Yin, 2012, 10.1029/2012GL052947; Bouttes et al., 2012, 10.1029/2012GL054207; IPCC AR5 WG1 chapter 13, Church et al., 2013; Slangen et al, 2014, 10.1007/s10584-014-1080-9). By applying flux perturbations from a range of CMIP5 models to the same AOGCM, previous analyses have shown that a substantial fraction, but not all, of the diversity of sea level projections arises from the spread in AOGCM projections of changes in surface fluxes of momentum (windstress), heat and freshwater (Bouttes et al., 2012, cited above; Bouttes et al., 2014, 10.1007/s00382-013-1973-8; Bouttes and Gregory, 2014, 10.1088/1748-9326/9/3/034004).

In the FAFMIP experiments, a prescribed set of surface flux perturbations will be applied to the ocean. These perturbations will be obtained from the ensemble-mean changes simulated at time time of doubled CO_2 by CMIP5 AOGCMs under the 1pctCO2 scenario, so they are representative of projected anthropogenic climate change. The aims of the experiments are:

- to quantify the difference in the geographical patterns of sea level change due to ocean density and circulation change simulated by the models, when given common surface flux perturbations.
- to provide information about the efficiency and interior distribution of ocean heat uptake in response to climate change; the AOGCM spread in these phenomena contributes to their spread in transient climate response and global mean sea level rise due to thermal expansion.
- to provide information about the sensitivity of the Atlantic meridional overturning circulation (AMOC) to prescribed buoyancy forcing of the character expected for CO_2 forcing, rather than idealised freshwater forcing such as has been used in previous AMOC intercomparisons; change in the AMOC is of relevance to both regional and global sea level rise, as well as to regional climate change.

The FAFMIP experiments are aimed at increased physical understanding. They are not themselves policy-relevant scenarios, but obviously the uncertainties in projection of global and regional sea level and AMOC change are of great policy relevance.

The steering committee undertakes to ensure that a paper on the FAFMIP design will be prepared, and all participants will be encouraged to collaborate in producing a paper on the results. At the time of writing (20 May 2015) there are ten groups who plan to run FAFMIP experiments (ACCESS, CanESM, CNRM/CERFACS, GFDL, GISS, IPSL, MIROC6, MPI, MRI, UKESM).

Design of experiments (see http://www.met.reading.ac.uk/~jonathan/FAFMIP)

All the experiments will add anomalies to the surface fluxes computed by the AOGCM (like a flux adjustment). The fluxes themselves will not be replaced because this would typically cause a very large climate drift and possible instability, and is technically more complicated than adding an anomalous flux. The surface flux anomalies are a function of (longitude, latitude, time of year) and constant throughout the experiments, which are proposed to be 70 years long (but shorter experiments would still be useful if 70 years cannot be afforded). The experiments will branch from and be analysed by comparison with the standard CMIP DECK pre-industrial control. All the experiments have pre-industrial atmospheric conditions.

There are three tier-1 experiments, most important first. The **bold** word is the name of the experiment.

stressFAF: Impose a perturbation in surface zonal and meridional windstress. We propose this experiment first because the windstress change appears to have the largest effect on sea level in CMIP5 scenario experiments. In addition to its relevance to sea level, this experiment will also be of interest regarding the phenomenon of eddy saturation (relative insensitivity of the circumpolar circulation to windstress change), especially in eddy-resolving models, and to study the influence of windstress change on advecting circumpolar deep water towards the Antarctic continental shelf, where it could affect ice-shelf melting and hence sea-level rise through the effect on ice-sheet dynamics (a different aspect of the Grand Challenge). The perturbation is made to windstress, rather than to wind speed in the atmosphere, because windstress is the flux experienced by the ocean. AOGCMs typically use other diagnostics of wind speed to supply turbulent mixing energy to the ocean in addition to windstress. Perturbing these quantities is not included in the proposed design at present.

heatFAF: Impose a perturbation in surface heat flux, which is second in importance in its influence on patterns of sea level change. It has also been found in a previous analyses to be the main influence on AMOC change. In an AOGCM, imposing a heat flux perturbation is not straightforward, because it alters the SST, which affects the surface heat flux calculated by the atmosphere model and tends to cancel out the perturbation. In this experiment and in the **allFAF** experiment, we propose to use a passive tracer to avoid this feedback (see documents on website). The design allows us to distinguish the effects of added heat and redistribution of the control heat content.

waterFAF: Impose a perturbation in the surface freshwater flux (including the contribution from runoff change). This is the least influential surface flux.

There are two tier-2 experiments.

passiveheat: Add a surface flux of passive tracer at the same rate as the surface heat flux perturbation in the **heatFAF** experiment. This flux will be added to the top layer of a passive temperature tracer. Comparison of this experiment with the **heatFAF** experiment will allow the effect of ocean advection on surface heat flux feedback to be assessed (cf. Winton et al., 2013, 10.1175/JCLI-D-12-00296.1).

allFAF: In the **allFAF** experiment, the anomalous fluxes of windstress, heat and water are simultaneously applied, using the passive-tracer method for heat as in the **heatFAF** experiment.

Diagnostics

No changes to the standard CMIP set of diagnostics or CF, CMOR or ESG are required. The analysis of sea level change will mainly use zos, zostoga, thetao and so. Analyses of ocean heat uptake efficiency will use thetao. Analyses of the AMOC will use msftmyz, msftyyz, uo and vo. It is strongly recommended that 3D ocean diagnostics should be implemented for monthly-mean temperature and salinity tendencies ($\partial T/\partial t$ and $\partial S/\partial t$) due to the various physical processes which modify the state (advection, diffusion, etc.). These diagnostics have been included in Table 2.9 of the recommendations from the CLIVAR Ocean Model Development Panel committee on CMIP6 ocean model output for use in all CMIP6 experiments, including DECK for instance, but their usefulness for FAFMIP is particularly noted there. If the $\partial/\partial t$ diagnostics are not submitted for all experiments, for FAFMIP they are particularly requested for the portion of the DECK piControl which is parallel to the FAFMIP experiments, and for the DECK idealised climate change experiments abrupt4xCO2 and 1pctCO2.

Proposed timing

The input fields will be prepared by the end of June 2015 and experiments can be done thereafter by any interested groups. Interim versions are currently being tested by three groups.

The Geoengineering Model Intercomparison Project Phase 6 (GeoMIP6): Simulation Design and Preliminary Results

Ben Kravitz,[1*] Alan Robock,[2] Simone Tilmes,[3] Olivier Boucher[4], Jason M. English[5,6], Peter J. Irvine[7], Andy Jones[8], Mark G. Lawrence[7], Michael MacCracken[9], Helene Muri[10], John C. Moore[11], Ulrike Niemeier[12], Steven J. Phipps[13], Jana Sillmann[14], Trude Storelvmo[15], Hailong Wang[1], Shingo Watanabe[16]

In preparation for *Geoscientific Model Development*

[1]Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, WA, USA
[2]Department of Environmental Sciences, Rutgers University, New Brunswick, NJ, USA
[3]National Center for Atmospheric Research, Boulder, CO, USA
[4]Laboratoire de Météorologie Dynamique, IPSL, CNRS/UPMC, Paris, France
[5]NOAA Earth System Research Laboratory, Boulder, CO, USA
[6]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA
[7]Institute for Advanced Sustainability Studies, Potsdam, Germany
[8]Met Office Hadley Centre, Exeter, United Kingdom
[9]Climate Institute, Washington, D.C., USA
[10]Department of Geosciences, University of Oslo, Oslo, Norway
[11]State Key Laboratory of Earth Surface Processes and Resource Ecology, College of Global Change and Earth System Science, Beijing Normal University, Beijing, China
[12]Max Planck Institute for Meteorology, Hamburg, Germany
[13]ARC Centre of Excellence for Climate System Science and Climate Change Research Centre, University of New South Wales, Sydney, Australia
[14]Center for International Climate and Environmental Research, Oslo, Norway
[15]Department of Geology & Geophysics, Yale University, New Haven, CT, USA
[16]Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan

*To whom correspondence should be addressed: P.O. Box 999, MSIN K9-24, Richland, WA 99352, USA (ben.kravitz@pnnl.gov)

Abstract. We present a suite of new climate model experiment designs for the Geoengineering Model Intercomparison Project (GeoMIP). This set of experiments, named GeoMIP6 (to be consistent with the Coupled Model Intercomparison Project Phase 6), builds on the previous GeoMIP project simulations, and has been expanded to address several further important topics, including key uncertainties in extreme events, the use of geoengineering as part of a portfolio of responses to climate change, and the relatively new idea of cirrus cloud thinning to allow more longwave radiation to escape to space. We discuss experiment designs, as well as the rationale for those designs, showing preliminary results from individual models when available. We also introduce a new feature, called the GeoMIP Testbed, which provides a platform for simulations that will be performed with a few models and subsequently assessed to determine whether the proposed experiment designs will be adopted as core (Tier 1) GeoMIP experiments. This is meant to encourage various stakeholders to propose new targeted experiments that address their key open science questions, with the goal of making GeoMIP more relevant to a broader set of communities.

1. Introduction

As anthropogenic climate change continues largely unabated, society is exploring research into options for addressing the effects of greenhouse gas emissions. Along with mitigation and adaptation, a further option that is under consideration is solar radiation management (SRM). SRM involves deliberate modification of the climate system to offset the radiative effects of increasing anthropogenic greenhouse gases by either increasing the reflection of solar radiation back to space or increasing the outgoing flux of terrestrial radiation. SRM is sometimes considered under the larger umbrella of "geoengineering" or "climate engineering", which also includes proposals for carbon dioxide removal (CDR). In this paper we will use the term "geoengineering", in the context of the Geoengineering Model Intercomparison Project (GeoMIP), to specifically refer to the broad range of proposed SRM techniques. Better understanding the potential role that SRM might have in addressing climate change requires research on the climate effects and impacts, as well as the underlying processes involved and their uncertainties.

The goal of GeoMIP is to understand the robust climate model responses to geoengineering (Kravitz et al., 2011a). So far, there have been seven core climate model experiments designed for analyzing the effects of solar irradiance reduction, an increase in the loading of stratospheric sulfate aerosols, and marine cloud (or sky) brightening (Kravitz et al., 2011a, 2013a), as well as several additional experiments proposed by various groups. Table 1 lists all of the proposed experiments to date. GeoMIP has achieved success on a number of fronts: fifteen modeling groups have participated in one or more experiments. As of the writing of this paper, GeoMIP has resulted in 23 peer-reviewed publications; and results from GeoMIP were featured in the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (Boucher et al., 2013), the recent National Academy of Sciences report on SRM (NAS, 2015), and the final report from the European Transdisciplinary Assessment of Climate Engineering (EuTRACE; Schäfer et al., 2015).

These past efforts targeted specific areas. However, they were not designed to answer all questions about the potential climate effects of geoengineering, including questions about geoengineering methods that have been proposed, and remaining unanswered questions about conduct and design of research activities (Schäfer et al., 2015). The Coupled Model Intercomparison Project is beginning its sixth phase (CMIP6), and one of its focus areas is geoengineering (Meehl et al., 2014). Now is an opportune moment to address some of the key uncertainties regarding geoengineering by introducing designs for a new suite of climate modeling experiments. Pressing questions we propose to address include:

1. How would geoengineering affect changes in less easily detectable climate features, such as extreme events, modes of natural variability, regional impacts, and long timescale processes?
2. Cirrus cloud thinning is a newly proposed geoengineering method. What are the common responses in its simulation?

3. How would the climate response to geoengineering differ if it were used to slow rather than halt climate change? That is, what are common responses in climate models if geoengineering were to be used to only partially offset climate change?
4. What are robust differences in the climate model response between stratospheric sulfate aerosol injection and solar irradiance reduction?

In this paper, we outline four Tier 1 experiments for the next phase of GeoMIP, which, to be consistent with the numbering convention of CMIP, we call GeoMIP6. The experiment design for GeoMIP6 is based on discussions held at the Fourth GeoMIP Workshop (Paris, April 2014; Kravitz et al., 2014a), the SCRiM All Hands Meeting (State College, May 2014), and the Exploring the Potential and Side Effects of Climate Engineering (EXPECT) workshop (Oslo, June 2014), as well as an experiment proposed for inclusion in the Chemistry Climate Model Initiative (CCMI; Tilmes et al., 2014). All of the proposed experiments are listed in Table 1 along with all previous GeoMIP and GeoMIP-affiliated experiments.

The guiding science questions in GeoMIP6 are directly relevant to the core questions of CMIP6. Geoengineering simulations have repeatedly been shown to be a novel method of uncovering fundamental climate behavior (e.g., Kleidon et al., 2015; Kravitz et al., 2013b), and continue to be relevant for addressing the question, "How does the Earth System respond to forcing?" Experiment G1 has already proven quite useful in this regard, particularly in its ability to separate mechanistic changes that contribute to the fast and slow responses of the climate system (e.g., Kravitz et al., 2013b; Tilmes et al., 2013); G1ext will likely provide even more information about mechanistic changes in the climate system slow response. Experiments G6sulfur and G6solar (below) will provide a useful multi-model comparison of the Earth System response to different forcing agents in a controlled protocol. GeoMIP has also been successful in identifying both model commonalities and the effects of different stratospheric aerosol parameterizations on the climate effects of geoengineering (e.g., Berdahl et al., 2014; Yu et al., 2015). These efforts are continuing for sea spray geoengineering experiments (Kravitz et al., 2013a). Our experimental design, particularly for G6sulfur and G7cirrus (below), will aid in uncovering the origins and consequences of different model parameterizations and how they contribute to model biases. Geoengineering simulations have been shown to actually reduce certain aspects of climate uncertainty and sources of model bias (Kravitz et al., 2013c; MacMartin et al., submitted). As such, we see our efforts as highly synergistic with those of CMIP6, potentially providing relevant information to the driving science questions via relatively underexplored means.

2. Tier 1 experiments in GeoMIP6

In this section, we outline the four Tier 1 experiments that are proposed for GeoMIP6. These same experiments have also been proposed for inclusion in CMIP6, with GeoMIP serving as an officially endorsed model intercomparison project.

The general experimental protocol is somewhat different from that of the previous experiments (Kravitz et al., 2011a; Kravitz et al., 2013a; also see Table 1). There has recently been interest in conducting geoengineering studies that examine phenomena for which previous experiments have generated only a low signal-to-noise ratio: for example, extreme temperature and precipitation events (Curry et al., 2014) and modes of internal variability (Gabriel and Robock, 2015). To obtain more robust estimates of potential changes in extreme events and regional climate, we are now requesting that all simulations be conducted for longer than 50 years. Cessation or termination (in which the background scenario continues, but geoengineering is no longer conducted) is no longer part of the experimental protocol. Many of the broad messages associated with the so-called termination effect were well captured by Jones et al. (2013), so additional efforts to represent termination are not currently a high priority.

The monthly average output requested for each experiment should be the same as is requested for the core CMIP6 experiments (see below). In addition, we request that all modeling groups produce the following at daily frequency: minimum and maximum near-surface air temperature (reference height; usually 1.5-2 m), total surface precipitation, surface convective precipitation, near-surface (usually 10 m) wind speed, and hourly surface ozone concentration, if available. If possible, precipitation and convective precipitation should be reported as a cumulative value at 6-hourly frequency, and wind speed should be reported as an instantaneous value at 6-hourly frequency. Each modeling group should produce a minimum of three ensemble members for each experiment; ideally, groups would complete five or more ensemble members.

As before, the Tier 1 experiments will be based on core experiments in CMIP. The newest version of the core CMIP6 experiments is called the CMIP Diagnostic, Evaluation and Characterization of Klima (DECK) experiment portfolio (Meehl et al., 2014). This will include many different simulations, but the DECK simulations that are relevant for GeoMIP6 are piControl, historical, and abrupt4xCO2, each of which was also included in CMIP5. Additionally, simulations involving future projections of climate change scenarios will be based on the Tier 1 simulations of ScenarioMIP (O'Neill et al., 2014). Tier 1 of ScenarioMIP will consist of high, medium, and low forcing scenarios, referring to the magnitude of anthropogenic radiative forcing applied in that scenario.

2.1. G1ext

This experiment is planned as an extended version of Experiment G1 (*Kravitz et al.*, 2011a). G1ext proposes that, beginning from a preindustrial simulation (piControl), the net top of atmosphere (TOA) radiative flux imbalance due to an abrupt quadrupling of the CO_2 concentration (abrupt4xCO2) would be balanced via a reduction in total solar irradiance (Figure 1). Here, "balance" is defined as the global mean value top-of-atmosphere net radiative flux being within ±0.1 W m^{-2} of the piControl experiment over an average of years 1-10 of the simulation. The original G1 was conducted for 50 simulation years, so this will be a simple extension of the previous experiment. Modeling groups that have already moved on to a new model version, or for whatever reason are not able to extend their previous model run, should run experiment G1ext for the full 100 years with their new version.

G1 has proven quite successful in revealing the underlying climate behavior in response to solar irradiance reduction; it also received the highest participation of all GeoMIP experiments thus far. Most models have been modified since CMIP5, so evaluating climate response to G1 with the new model versions could serve as a useful comparison. A longer simulation will also improve the detection of changes in extreme events and modes of climate variability, particularly as related to regional changes. Moreover, some processes of interest, such as changes in ice sheet dynamics or deep ocean circulation, take longer than 50 years to resolve. Although 100 years is probably an insufficient length of time to fully assess changes in these fields, it may nevertheless allow enough time for an early indication of features that emerge above the noise level of the climate system; early detection will be aided by having multiple ensemble members.

G1ext will be highly synergistic with single-forcing experiments to be included in the Cloud Feedback Model Intercomparison Project (CFMIP), in which total solar irradiance is abruptly increased or decreased. Through comparisons between these CFMIP experiments and G1ext, we will be able to gain a better understanding of how the Earth System responds to radiative forcing. It will also reveal key information on the differences in cloud responses to single vs combined forcings, which has strong implications for diagnosing transient and equilibrium climate sensitivity.

G1 is the only original experiment from Kravitz et al. (2011a) that is proposed to be lengthened. The climate response in G2 is very similar to that of G1, but with a lower signal-to-noise ratio, so extending G2 is unlikely to provide substantial additional information. A new experiment (G6sulfur, below) has been proposed that will accomplish similar goals to G3, but without some of the ambiguities in experimental design that caused difficulties in interpreting results from G3 in certain cases (i.e., inconsistent experimental protocols; see Kravitz et al., 2011a, 2011b for further details). G4 may be extended in the future, but such a simulation is not a high priority at this time.

2.2. G6sulfur

Previous GeoMIP experiments (G3 and G4) used RCP4.5 as a background scenario. To maintain relevance to the newly designed experiments in CMIP6, our background scenario is changed to follow the ScenarioMIP Tier 1 scenarios, described above.

Under experiment G6sulfur (Figure 2), stratospheric sulfate aerosol precursors will be injected into the model with the goal of reducing the magnitude of the net anthropogenic radiative forcing from the ScenarioMIP Tier 1 high forcing scenario to match that of the ScenarioMIP Tier 1 medium forcing scenario (within ±0.1 W m^{-2}). The motivation for this choice is to evaluate a climate in which geoengineering is used to only partially offset climate change, which would hopefully reduce the burden of adaptation. The choice of the medium forcing scenario as the target, instead of the low forcing scenario (as in Section 4.1), is because the required amount of sulfate aerosol injection to achieve a low anthropogenic forcing is quite large. Representing such large values of injection in a variety of climate models will likely lead to highly variable inter-model results that are overly sensitive to individual parameterizations.

For this experiment, geoengineering will be simulated over years 2020-2100. All atmospheric constituents in the ScenarioMIP Tier 1 scenarios are well defined through the year 2100. Modeling groups that have an internal sulfate aerosol treatment should calibrate the radiative response to sulfate aerosols individually so that the results will be internally consistent. This procedure will be more difficult for models that have a complex microphysical treatment of the aerosols, which may require more sophisticated methods of meeting the goals of G6sulfur. One method to calculate the necessary amount of sulfate aerosol is a double radiation call, once with and once without the stratospheric aerosols. Another potential method involves using feedback methods (Jarvis and Leedal, 2012; Kravitz et al., 2014b; MacMartin et al., 2014). For models that have no dynamical treatment of sulfate aerosols, GeoMIP will provide a data set of aerosol optical depth, as well as ozone fields that are consistent with this aerosol distribution; these fields will be consistent with the fields generated for G4SSA (see Section 3.2 for further details). The amount of sulfate injection needed for a given model to achieve the goals of this experiment may vary, so modeling groups should scale the aerosol and ozone perturbation fields as necessary.

Of notable importance is that the lifecycle of stratospheric sulfate aerosols is very complex. To date, there are no comprehensive simulations of stratospheric sulfate aerosol geoengineering that include aerosol microphysical processes, explicit size representation, interactive chemistry, clouds, and radiation. Of the more comprehensive simulations conducted, some studies include aerosol microphysics and explicit size representation but do not allow oxidants to evolve (e.g., Heckendorn et al., 2009) or do not allow aerosol heating to interact with radiation and dynamics (e.g., English et al., 2012). Other studies include aerosol microphysics and heating, but represent the aerosol size distribution in assumed lognormal modes of prescribed constant width (e.g., Niemeier et al., 2011, 2013). Because geoengineering has not been conducted in the real world, there are no observations to constrain these particular physical processes in models. Kokkola et al. (2009) showed that even for volcanic eruptions, capturing the evolution of the aerosol size distribution is more difficult for larger amounts of stratospheric SO_2 injection. An additional complicating factor is that stratospheric aerosol geoengineering would be expected to modify the quasi-biennial oscillation (Aquila et al., 2014). This is important for the direct effects on circulation as well as the fact that the phase of the quasi-biennial oscillation would affect the rate of meridional transport of stratospheric aerosols (Plumb and Bell, 1982). Development of models that can represent these processes and thus constrain the uncertainties that may arise is ongoing, and we expect that substantial progress will be made by the time the GeoMIP6 experiments will begin. Nevertheless, the goal of GeoMIP is to use the best available models and attempt to characterize uncertainties introduced by structural uncertainties in those models.

All simulations will be conducted as if the aerosols or aerosol precursors are emitted in a line from 10°S to 10°N along a single longitude band (0°). This setup differs somewhat from a single point source injection in that it allows models with a strong stratospheric transport barrier to achieve a reasonable global distribution of sulfate aerosol rather than an aerosol optical depth maximum in the tropics. The size of the injection zone can substantially alter the resulting aerosol size distribution (English et al., 2012). However, we do not wish to add additional complications to the simulation design at this time, so our design does not strongly deviate from the design of a point-source injection. Injected aerosols or aerosol precursors should be evenly spread across model layers between 18 and 20 km. This is a slightly different setup from that of the original sulfate aerosol experiments (Kravitz et al., 2011a), but sedimentation processes and self-lofting due to heating are likely to result in the aerosols being distributed between 16-25 km in altitude, which is the specification of the original experiments. Models will use their own individual treatments of aerosol optical properties, as this would be too difficult to specify in a consistent way across all participating models.

2.3. G6solar

Experiment G3solar was proposed as an unofficial counterpart to experiment G3 (Kravitz et al., 2011a; Table 1); in G3solar, the goals of G3 were achieved using a solar irradiance reduction rather than stratospheric sulfate aerosol injections. Comparison of these two simulations would reveal differential effects of sulfate aerosols and solar irradiance reduction. Preliminary results from a limited set of models show some differences in the results of the two experiments, particularly related to the hydrological cycle response (Niemeier et al., 2013).

We propose G6solar as a parallel experiment to G6sulfur, to compare the effects of solar reduction with those of stratospheric aerosols. G6solar uses the same setup as G6sulfur, but geoengineering is performed using solar irradiance reduction (Figure 2). Because of the difficulties in setting up experiment G3, few groups performed either G3 or G3solar. The proposed G6solar is better specified than G3 and better aligned with the core simulations of CMIP, so it should garner substantially greater participation.

2.4. G7cirrus

A recent proposal in the geoengineering literature is the idea of seeding cirrus clouds, thinning them and thus allowing more longwave radiation to escape to space (Mitchell et al., 2009; Storelvmo et al., 2013). Encapsulated in this idea are two complementary areas of investigation: (1) the experimental design should capture the dominant effect of a drying of the upper troposphere (Muri et al., 2014), and (2) the experiment show allow for a determination of the effects on future climate response to geoengineering via cirrus thinning.

Because different models have different treatments of cirrus clouds, the description of the experimental design (below) consists of a simple treatment of cirrus clouds, allowing all models to simulate this experiment in the same way. Therefore, this experiment can be seen as assessing the spread of model response to a simple sensitivity test that mimics a proposed geoengineering technique. As such, this concept is directly relevant to answering questions about the sensitivity of ice clouds to perturbations, which directly impacts changes in convection, circulation, and ultimately climate sensitivity. In particular, by simulating this experiment in fully-coupled general circulation models, we can ascertain both how forced changes in high clouds affect circulation and the radiation budget and, in turn, how those effects feedback onto changes in high cloud coverage. This experiment will complement results obtained through CFMIP dealing with isolating the effects of cloud-radiation interactions in ice clouds.

The goal of cirrus seeding in the real world would be to cause cirrus clouds to consist of fewer but larger ice crystals, thus increasing the fall speed and so reducing the IR opacity of these clouds. A first attempt at representing the effects of cirrus cloud thinning was to multiply cirrus cloud optical depth in the radiation code by a factor $\varepsilon < 1$ without modifying the actual cirrus fields. However, this approach could be difficult in some models, as many models only distinguish between liquid and ice clouds, and the factor ε could only be implemented for ice clouds with temperature below -35°C and pressures lower than 600 hPa. Other models formulate the effects of cirrus clouds in the infrared as a modification to atmospheric emissivity, not optical depth.

Figure 3 shows results from GISS ModelE2 (Schmidt et al., 2014) for various values of ε. Global mean surface air temperature changes appear to be linear with ε, but the required cooling is not nearly substantial enough to achieve the goal of G7cirrus. We hypothesize that these results are due to cirrus clouds being very efficient absorbers of longwave radiation, even if they are optically thin. To achieve substantial cooling, it appears necessary to reduce cirrus cloud coverage, not just optical depth. Single model simulations of cirrus thinning that incorporate a treatment of cloud microphysics show more substantial surface cooling. Storelvmo and Herger (2014) found global cooling of 0.25°C with regional cooling by as much as 3°C; Muri et al. (2014) found global mean cooling of ~1°C; and Storelvmo et al. (2014) found global mean cooling of 1.4°C in coupled simulations of high latitude cirrus cloud thinning. As such, we conclude that the simplistic method of decreasing cirrus cloud optical depth does not capture the relevant effects necessary to represent cirrus cloud thinning.

A more complicated representation of cirrus cloud thinning, yet one that many models can still reproduce, would be to double the ice crystal fall speed. Ice sedimentation velocity is used as a model tuning parameter in the ECHAM family of models (Roeckner et al., 2003). This approach of increasing ice crystal fall speed was adopted by Muri et al. (2014) in their simulations using the Community Earth System Model version 1.0.3 (Hurrell et al., 2013). Storelvmo et al. (2013, 2015) and Storelvmo and Herger (2014) conducted similar cirrus thinning simulations in the same atmosphere model but using two different, more complicated cirrus parameterizations (Gettelman et al., 2008, 2010; Liu et al., 2012a; Barahona and Nenes, 2008, 2009). Both of these parameterizations have been validated by Liu et al. (2012b). They found that the prominent climate effects of cirrus thinning are well approximated by simply increasing the cirrus fall speed, lending credence to our chosen method. This idealized representation serves as a sensitivity test involving parameter perturbation, potentially providing cobenefits to other model intercomparison projects like CFMIP.

Figure 4 shows simulations from a different model, NorESM1-ME (Tjiputra et al., 2013), in which the fall speed was increased by varying amounts. Reductions in global mean temperature occur in all simulations. This representation is also not ideal, as fall speed is greater for large crystals. Actually introducing ice nuclei (IN) would result in large ice crystals (although not so large as to fall out quickly), but increasing the fall speed causes all large crystals to fall out quickly, resulting in an unrealistically small size distribution of crystals. Doubling the size of the ice crystals would be a better representation of cirrus cloud seeding, but how best to double a size distribution is not well-defined. Moreover, a change in size of the ice crystals would change the scattering properties of the crystals; accounting for this effect in a way that is consistent across all participating models would be quite complicated. Figure 5 shows that for an eight-fold increase of the ice crystal fall speed against a background of RCP8.5, relative humidities in the upper troposphere are reduced by over 30% in the tropical upper troposphere, which is consistent with the aims of cirrus cloud thinning. As such, we conclude that despite the shortcomings listed previously, increasing the sedimentation velocity of the ice crystals captures many of the hypothesized effects of cirrus thinning, particularly upper troposphere humidity changes.

For the design of this experiment, we recommend that all simulations of G7cirrus follow the simple approximation similar to that of Muri et al. (2014). All modeling groups should add a new local variable that replaces (in all locations where temperature is colder than 235 K) the ice mass mixing ratio in the calculation of the sedimentation velocity with a value that is eight times the original ice mass mixing ratio. We acknowledge that this approach has many shortcomings. Increasing the sedimentation velocity may not capture part of the cooling effect due to the increase in crystal size. It would also artificially increase fall speed without having larger ice crystals. However, this method captures many of the broad effects of cirrus thinning and avoids the very difficult task of including a doubling of the ice crystal size in both radiative transfer and fall speed calculations; this more complicated approach is not straightforward to incorporate in all models.

Storelvmo and Herger (2014) found that the majority of the cirrus thinning effects on net cloud forcing and surface temperatures are due to cirrus seeding outside of the tropics; including the tropics in the regions that are seeded caused a modest additional effect. However, so as not to introduce artificial boundaries in the regions where cirrus clouds are altered, cirrus clouds will be modified at all latitudes.

The design of G7cirrus (Figure 6) is comparable to previous GeoMIP experiments. Against a background of the ScenarioMIP Tier 1 high forcing scenario, cirrus seeding will begin in 2020 and continue through the year 2100. The goal of this experiment is to seed cirrus by a constant amount that reduces average global mean temperature in the decade 2020-2029 to that of the decade 1970-1979 (as calculated in a historical run), offsetting a radiative forcing of approximately 1.0 W m^{-2}. The decade 1970-1979 was chosen to avoid the climate effects of the 1982 El Chichón eruption, the 1991 Mount Pinatubo eruption, and the unusually large El Niño events in 1982 and 1998. Unlike G6sulfur or G6solar, G7cirrus does not propose to return net radiative forcing from one ScenarioMIP Tier 1 scenario to another, as it is yet unclear what levels of forcing could be achieved through cirrus seeding.

Because the goal of G7cirrus is to simulate cooling, it is best run against a background of a warming climate. However, this prevents diagnoses of the sensitivity to forcing from Atmosphere Model Intercomparison Project (AMIP)-style simulations, in which sea surface temperatures are fixed at present-day values. Because of the usefulness of these sorts of simulations in providing complementary information about radiative forcing, we recommend timeslice simulations (Section 3.1) also be performed, where sea surface temperatures are fixed at values corresponding to time periods more relevant to the proposed scenario than present day values.

It is important to reiterate that cirrus cloud processes are poorly understood and poorly represented in climate models. As an example, comparisons between observed and modeled ice water path in CMIP5 models reveal model biases of a factor of 2-10 (Li et al., 2012). Nevertheless, preliminary results and recent studies indicate that G7cirrus will reveal commonalities among model responses. Therefore, in addition to providing relevant information about the potentials and limitations of cirrus thinning, exploring inter-model differences in the results can reveal sources of model biases, directly addressing one of the core scientific questions in CMIP6.

3. Tier 2 Experiments in GeoMIP6

In addition to the four Tier 1 experiments, we propose another set of experiments that will aid in diagnosing climate model response.

3.1. Timeslice Simulations

Separately calculating the rapid adjustments and the feedback response (also called the fast and slow responses, respectively) can reveal fundamental climate behavior. This has been shown to be particularly useful for geoengineering simulations (Tilmes et al., 2013; Kravitz et al., 2013b; Huneeus et al., 2014). As such, we are requesting that all participating modeling groups conduct timeslice simulations (e.g., Cubasch et al., 1995) for each of the Tier 1 experiments to aid in diagnosing radiative forcing for the scenarios proposed here. These simulations will provide key information about how the climate system response to radiative forcing, as well as the relative sensitivities of climate responses and model biases to changes in aerosol and cloud microphysical properties, thus directly addressing several of the core science questions in CMIP6.

These timeslice experiments involve fixed sea surface temperature (SST) simulations for a period of 10 years; these are similar to Radiative Flux Perturbation simulations (Haywood et al., 2009). In these simulations, SSTs, sea ice, and all boundary conditions (greenhouse gas concentrations, aerosols, and other climate forcing agents) are to be prescribed at a constant climatology for the entire 10-year simulation. In most of the timeslice simulations, an external forcing is applied. For this forcing, the climatology is derived from the appropriate geoengineering experiment. For all the other boundary conditions, the climatologies are derived from the appropriate reference scenarios, in which no geoengineering is applied. Each Tier 1 experiment will have two associated timeslice simulations, one at the beginning of the coupled simulation and one at the end. The timeslice simulations are described in more detail in Table 2.

3.2. G4-Specified Stratospheric Aerosol experiment (G4SSA)

There are several issues in simulations of geoengineering with prognostic stratospheric sulfate aerosols, as differences in the resulting aerosol distribution can have prominent effects on the climate impacts of geoengineering and thus can produce large differences in the response between the models. To remove this difference between the models, Tilmes et al. (2014a) have designed an experiment for chemistry climate models (CCMs) called G4SSA. This experiment is designed so that all models would use the same prescribed stratospheric sulfur distribution, allowing for assessments of the range of climate responses for different representations of aerosol-chemistry and climate interactions. This experiment is connected to the other experiments in the Chemistry Climate Model Initiative (CCMI).

The experiment design takes inspiration from GeoMIP experiment G4. Against a background of RCP6.0, a layer of stratospheric aerosols will be injected into the model at a rate of 8 Tg SO_2 per year. Instead of allowing the models to calculate their aerosol distributions, a distribution of surface area density and other aerosol parameters will be provided to all models. The described distribution can also be scaled so as to apply to other scenarios, such as the ScenarioMIP scenarios (this is relevant for Experiment G6sulfur). We will provide time series of aerosol optical depth and ozone concentration that are consistent with the aerosol distribution at the website https://www2.acd.ucar.edu/gcm/geomip-g4-specified-stratospheric-aerosol-data-set.

Although G4SSA was developed for CCMs, it would be useful to obtain results from general circulation models (GCMs) or Earth system models (ESMs) as well, hence the inclusion in GeoMIP6. These two classes of models have very different treatments of the atmosphere, including stratospheric chemistry, aerosol microphysics, and representation of the quasi-biennial oscillation. Comparing results from these two groups would reveal some of the mechanisms behind the climate model response to stratospheric aerosol geoengineering, as well as provide a guideline for identifying which model representations of physical processes need improvement.

3.3 Overshoot Scenarios: G6sulfurExt and G6solarExt

ScenarioMIP includes an overshoot scenario (Boucher et al., 2012). In this experiment, beginning from the ScenarioMIP Tier 1 highest forcing scenario, aggressive emissions reductions beginning in the year 2100 would linearly reduce net anthropogenic emissions from those of the highest forcing scenario to those of the lowest forcing scenario. Analysis of this scenario will provide information on any potential hystereses in the simulated Earth system response and could provide warnings about potential tipping points or irreversible changes. As emissions reductions occur over the 22nd and 23rd centuries, the overshoot scenario would be an extension of the Tier 1 high forcing scenario through the year 2300. It is worth noting that the decline in forcing over the 22nd and 23rd centuries will not be linear, and the forcing level would be higher than in the lowest forcing scenario. Details on the actual forcing will be provided by the coordinators of ScenarioMIP.

Here we propose extensions of G6sulfur and G6solar that parallel the ScenarioMIP overshoot scenario; these simulations are similar to those described by Wigley (2006). The general principle behind these proposed extensions is that, at any time that the net forcing is greater in magnitude than that of the ScenarioMIP Tier 1 medium forcing scenario, geoengineering is used to reduce the net forcing. This would effectively result in a situation in which the magnitude of geoengineering is ramped up at the beginning of the simulation (before 2100, when the overshoot scenario starts). It is then ramped down near the end of the simulation once emissions reductions have sufficiently reduced the forcing from the level in the high forcing scenario, such that geoengineering would no longer be required to meet the forcing objective. This scenario will illuminate the extent to which geoengineering may help in preventing irreversible changes in the climate and avoiding tipping points.

4. The GeoMIP Testbed

A new feature of GeoMIP is termed the *GeoMIP Testbed*. This is a set of experiments that are potentially useful geoengineering studies that have been proposed by individual groups. The idea is that each group understands the key problems in its own sector and is thus uniquely posed to design a simulation that would best address those problems. That simulation design would then be vetted by individual models before a decision would be made as to whether the simulation should be undertaken by the full model suite.

4.1. G6sulfur_limits

Experiment G6sulfur is designed to reduce radiative forcing in a high emissions scenario to that of a moderate emissions scenario via simulating stratospheric sulfate aerosol injection. This experiment would be useful in assessing the effectiveness of geoengineering as part of a portfolio of responses to climate change. However, this experiment does not address feasibility or limits of stratospheric sulfate aerosol injection. As was stated in Section 2.2, increasing amounts of stratospheric SO_2 injection would cause particles to coagulate and fall out more rapidly. Therefore, the relationship between the amount of injection and the resulting radiative forcing is projected to be sublinear. This problem prompts a natural question: What is the limit of achievable radiative forcing from stratospheric sulfate aerosol injection?

A natural first step in addressing this problem would involve a similar setup to that of G6sulfur. Against a background of the ScenarioMIP Tier 1 high forcing scenario, sulfate aerosol precursors would be injected into the stratosphere in sufficient amounts to reduce anthropogenic radiative forcing from the levels in the high forcing scenario to levels in the low forcing scenario. As the low forcing scenario is a ScenarioMIP Tier 1 experiment, it would likely be conducted by all GeoMIP participants, and the extra simulations would be done with relatively little preparation.

Figure 7 shows the required amount of stratospheric aerosol injection to achieve given amounts of radiative forcing; these simulations were performed in ECHAM-HAM (Stier et al., 2005; Niemeier et al., 2011), a general circulation model coupled to an aerosol microphysical model that simulates the physical evolution and particle growth of sulfate aerosols. The sublinear relationship between injection amount and radiative forcing is clearly illustrated. The difference between RCP8.5 and RCP2.6 in the year 2100 is 5.9 Wm^{-2}, or the approximate radiative forcing of a tripling of the preindustrial CO_2 concentration; this difference is similar to the expected difference in forcing between the ScenarioMIP Tier 1 high forcing scenario and the Tier 1 low forcing scenario, when those scenarios are finalized. Extrapolating from the results of Figure 7, achieving this radiative forcing would require an injection of 40-50 Tg S (80-100 Tg SO_2) per year. This injection rate is equivalent to four to five 1991 Mount Pinatubo eruptions per year. Some efforts to evaluate the climate effects of such a scenario are already underway (Niemeier and Timmreck, submitted).

4.2. GeoSulfur10, GeoSulfur20, GeoSulfur50

A different way of quantifying the effects of stratospheric aerosol geoengineering is to perform a series of experiments in which the hypothetical rate of injection of stratospheric sulfate aerosols is constrained. Such a simulation would be well suited to ascertain the range of model responses to a fixed amount of SO_2 injection, highlighting model diversity. Against a background of the ScenarioMIP Tier 1 high forcing scenario, the modeling groups will inject 10, 20, or 50 Tg of sulfur dioxide per year into the lower stratosphere, in a similar setup to Experiment G4 (Kravitz et al., 2011a).

4.3. GeoLandAlbedo

Experiment G1ocean-albedo has simulated the effects of marine cloud brightening by increasing ocean albedo by a constant multiplication factor (Kravitz et al., 2013). However, GeoMIP has not yet explored land-based approaches towards solar radiation management. Such approaches could readily be implemented on the regional scale, as human activities already control the albedo of a significant fraction of the land surface. We therefore propose an alternative experiment in which the land surface albedo is increased, against a background of the CMIP5 abrupt4xCO2 experiment.

Under experiment GeoLandAlbedo, the land surface albedo would be increased by a uniform amount of 0.1 across all urban and agricultural areas. Such an increment represents a reasonable estimate of the maximum large-scale albedo increase that could be achieved in practice (Lobell et al., 2006; Lenton and Vaughan, 2009; Davin et al., 2014). The aim of experiment GeoLandAlbedo would not be to achieve global energy balance, but rather to determine the extent to which land surface albedo changes could offset the effects of increasing greenhouse gases on a regional basis.

To some degree, different aspects of this problem have been explored. Irvine et al. (2011) determined that different types of surface albedo geoengineering were incapable of offsetting the radiative forcing from a doubling of the CO_2 concentration, and the adverse side effects of such attempts could be large. Focusing only on bio-engineering crops to increase crop canopy albedo (Ridgwell et al., 2009) could cause local cooling effects (Doughty et al., 2011) but would likely have a small global impact (Singarayer et al., 2009; Singarayer and Davies-Barnard, 2012).

All of the previous studies on terrestrial-based albedo increases were conducted with single models, so the robustness of the effectiveness of this particular method of geoengineering, as well as the side effects, have not yet been tested. Assessing the range of responses to terrestrial-based geoengineering is especially important, given the wide range of structural and parametric uncertainties associated with modeling land surface processes.

5. Conclusions

The climate model experiment designs presented here mark the beginning of a concerted effort to include broader perspectives within GeoMIP. The extension of all experiments to at least 80 years is recommended to obtain more robust estimates of changes in extremes and modes of variability; it will be particularly interesting to discover what results can be obtained from G1ext that could not be obtained through analyses of Experiment G1, particularly in relation to extreme events (Curry et al., 2014) and modes of climate variability (Gabriel and Robock, 2015). The two G6 experiments were designed to open the door toward possible conversations with designers of climate change scenarios. We have begun to explore potential synergies with ScenarioMIP, on which our core simulations are based. Experiment G7cirrus is the first model intercomparison of the new idea of cirrus thinning and is designed to open avenues of investigation in both geoengineering and cirrus cloud microphysical representations. G4SSA was designed to explore commonalities and differences between general circulation models and CCMs, potentially highlighting processes that are important in representing the effects of aerosols not only on atmospheric chemistry, but also on dynamics and climate.

Geoengineering has the potential to impact climate systems at all scales, so by incorporating requirements from communities studying these different systems, we can broaden the usefulness of GeoMIP to a wider variety of scientists, policy makers, and other stakeholders. The GeoMIP Testbed is a key part of this effort. Under this new framework, individual communities can propose and test experiments that are designed to address problems in their sectors, providing invaluable information as to whether simulations by the full GeoMIP community are warranted.

Nevertheless, there remain some key gaps in GeoMIP that can provide a roadmap for future experiment design. One notable area is in impacts assessment. GeoMIP is quite adept at calculating expected climate effects from particular geoengineering scenarios, but translating those effects into impacts on people has only been explored in a limited set of studies (e.g., Xia et al., 2014). Interaction with the impacts assessment communities is one of the highest priorities for future directions of GeoMIP. This is particularly applicable for effects on developing countries, many of which will be most affected by climate change, and thus might also be most affected by geoengineering.

Another notable gap is the effect of geoengineering on carbon cycle feedbacks. Studies with intermediate complexity ESMs suggest that geoengineering could have a profound effect on the global carbon cycle through, for example, an enhancement of the land carbon sink (Keller et al., 2014). While much can be learned about the response of the carbon cycle to geoengineering from the experiments proposed in this article, the atmospheric carbon concentration does not evolve freely in all experiments. Multi-model studies driven by emissions which allow the atmospheric CO_2 concentration to evolve freely would provide valuable insights into the effect of SRM on this important feedback (e.g., the Coupled Carbon Cycle Climate Model Intercomparison Project, or C4MIP; Friedlingstein et al., 2006).

Although we expect that this new suite of climate model experiments will be useful in addressing many uncertainties in the physical science of geoengineering, there will remain many key questions. These experiment designs are idealized and are not representative of how geoengineering might be done in the real world, if society were to decide to deploy it. These designs also do not include studies of feasibility; some of the designed strategies might be more easily implemented in the real world than others. Moreover, while physical science studies are necessary for gaining information about the effects and impacts of geoengineering, they are only one aspect among a multitude of concerns, relating to both natural and social sciences, that are crucial for making informed decisions about geoengineering (e.g., Robock, 2014).

Acknowledgments. We thank Thorsten Mauritsen and Erich Roeckner for input on the protocol for G7cirrus. We also thank Bjorn Stevens and the CMIP6 organizing committee for more general comments on the experimental protocol. Ben Kravitz and Hailong Wang are supported by the Fund for Innovative Climate and Energy Research (FICER). The Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by Battelle Memorial Institute under contract DE-AC05-76RL01830. Simulations performed by Ben Kravitz were supported by the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at Goddard Space Flight Center. Alan Robock is supported by NSF grants AGS-1157525 and GEO-1240507. The National Center for Atmospheric Research is funded by the National Science Foundation. Andy Jones was supported by the Joint UK DECC/Defra Met Office Hadley Centre Climate Programme (GA01101). Helene Muri is supported by the Norwegian Research Council project EXPECT (grant no. 229760/E10) and computing time was provided by NOTUR. U. Niemeier is supported by the priority program 1689 of the German Research Foundation within project CEIBRAL. Jana Sillmann is supported by the Norwegian Research Council project NAPEX (229778). Shingo Watanabe is supported by the SOUSEI program, MEXT, Japan.

References

Aquila, V., C. I. Garfinkel, P. A. Newman, L. D. Oman, and D. W. Waugh (2014), Modifications of the quasi-biennial oscillation by a geoengineering perturbation of the stratospheric aerosol layer, *Geophys. Res. Lett.*, *41*, 1738-1744, doi:10.1002/2013GL058818.

Barahona, D. and A. Nenes (2008), Parameterization of cirrus cloud formation in large-scale models : Homogeneous nucleation, *J. Geophys. Res.*, *113*, D11211, doi :10.1029/2007JD009355.

Barahona, D. and A. Nenes (2009), Parameterizing the competition between homogeneous and heterogeneous freezing in cirrus cloud formation—Monodisperse ice nuclei, *Atmos. Chem. Phys.*, *9*, 369-381, doi :10.5194/acp-9-369-2009.

Berdahl, M., A. Robock, D. Ji, J. C. Moore, A. Jones, B. Kravitz, and S. Watanabe (2014), Arctic cryosphere response in the Geoengineering Model Intercomparison Project G3 and G4 scenarios, *J. Geophys. Res.*, *119*, 1308-1321, doi :10.1002/2013JD020627.

Boucher, O., P. R. Halloran, E. J. Burke, M. Doutriaux-Boucher, C. D. Jones, J. Lowe, M. A. Ringer, E. Robertson, and P. Wu (2012), Reversibility in an Earth System model in response to CO_2 concentration changes, *Environ. Res. Lett.*, *7*, 024013, doi :10.1088/1748-9326/7/2/024013.

Boucher, O., et al. (2013), Clouds and Aerosols. 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 (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 571–658, doi:10.1017/CBO9781107415324.016.

Cubasch, U., J. Waszkewitz, G. Hegerl, and J. Perlwitz (1995), Regional climate changes as simulated in time-slice experiments, *Climatic Change*, *31*, 372-304, doi :10.1007/BF01095150.

Curry, C. L., et al. (2014), A multi-model examination of climate extremes in an idealized geoengineering experiment, *J. Geophys. Res.*, *119*, 5226-5239, doi:10.1002/2013JD020648.

Davin, E.L., S.I. Seneviratne, P. Ciais, A. Olioso, and T. Wang (2014), Preferential cooling of hot extremes from cropland albedo management. *Proc. Natl Acad. Sci.* Published online, doi:10.1073/pnas.1317323111.

Doughty, C. E., C. B. Field, and A. M. S. McMillan (2011), Can crop albedo be increased through the modification of leaf trichomes, and could this cool regional climate? *Climatic Change*, *104*, 379-387.

English, J. M., O. B. Toon, and M. J. Mills (2012), Microphysical simulations of sulfur burdens from stratospheric sulfur geoengineering, *Atmos. Chem. Phys.*, *12*, 4775-4793, doi:10.5194/acp-12-4775-2012.

Friedlingstein, P., et al. (2006), Climate-carbon cycle feedback analysis: Results from the C4MIP Model Intercomparison, *J. Climate*, *19*, 3337-3353, doi:10.1175/JCLI3800.1.

Gabriel, C. J. and A. Robock (2015), Stratospheric geoengineering impacts on El Niño/Southern Oscillation, *Atmos. Chem. Phys. Discuss.*, *15*, 9173-9202, doi:10.5194/acpd-15-9173-2015.

Gettelman, A., H. Morrison, and S. J. Ghan (2008), A new two-moment bulk stratiform cloud microphysics scheme in the Community Atmosphere Model, Version 3 (CAM3). Part II: Single-column and global results, *J. Climate*, *21*, 3660-3679.

Gettelman, A., et al. (2010), Global simulations of ice nucleation and ice supersaturation with an improved cloud scheme in the Community Atmosphere Model, *J. Geophys. Res.*, *115*, doi:10.1029/2009JD013797.

Haywood, J., L. J. Donner, A. Jones, and J.-C. Golaz (2009), Global indirect radiative forcing caused by aerosols: IPCC (2007) and beyond, *Clouds in the Perturbed Climate System* (J. Heintzenberg and R. J. Charlson, eds.), 451-467, MIT Press, Cambridge.

Heckendorn, P., D. Weisenstein, S. Fueglistaler, B. P. Luo, E. Rozanov, M. Schraner, L. W. Thomason, and T. Peter (2009), The impact of geoengineering aerosols on stratospheric temperature and ozone, *Environ. Res. Let.*, *4*, 045108, doi:10.1088/1748-9326/4/4/045108.

Huneeus, N. et al. (2014), Forcings and feedbacks in the GeoMIP ensemble for a reduction in solar irradiance and increase in CO_2, *J. Geophys. Res.*, *119*, 5226-5239, doi:10.1002/2013JD021110.

Hurrell, J. W., et al. (2013), The Community Earth System Model: A framework for collaborative research, *Bull. Amer. Meteor. Soc.*, *94*, 1339-1360, doi:10.1175/BAMS-D-12-00121.1.

Irvine, P. J., A. Ridgwell, and D. J. Lunt (2011), Climatic effects of surface albedo geoengineering, *J. Geophys. Res.*, *116*, D24112, doi:10.1029/2011JD016281.

Jarvis, A. and D. Leedal (2012), The Geoengineering Model Intercomparison Project (GeoMIP): A control perspective, *Atmos. Sci. Lett.*, *13*, 157-163, doi:10.1002/asl.387.

Jones, A., et al. (2013), The impact of abrupt suspension of solar radiation management (termination effect) in experiment G2 of the Geoengineering Model Intercomparison Project (GeoMIP), *Journal of Geophysical Research*, *118*, 9743-9752, doi:10.1002/jgrd.50762.

Keller, D. P., E. Y. Feng, and A. Oschlies (2014), Potential climate engineering effectiveness and side effects during a high carbon dioxide-emission scenario, *Nature Communications*, *5*, 3304, doi:10.1038/ncomms4304.

Kleidon, A., B. Kravitz, and M. Renner (2015), The hydrological sensitivity to global warming and solar geoengineering derived from thermodynamic constraints, *Geophys. Res. Lett.*, *42*, 138-144, doi:10.1002/2014GL062589.

Kokkola, H., R. Hommel, J. Kazil, U. Niemeier, A. –I. Partanen, J. Feichter, and C. Timmreck (2009), Aerosol microphysics modules in the framework of the ECHAM5 climate model – Intercomparison under stratospheric conditions, *Geosci. Model Dev.*, *2*, 97-112, doi:10.5194/gmd-2-97-2009.

Kravitz, B., A. Robock, O. Boucher, H. Schmidt, K. E. Taylor, G. Stenchikov, and M. Schulz (2011a), The Geoengineering Model Intercomparison Project (GeoMIP), *Atmos. Sci. Lett.*, *12*, 162-167, doi:10.1002/asl.316.

Kravitz, B., A. Robock, O. Boucher, H. Schmidt, and K. E. Taylor (2011b), Specifications for GeoMIP experiments G1 through G4, Version 1.0, available online at http://climate.envsci.rutgers.edu/GeoMIP/docs/specificationsG1_G4_v1.0.pdf.

Kravitz, B., *et al.* (2013a), Sea spray geoengineering experiments in the Geoengineering Model Intercomparison Project (GeoMIP): Experimental design and preliminary results, *J. Geophys. Res.*, *118*, 11175-11186, doi:10.1002/jgrd.50856.

Kravitz, B., et al. (2013b), An energetic perspective on hydrological cycle changes in the Geoengineering Model Intercomparison Project, *J. Geophys. Res.*, *118*, 13087-13102, doi:10.1002/2013JD020502.

Kravitz, B., et al. (2013c), Climate model response from the Geoengineering Model Intercomparison Project (GeoMIP), *J. Geophys. Res.*, *118*, 8320-8332, doi:10.1002/jgrd.50646.

Kravitz, B., A. Robock, and O. Boucher (2014a), Future directions in simulating solar geoengineering, *Eos Trans. Amer. Geophys. Union*, *95*, 280, doi:10.1002/2014EO310010.

Kravitz, B., D. G. MacMartin, D. T. Leedal, P. J. Rasch, and A. J. Jarvis (2014b), Explicit feedback and the management of uncertainty in meeting climate objectives with solar geoengineering, *Environ. Res. Lett.*, *9*, 044006, doi:10.1088/1748-9326/9/4/044006.

Lenton, T. M., and N. E. Vaughan (2009), The radiative forcing potential of different climate geoengineering options, *Atmos. Chem. Phys.*, 9, 5539-5561.

Li, J.-L. F., et al. (2012), An observationally based evaluation of cloud ice water in CMIP3 and CMIP5 GCMs and contemporary reanalyses using contemporary satellite data, *J. Geophys. Res.*, *117*, doi :10.1029/2012JD017640.

Liu, X., et al. (2012a), Toward a minimal representation of aerosols in climate models : Description and evaluation in the Community Atmosphere Model CAM5, *Geosci. Model Dev.*, *5*, 709-739.

Liu, X., et al. (2012b), Sensitivity studies of dust ice nuclei effect on cirrus clouds with the Community Atmosphere Model CAM5, *Atmos. Chem. Phys.*, *12*, 12061-12079, doi :10.5194/acp-12-12061-2012.

Lobell, D.B., et al. (2006), Biogeophysical impacts of cropland management changes on climate, *Geophys. Res. Lett.*, 33, L06708, doi:10.1029/2005GL025492.

MacMartin, D. G., B. Kravitz, D. W. Keith, and A. Jarvis (2014), Dynamics of the coupled human-climate system resulting from closed-loop control of solar geoengineering, *Clim. Dynam.*, *43*, 243-258, doi:10.1007/s00382-013-1822-9.

MacMartin, D. G., B. Kravitz, and P. J. Rasch, On solar geoengineering and climate uncertainty, *Geophys. Res. Lett.*, submitted.

Meehl, G. A., R. Moss, K. E. Taylor, V. Eyring, R. J. Stouffer, S. Bony and B. Stevens (2014), Climate Model Intercomparisons: Preparing for the Next Phase, *Eos Trans. AGU*, *95*, 77, doi:10.1002/2014EO090001.

Mitchell, D. L. and W. Finnegan (2009), Modification of cirrus clouds to reduce global warming, *Environ. Res. Lett.*, *4*, 045102, doi:10.1088/1748-9326/4/4/045102.

Muri, H., J. E. Kristjánsson, T. Storelvmo, and M. A. Pfeffer (2014), The climatic effects of modifying cirrus clouds in a climate engineering framework, *J. Geophys. Res.*, *119*, 4174-4191, doi:10.1002/2013JD021063.

NAS (2015), Climate Intervention: Reflecting Sunlight to Cool Earth; Committee on Geoengineering Climate: Technical Evaluation and Discussion of Impacts; Board on Atmospheric Sciences and Climate; Ocean Studies Board; Division on Earth and Life Studies; National Research Council, available online at http://www.nap.edu/catalog/18988/climate-intervention-reflecting-sunlight-to-cool-earth.

Niemeier, U., H. Schmidt, and C. Timmreck (2011), The dependency of geoengineered sulfate aerosol on the emission strategy, *Atmos. Sci. Lett.*, *12*, 189-194, doi:10.1002/asl.304.

Niemeier, U., H. Schmidt, K. Alterskjær, and J. E. Kristjánsson (2013), Solar irradiance reduction via climate engineering–impact of different techniques on the energy balance and the hydrological cycle, *J. Geophys. Res.*, *118*, 11905-11917, doi:10.1002/2013JD020445.

Niemeier, U. and C. Timmreck, What is the limit of stratospheric sulfur climate engineering, *Atmospheric Chemistry and Physics*, submitted, doi:10.5194/acpd-15-10939-2015.

O'Neill, B. C., et al. (2014), A new scenario framework for climate change research: The concept of shared socioeconomic pathways, *Climatic Change*, *122*, 387-400, doi:10.1007/s10584-013-0905-2.

Plumb, R. A. and R. C. Bell (1982), A model of quasibiennial oscillation on an equatorial beta plane, *Q. J. Roy. Meteor. Soc.*, *108*, 335-352.

Ridgwell, A., J. S. Singarayer, A. M. Hetherington, and P. J. Valdes (2009), Tackling regional climate change by leaf albedo bio-geoengineering, *Current Biology*, *19*, 146-150, doi:10.1016/j.cub.2008.12.025.

Robock, A. (2014) Stratospheric aerosol geoengineering, *Issues Env. Sci. Tech.* (special issue "Geoengineering of the Climate System"), *38*, 162-185.

Roeckner, E., et al. (2003), The Atmospheric General Circulation Model ECHAM5, Part I : Model description, Max Planck Institute for Meteorology Report No. 349, available online at http://www.mpimet.mpg.de/fileadmin/models/echam/mpi_report_349.pdf.

Schäfer, S., et al. (2015), The European Transdisciplinary Assessment of Climate Engineering (EuTRACE): Removing Greenhouse Gases from the Atmosphere and Reflecting Sunlight away from Earth.

Schmidt, G. A., et al. (2014), Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive, *J. Adv. Model. Earth Syst.*, *6*, 141-184, doi:10.1002/2013MS000265.

Singarayer, J. S., A. Ridgwell, and P. Irvine (2009), Assessing the benefits of crop albedo bio-geoengineering, *Environ. Res. Lett.*, *4*, 045110, doi:10.1088/1748-9326/4/4/045110.

Singarayer, J. S. and T. Davies-Barnard (2012), Regional climate change mitigation with crops: Context and assessment, *Phil. Trans. Roy. Soc. A*, *370*, 4301-4316, doi: 10.1098/rsta.2012.0010.

Stier, P., et al. (2005), The aerosol-climate model ECHAM5-HAM, *Atmos. Chem. Phys.*, *5*, 1125-1156, doi:10.5194/acp-5-1125-2005.

Storelvmo, T., J. E. Kristjansson, H. Muri, M. Pfeffer, D. Barahona and A. Nenes (2013), Cirrus cloud seeding has potential to cool climate, *Geophys. Res. Lett.*, *40*, 178–182, doi:10.1029/2012GL054201.

Storelvmo, T., and N. Herger (2014), Cirrus cloud susceptibility to the injection of ice nuclei in the upper troposphere, *J. Geophys. Res. Atmos.*, *119*, 2375–2389, doi:10.1002/2013JD020816.

Storelvmo, T., W. R. Boos, and N. Herger (2015), Cirrus cloud seeding: A climate engineering mechanism with reduced side effects? *Phil. Trans. Roy. Soc. A.*, *20140116,* doi:10.1098/rsta.2014.0116.

Tilmes, S., et al. (2013), The hydrological impact of geoengineering in the Geoengineering Model Intercomparison Project (GeoMIP), *J. Geophys. Res.*, *118*, 11036-11058, doi :10.1002/jgrd.50868.

Tilmes, S., et al. (2014), A new Geoengineering Model Intercomparison Project (GeoMIP) experiment designed for climate and chemistry models, *Geosci. Model Dev. Discuss.*, ***7***, 5447-5464, doi:10.5194/gmdd-7-5447-2014.

Tjiputra, J. F., Roelandt, C., Bentsen, M., Lawrence, D. M., Lorentzen, T., Schwinger, J., Seland, Ø., and Heinze, C.: Evaluation of the carbon cycle components in the Norwegian Earth System Model (NorESM), *Geosci. Model Dev.*, *6*, 301-325, doi:10.5194/gmd-6-301-2013, 2013.

Wigley, T. M. L. (2006), A combined mitigation/geoengineering approach to climate stabilization, *Science*, *314*, 452-454, doi:10.1126/science.1131728.

Xia, L., et al. (2014), Solar radiation management impacts on agriculture in China: A case study in the Geoengineering Model Intercomparison Project (GeoMIP), *J. Geophys. Res.*, *119,* 8695-8711, doi:10.1002/2013JD020630.

Yu, X., J. C. Moore, X. Cui, A. Rinke, D. Ji, B. Kravitz, and J.-H. Yoon (2015), Impacts, effectiveness and regional inequalities of the GeoMIP G1 to G4 solar radiation management scenarios, *Global Planet. Change*, *129*, 10-22, doi:10.1016/j.gloplacha.2015.02.010.

Table 1. All core GeoMIP experiments up to this point, including the additional proposed Tier 1 GeoMIP6 experiments. Only the timeslice Tier 2 experiments are listed in Table 2. For each experiment, the name is given, along with a short description and reference. Newly proposed experiments are printed in boldface. G5 is not a core GeoMIP experiment but is included for completeness.

Experiment name	**Description**	**Reference**
G1	Balance 4xCO_2 via solar irradiance reduction	Kravitz et al. (2011)
G1ext	**Same as G1 but extended an extra 50 years**	**This document**
G1ocean-albedo	Balance 4xCO_2 via ocean albedo increase	Kravitz et al. (2013)
G2	Balance 1% CO_2 increase per year via solar irradiance reduction	Kravitz et al. (2011)
G3	Keep TOA radiative flux at 2020 levels against RCP4.5 via stratospheric sulfate aerosols	Kravitz et al. (2011)
G4	Injection of 5 Tg SO2 into lower stratosphere per year	Kravitz et al. (2011)
G4cdnc	Increase CDNC in marine low clouds by 50% against a background of RCP4.5	Kravitz et al. (2013)
G4sea-salt	Inject sea salt aerosols into tropical marine boundary layer to achieve ERF of -2.0 W m^{-2} against a background of RCP4.5	Kravitz et al. (2013)
G5	Identical setup as G3 but using sea salt injection into marine low clouds (IMPLICC experiment; named SALT in Niemeier et al., 2013)	Alterskjær et al. (2013); Niemeier et al. (2013)
G6sulfur	**Reduce forcing from ScenarioMIP Tier 1 high forcing scenario to the medium forcing scenario with stratospheric sulfate aerosols**	**This document**
G6solar	**Reduce forcing from ScenarioMIP Tier 1 high forcing scenario to the medium forcing scenario with solar irradiance reduction**	**This document**
G7cirrus	**Reduce forcing by constant amount (against a baseline of the ScenarioMIP Tier 1 high forcing scenario) via increasing cirrus ice crystal fall speed**	**This document**

Table 2. Timeslice simulations associated with each of the four Tier 1 experiments. Further description of the timeslice simulations is given in Section 3.1. Each tier 1 has two associated timeslice simulations: one for the beginning of the coupled simulation and one at the end of the coupled simulation. Note that the first timeslice simulations for G6sulfur and G6solar is identical, as no geoengineering has been applied yet. As such, this simulation is simply called G6Slice1.

Experiment Name	Applied forcing	Boundary conditions
G1extSlice1	4xCO2	piControl
G1extSlice2	4xCO2	abrupt4xCO2 after 100 years
G6Slice1	None	ScenarioMIP Tier 1 high forcing scenario in year 2020
G6sulfurSlice2	G6sulfur in year 2100	ScenarioMIP Tier 1 high forcing scenario in year 2100
G6solarSlice2	G6solar in year 2100	ScenarioMIP Tier 1 high forcing scenario in year 2100
G7cirrusSlice1	G7cirrus in year 2020	ScenarioMIP Tier 1 high forcing scenario in year 2020
G7cirrusSlice2	G7cirrus in year 2100	ScenarioMIP Tier 1 high forcing scenario in year 2100

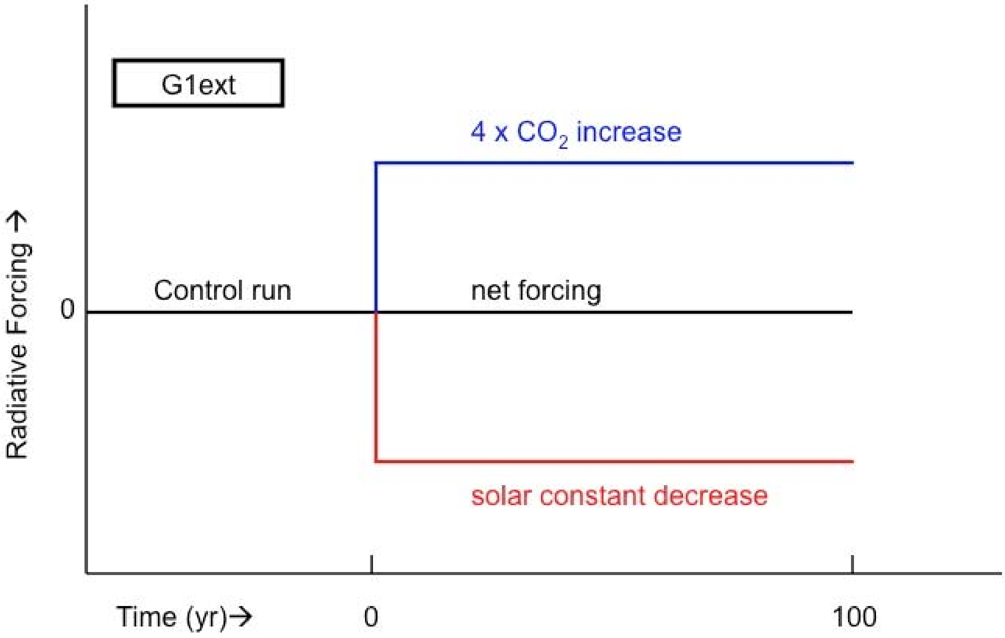


Figure 1. Schematic of experiment G1ext. The experiment is started from a preindustrial control run. The instantaneous quadrupling of the CO_2 concentration from its preindustrial value is balanced by a reduction in solar irradiance for 100 years.

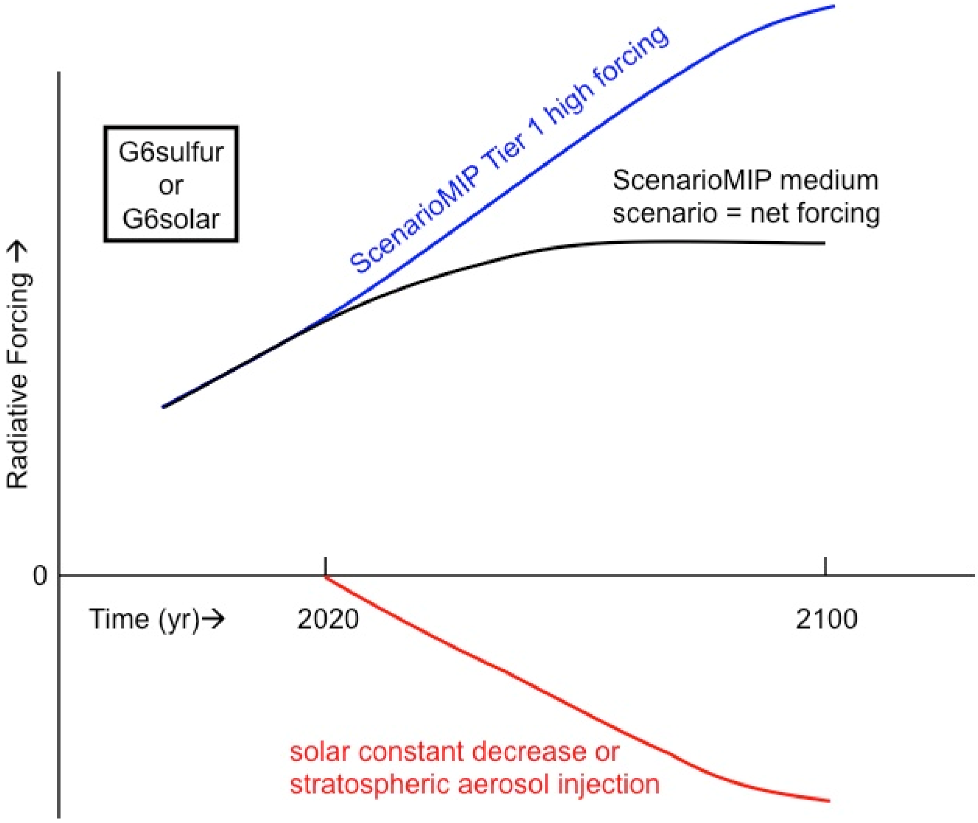


Figure 2. Schematic of experiments G6sulfur and G6solar. Against a background of the ScenarioMIP Tier 1 high forcing scenario, geoengineering will be conducted at time-varying amounts to return net anthropogenic radiative forcing to the levels of the ScenarioMIP Tier 1 medium forcing scenario. Geoengineering will be accomplished by stratospheric aerosol injection (G6sulfur) or solar irradiance reduction (G6solar).

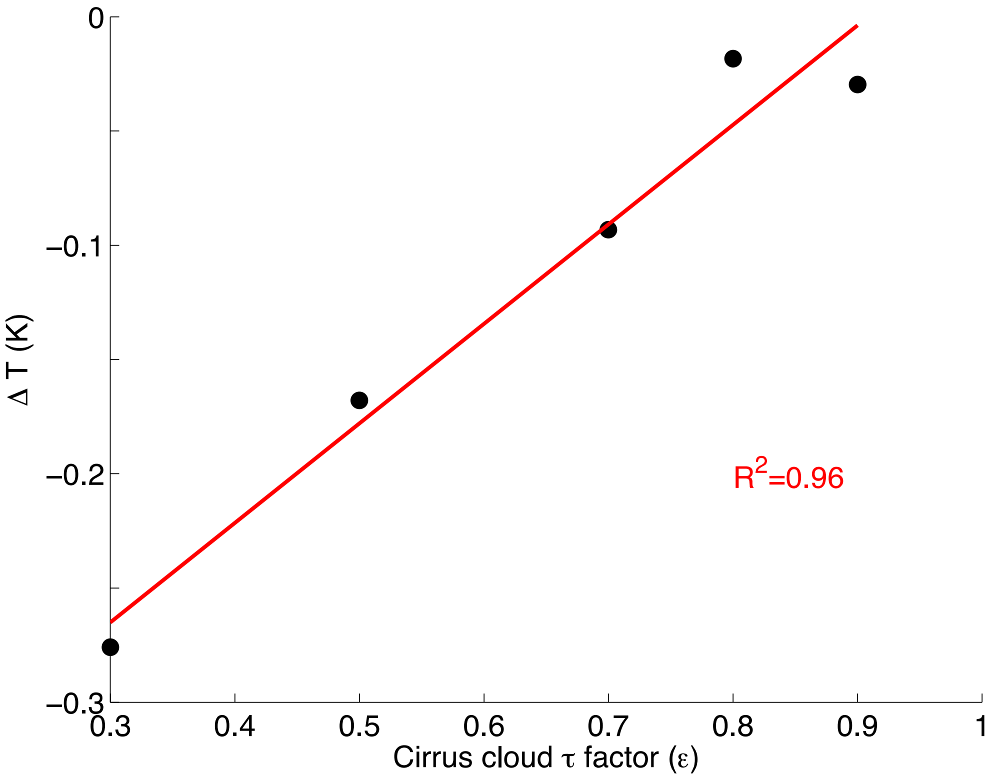


Figure 3. Test simulations of reducing cirrus cloud optical depth (τ) as described in Section 2.4. τ was scaled by a factor ε < 1 (x-axis). The amount of surface air temperature change due to this scaling (y-axis) was measured over a 4 year average; 0 indicates the global mean surface air temperature over years 2020-2023 in an RCP8.5 simulation. All simulations were performed using GISS ModelE2 (Schmidt et al., 2014).

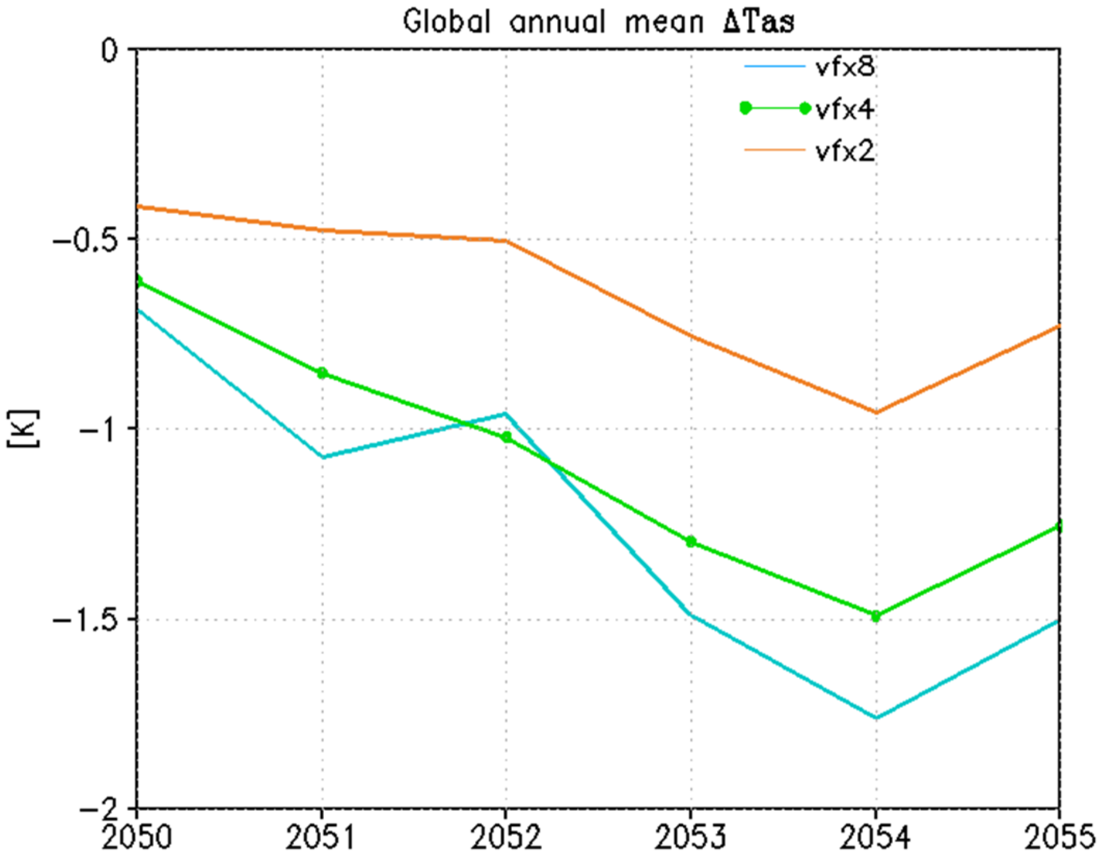


Figure 4. A sensitivity study of the effects of changing cirrus ice crystal sedimentation velocity in NorESM1-ME. vfx2, vfx4, and vfx8 indicate an increase in the sedimentation velocity by 2, 4, and 8 times, respectively. y-axis shows the global mean temperature change as a function of year (x-axis); differences are calculated with respect to an average over years 2050-2055 under an RCP8.5 scenario.

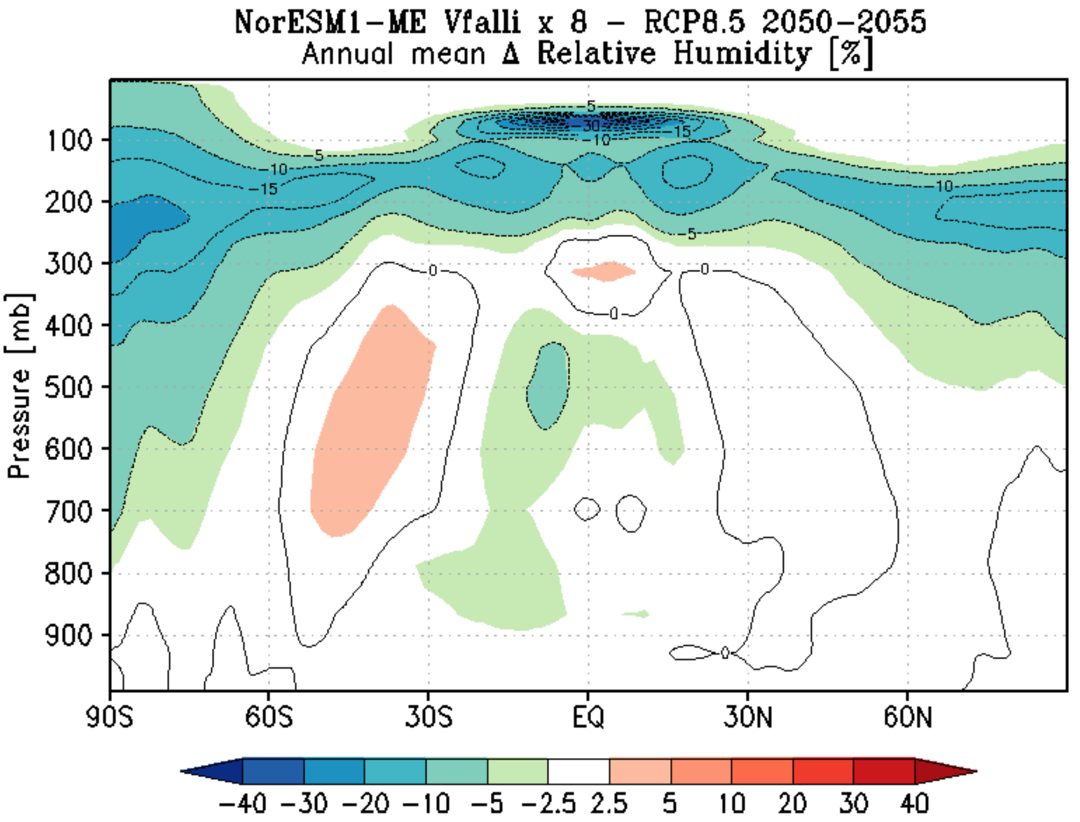


Figure 5. Zonally averaged annual mean of the difference in relative humidity (%) from NorESM1-ME for an octupling of the cirrus ice crystal fall speed. Differences are calculated as an average over years 2050-2055 against a background of RCP8.5.

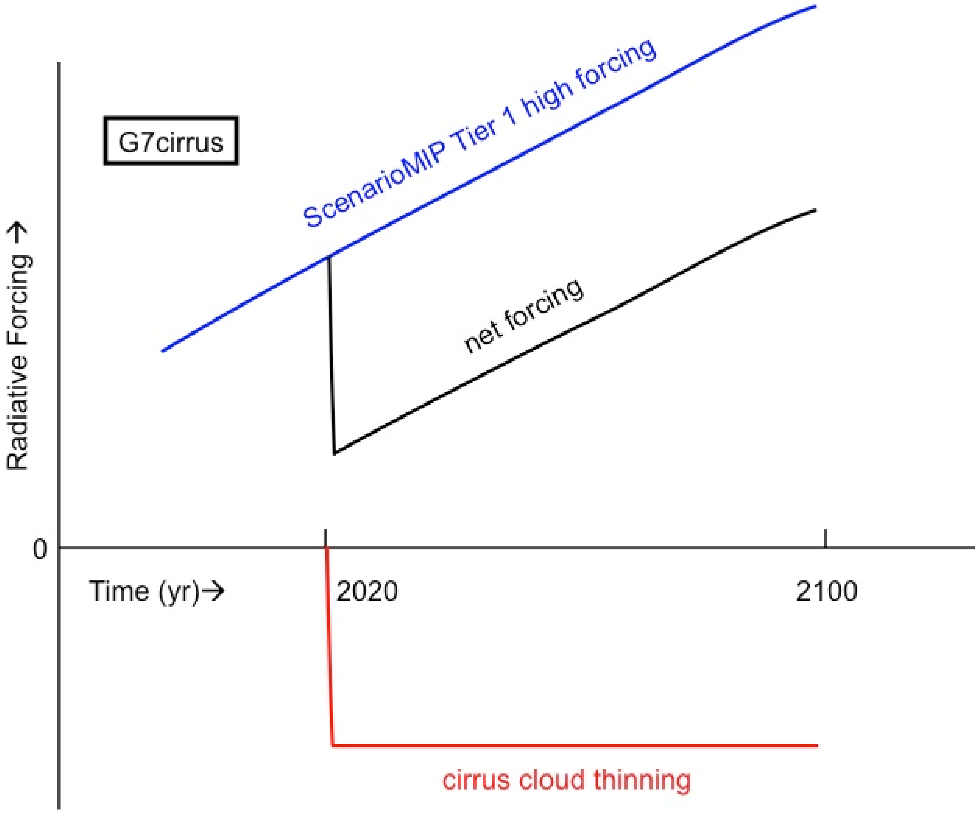


Figure 6. Schematic of experiment G7cirrus. Against a background scenario of the ScenarioMIP Tier 1 high forcing scenario, a representation of cirrus cloud seeding will reduce net forcing by a constant amount. This simulation will begin in 2020 and will be conducted for 80 years.

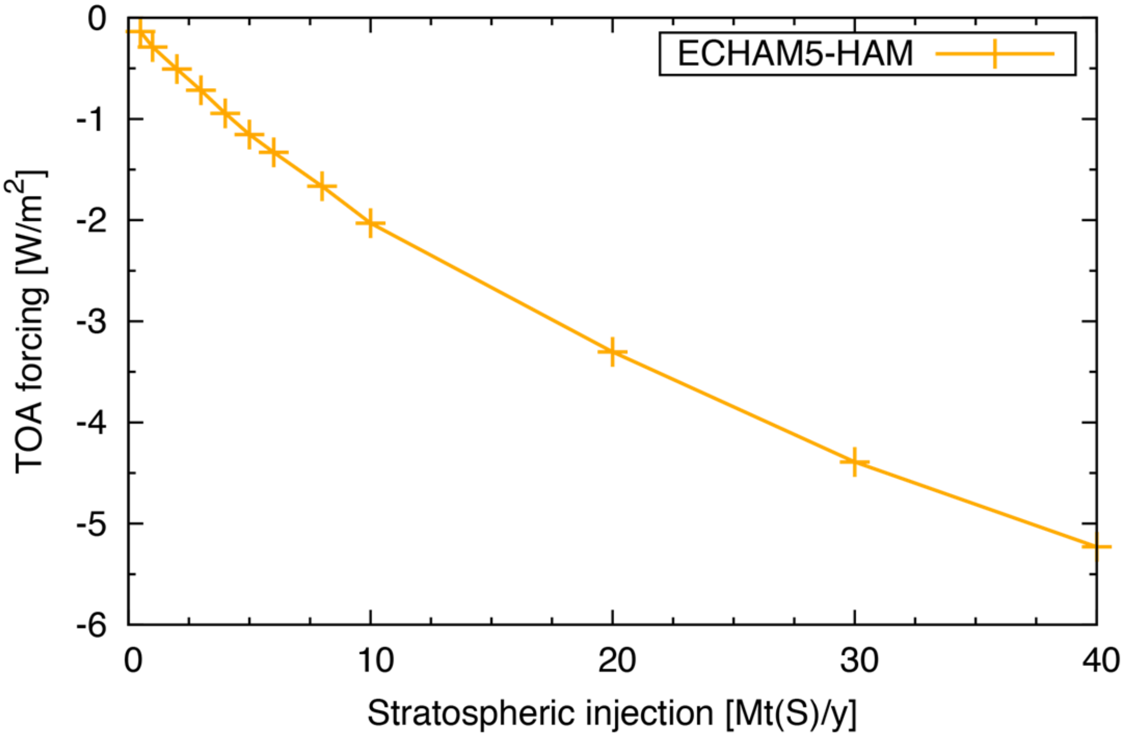


Figure 7. This figure shows the amount of annual stratospheric injection (x-axis) required to offset a given level of TOA net radiative flux imbalance (y-axis) in ECHAM5-HAM, an atmospheric general circulation model with a treatment of the microphysical evolution of sulfate aerosols. Maintaining 2020 values of net TOA radiative flux imbalance against a background of RCP8.5 requires an injection of approximately 70 Tg(S)/year in 2100 (based on extrapolation of the above values). All values were calculated for injection of SO_2 into one grid box over the equator; other injection strategies would likely require a different injection rate to achieve the same radiative forcing.

Global Monsoons Modeling Inter-comparison Project (GMMIP)

Application for CMIP6-Endorsed MIPs

Last updated: 31 March 2015

Chairs:
Tianjun Zhou, Institute of Atmospheric Physics, Chinese Academy of Sciences, China (zhoutj@lasg.iap.ac.cn)
Andy Turner, University of Reading, UK (a.g.turner@reading.ac.uk)
James Kinter, COLA, George Mason University, USA (ikinter@gmu.edu)

Suggested Members of the Scientific Steering Committee (*new members from African and American monsoon communities, and GMMIP partners will be included*)
Bin Wang, University of Hawaii, USA (wangbin@hawaii.edu)
Yun Qian, Atmospheric Sciences & Global Change Division, Pacific Northwest National Laboratory, USA (yun.qian@pnnl.gov)
Bin Wang, Institute of Atmospheric Physics, Chinese Academy of Sciences, China (wab@lasg.iap.ac.cn)
Annalisa Cherchi, CMCC, Italy (annalisa.cherchi@ingv.it)

Website:
http://www.lasg.ac.cn/gmmip

Proposed by:
CLIVAR AAMP, CLIVAR-GEWEX MP, CLIVAR/C20C+, in collaboration with LASG/IAP, China and PNNL, USA

Goal of GMMIP:

Changes in the precipitation and atmospheric circulation in the global monsoons are of great scientific and societal importance owing to their impacts on more than two-thirds of the world’s population. Monsoons occur in various regions around the world. Prediction of the monsoon rainfall change in the coming decades is of deep societal concern and vital for infrastructural planning, water resource management, and sustainable economic development.

The dominant monsoon systems in the world include the Asian-Australian, African, and the American monsoons. Each monsoon system generally has its own unique and specific characteristics in terms of variability. At the same time, the connections in the global divergent circulation necessitated by mass conservation link the various regional monsoons as they evolve through the season. On interannual-to-multidecadal time scales, there is evidence that monsoon precipitation in the Northern Hemisphere (NH) and Southern Hemisphere (SH) varies coherently, driven by ENSO and other global modes of climate variability at the lower boundary of the atmosphere.

The combination of changes in monsoon area and rainfall intensity has led to an overall weakening trend of global land monsoon rainfall accumulation since the 1950s. This decreasing tendency is dominated by the African and South Asian monsoons, due to the significant decreasing tendencies of both rainfall intensity and monsoon coverage. Beginning in the 1980s, however, the NH global monsoon precipitation has shown an upward trend. Understanding the mechanisms of precipitation changes in the global monsoons and identifying the roles of natural and anthropogenic forcing agents have been foci of the monsoon research community.

While all monsoons are large-scale cross-equatorial overturning circulations, major differences between characteristics of the different regional monsoons arise because of the different orography. This is most apparent for the Asia region, due to the TIP/Himalaya.

Climate models are useful tools in climate variability and climate change studies. However, the performance of the current state-of-the-art climate models is very poor and needs to be greatly improved over the monsoon domains. The Global Monsoons Model Inter-comparison Project (hereafter **GMMIP**) aims to improve our understanding of physical processes in global monsoon systems and to better simulate the mean state, interannual variability and long-term change of global monsoons by performing multi-model inter-comparisons. The contributions of internal variability (IPO-Interdecadal Pacific Oscillation, AMO-Atlantic Multidecadal Oscillation) and external anthropogenic forcing to the historical evolution of global monsoons in the 20th and 21st century will be addressed.

Primary Science Questions:

1) What are the relative contributions of internal processes and external forcing that are driving the 20th century historical evolution of global monsoons?
2) To what extent and how does the atmopshere-ocean interaction contribute to the interannual variability and predictability?
3) What are the effects of Eurasian orography, in particular the Himalaya/Tibetan Plateau, on the regional/global monsoons?
4) How well can developing high-resolution models and improving model dynamics and physics help to reliably simulate monsoon precipitation and its variability and change?

By focusing on addressing these four questions we expect to deepen our understanding of models' capability in reproducing the monsoon mean state and its natural variability as well as the forced response to natural and anthropogenic forcing, which ultimately will help to reduce model uncertainty and improve the credibility of models in projecting future changes in the monsoon. The coordinated experiments will also help advance our physical understanding and prediction of monsoon changes.

Due to the uncertainties in the physical parameterizations in current models, the best way to address these questions is through a multi-model framework. CMIP6 provides a good opportunity for advancement of monsoon modeling and understanding. GMMIP will contribute to four of the five grand challenges of the WCRP, viz. Regional Climate Information, Water Availability, Climate Extremes, and Clouds, Circulation and Climate Sensitivity.

Proposed Experiments:

The main experiments of GMMIP will be divided into Tier 1 and Tier 2, with further optional ideas in Tier 3. The total experiments of GMMIP are summarized in Table 1. The **Tier-1** experiments will be extended AMIP runs. This is the ***entry card for GMMIP***.

Table 1: Experiment list of GMMIP

	EXP name	**Integration time**	**Short description and purpose of the EXP design**	**Model type**
Tier-1	AMIP20C	1870-2013	Extended AMIP run that covers 1850-2014. All natural and anthropogenic historical forcings as used in *CMIP6 Historical Simulation* will be included. AGCM resolution as *CMIP6 Historical Simulation*. The HadISST data will be used. Minimum number of integrations is 1, realizations more than 3 are encouraged.	AGCM
Tier-2	HIST-IPO	1870-2013	Pacemaker 20th century historical run that includes all forcing as used in *CMIP6 Historical Simulation*, and the observational historical SST is restored in the tropical lobe of the IPO domain (20°S-20°N, 175°E-75°W); to understand the forcing of IPO-related tropical SST to global monsoon changes. Models resolutions as *CMIP6 Historical Simulation*. The HadISST data will be used. Minimum number of integrations is 3, more realizations are encouraged.	CGCM with SST restored to the model climatology plus observational historical anomaly in the tropical lobe of IPO domain
Tier-2	HIST-AMO	1870-2013	Pacemaker 20th century historical run that includes all forcing as used in *CMIP6 Historical Simulation*, and the observational historical SST is restored in the AMO domain (0°-70°N, 70°W-0°); to understand the forcing of AMO-related SST to global monsoon changes. Models resolutions as *CMIP6 Historical Simulation*. The HadISST data will be used. Minimum number of integrations is 3, more realizations are encouraged.	CGCM with SST restored to the model climatology plus observational historical anomaly in the AMO domain
Tier-3	DTIP	1979-2013	The topography of the TIP is modified by setting surface elevations to 500m; to understand the combined thermal and mechanical forcing of the TIP. Same model as DECK. Minimum number of integrations is 1.	AGCM
Tier-3	DTIP-DSH	1979-2013	Surface sensible heat released at the elevation above 500m over the TIP is not allowed to heat the atmosphere; to compare of impact of removing thermal effects. Same model as DECK. Minimum number of integrations is 1.	AGCM
Tier-3	DHLD	1979-2013	The topography of the highlands in Africa, N. America and S. America TP is modified by setting surface elevations to a certain height (500m). Same model as DECK. Minimum number of integrations is 1.	AGCM

The **Tier-2 HIST-IPO run** is Pacemaker 20th century historical climate simulation that includes all forcing, and the sea surface temperature (SST) restored to the model climatology plus observational historical anomaly in the tropical lobe of the Interdecadal Pacific Oscillation (IPO; Power et al. 1999; Folland et al. 2002) domain (20°S-20°N, 175°E-75°W): the weight=1 in the inner box (15°S-15°N, 180°-80°W), linearly reduced to zero in the buffer zone (zonal and meridional ranges are both 5°) from the inner to outer box.

The **Tier-2 HIST-AMO run** is Pacemaker 20th century historical climate simulation that includes all forcing, and the SST restored to the model climatology plus observational historical anomaly in the Atlantic Multidecadal Oscillation (AMO; Enfield et al. 2001; Trenberth and Shea 2006) domain (0°-70°N, 70°W-0°): the weight=1 in the inner box (5°N-65°N, 65°W-5°W), linearly reduced to zero in the buffer zone (zonal and meridional ranges are both 5°) from the inner to outer box.

In **Tier-3 DTIP run**, following Boos and Kuang (2011, 2013) and Wu et al. (2007, 2012), the topography of the Tibetan Plateau(hereafter TIP) (20-60^{o}N, 25-120^{o}E) in the model is modified by leveling off the TIP to a certain height (e.g. 500m), with the surface properties unchanged. Other settings of the integration are the same as the standard DECK AMIP run. This experiment represents perturbations to both thermal and mechanical forcing of the TIP with respect to the standard DECK AMIP run.

In **Tier-3 DTIP-DSH run**, the surface sensible heat flux at elevations above 500m over the TIP is not allowed to heat the atmosphere, i.e., the vertical diffusive heating term in the atmospheric thermodynamic equation is set to zero (Wu et al. 2012).Other settings of the integration are the same as the standard DECK AMIP run. The differences between the standard DECK AMIP run and the DTIP-DSH are considered to represent the removal of TIP thermal forcing only and thus the circulation pattern of DTIP-DSH reflects the impacts of mechanical forcing.

Description of the analysis of GMMIP experiments:

There are four tasks in the analysis of GMMIP:

1) **Task-1:** The global monsoons changes in the 20th century
2) **Task-2:** The role of Eurasian orography on the regional/global monsoons (Himalaya/Tibetan Plateau experiment)
3) **Task-3:** Interannual variability of global monsoon precipitations
4) **Task-4:** High resolution modeling of global monsoons

The analysis of fours tasks will use the outputs of GMMIP experiments, DAMIP (Detection and Attribution MIP) experiments, HighResMIP experiments, the CMIP6 Historical Simulation, and the AMIP experiments of DECK.

Connection with DECK and CMIP6 Historical Simulation

The DECK simulations will serve as an entry card for the CMIP6-Endorsed MIPS. The **DECK** experiments are:

- AMIP simulations
- Pre-industrial control simulations
- 1%/yr increase in CO2 concentration
- Switch-on 4XCO2

The **CMIP6 Historical Simulation** experiment is:

- Historical simulation of fully coupled models (1850-2014)

The AMIP DECK simulation with the standard CMIP6 resolution will be used in the analysis of GMMIP. The Tier-1 AGCM experiment of GMMIP will specify the specific forcings which are consistent with the historical simulation from 1850-2014, viz. the CMIP6 Historical Simulation.

Connection with other MIPs

DAMIP (Detection and Attribution MIP):

The histALL (enlarging ensemble size of historical ALL foring runs in DECK), histNAT (Historical natural-only run), histGHG (Historical well-mixed GHG-only run), histAER experiments (Historical anthropogenic-Aerosols-only run) of DAMIP will be used in the analysis of Task-1 of GMMIP.

Combinations of histALL, histNAT and histGHG will allow us to understand the observed the 20th century global monsoon precipitation and circulation changes in the context of contributions from GHG, the other anthropogenic factors and natural forcing. The contributions of these external forcings will be compared to those from internal variability modes such as IPO and AMO.

HighResMIP:

The Tier-1 experiments of HighResMIP, which are AMIP runs but with minimum 25-50 km at mid-latitudes for high resolution and a standard resolution configuration (1950-2014), will be used in the analysis of Task-4 of GMMIP, which aims to examine the performance of high-resolution models in reproducing both the mean state and year-to-year variability of global monsoons.

The Tier-2 experiments of HighResMIP, which are coupled runs consisting of pairs of both historic runs and control runs using fixed 1950s forcing, will be used in the analysis of Task-3 of GMMIP, which aims to understand the role of air-sea interaction process in the improvement of monsoon mean state and year-by-year variability.

WCRP Grand Challenges:

GMMIP will address the grand challenges of the WCRP in the following way:

Regional Climate Information *(rank 1)*

GMMIP will improve our understanding of the 20th climate changes in global monsoon domains. The contributions of external anthropogenic forcings (GHG, aerosol), natural forcing, and internal variability modes (IPO, AMO) will be indentified. These would provide useful information for climate prediction/projections in the highly populated global monsoon domains.

Water Availability *(rank 2)*

The water resources in global monsoon domains are greatly affected by the anomalous activities of monsoons. Understanding the mechanisms of monsoon variability as posed by GMMIP will lead to improvement of monsoon prediction/projection and provide useful information for policymakers in water availability-related decision making.

Climate Extremes *(rank 2)*

Extreme events such as mega-droughts and flooding have been frequently occurred in global monsoon domains. GMMIP is hopefully to identify the useful ways of improving the simulation/prediction of climate extremes in global monsoon domains.

Clouds, Circulation and Climate Sensitivity *(rank 2).*

A reasonable simulation of monsoon circulation and clouds is a prerequisite for a successful simulation of monsoon precipitation. By comparing the performances of climate models with high and normal resolutions, model simulations with/without air-sea interaction processes, the implementation of GMMIP will link the monsoon circulations to monsoon precipitation in the context of reducing model bias and improving model performances.

GEWEX and CLIVAR

Monsoon has been a research focus of GEWEX and CLIVAR. The scientific questions listed in GMMIP were originally identified by the CLIVAR Asian-Australian monsoon panel, the GEWEX/CLIVAR Monsoons Panel, and CLIVAR/C20C+ project. The questions have also been highlighted by the reports of CLIVAR Research Opportunities Tiger Team on "*Decadal Variability in the Climate System and its Predictability*", and CLIVAR Research Opportunities Tiger Team on "*Intra-seasonal, Seasonal and Interannual Variability and Predictability of Monsoon Systems*".

Participation:

Participation in GMMIP is voluntary and open. GMMIP will be coordinated by a small working group composed of engaged representatives from climate diagnosis, climate change attribution and climate modeling communities. This working group will engage the broadest degree of input and involvement from members of the scientific community.

The Scientific Steering Committee (SSC) of GMMIP will be composed of representatives from CLIVAR & GEWEX monsoon panels, relevant projects and the global monsoon community. The SSC will provide comments and instructions for the analysis of GMMIP with focus on the scientific questions listed in the proposal.

The following modeling centers have expressed their interests in participating in GMMIP:

- ACCESS, Australia, (Harun Rashid,harun.rashid@csiro.au)
- BCC, China (Xiaoge Xin, xinxg@cma.gov.cn)
- BNU, China (Duoying Ji, duoyingji@bnu.edu.cn)
- CAMS-CSM, China (rongur@cams.cma.gov.cn)
- CAS-ESM, China (He Zhang, zhanghe@mail.iap.ac.cn)
- CanESM, Canada (John Scinocca , john.scinocca@ec.gc.ca)
- CESM, USA (J. Fasullo)
- CESS-THU, China (Jianbin Huang, jbh@mail.tsinghua.edu.cn)
- CFS- IITM-ESM, India (R. Krishnan, P. Swapna and J. Sanjay)
- CMCC, Italy (Annalisa Cherchi, annalisa.cherchi@ingv.it)
- CNRM, France (herve.douville@meteo.fr, cassou@cerfacs.fr)
- HadGEM3, UK (Andrew Turner)
- IAP-LASG FGOALS, China (Lixia Zhang, lixiazhang@mail.iap.ac.cn)
- IPSL, France (Laurent.li@lmd.jussieu.fr, pascale.braconnot@lsce.ipsl.fr)
- FIO, China (songroy@fio.org.cn; baoying@fio.org.cn)
- GISS, USA (Sonali McDermid, sps2113@columbia.edu, *TBC*)
- GFDL, USA (Yi Ming, yi.ming@noaa.gov)
- MPI-ESM, Germany (Juergen Bader, juergen.bader@mpimet.mpg.de)

- MIROC, Japan (Masahiro Watanabe, hiro@aori.u-tokyo.ac.jp)
- MRI, Japan (Hirokazu Endo)
- NUIST-CSM, China (pangchi@nuist.edu.cn)

Proposed timing

Start of the experiments: Beginning of 2016
End of the experiments: No fixed date.

References

Boos, W. R. &Z. M. Kuang 2010: Dominant control of the South Asian monsoon by orographic insulation versus plateau heating. ***Nature*** 463, 218-223.

Boos, W.R. &Z. M. Kuang 2013: Sensitivity of the South Asian monsoon to elevated and non-elevated heating. ***Sci. Rep.*** 3, 1192; DOI:10.1038/srep01192.

Cook, B.I., and R. Seager, 2013: The response of the North American Monsoon to increased greenhouse gas forcing. ***J. Geophys. Res.***,118, 1690-1699

Cook, K. H., G. A. Meehl, and J. M. Arblaster, 2012: Monsoon regimes and processes in CCSM4. Part II: African and American monsoon systems. ***J. Climate***, 25, 2609-2621

Enfield, D., A. Mestas-Nuñez, and P. Trimble (2001), The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S., ***Geophys. Res. Lett.***, 28, 2077–2080, doi:10.1029/2000GL012745.

Folland, C. K., J. A. Renwick, M. J. Salinger, and A. B. Mullan (2002), Relative influences of the Interdecadal Pacific Oscillation and ENSO on the South Pacific Convergence Zone, ***Geophys. Res. Lett.***, 29(13), 1643, doi:10.1029/ 2001GL014201.

Kitoh, A., H. Endo, K. Krishna Kumar, I.F.A. Cavalcanti, P. Goswami, T. Zhou, 2013: Monsoons in a changing world: a regional perspective in a global context. ***J. Geophys. Res.***, 118, 3053-3065,doi:10.1002/jgrd.50258.

Lin R., T. Zhou, Y. Qian, 2014: Evaluation of Global Monsoon Precipitation Changes based on Five Reanalysis Datasets, ***Journal of Climate***,27(3),1271-1289

Liu, J., B. Wang, M.A. Cane, S.-Y. Yim, and J.-Y. Lee, 2013: Divergent global precipitation changes induced by natural versus anthropogenic forcing. ***Nature***, 493, 656-659

Power, S., T. Casey, C. Folland, A. Colman, and V. Mehta (1999), Interdecadal modulation of the impact of ENSO on Australia, ***Clim. Dyn.***, 15, 319–324.

Qiu J. 2013: Monsoon Melee. ***Science***, 340 (6139), 1400-1401; DOI:10.1126/ science.340.6139.1400.

Song, F., T. Zhou,and Y. Qian (2014), Responses of East Asian summer monsoon to natural and anthropogenic forcings in the 17 latest CMIP5 models, ***Geophys. Res. Lett.***, 41,doi:10.1002/2013GL058705

Song, F., T. Zhou, 2014: Interannual Variability of East Asian Summer Monsoon Simulated by CMIP3 and CMIP5 AGCMs: Skill Dependence on Indian Ocean–Western Pacific Anticyclone Teleconnection. ***J. Climate***, 27, 1679-1697

Song F., T. Zhou, 2014: The climatology and inter-annual variability of East Asian summer monsoon in CMIP5 coupled models: Does air-sea coupling improve the simulations ? ***Journal of Climate***, 27, 8761–8777

Sperber, K.R., H. Annamalai, I.-S. Kang, A. Kitoh, A. Moise, A. Turner, B. Wang and T. Zhou, 2013: The Asian summer monsoon: an intercomparison of CMIP5 vs. CMIP3 simulations of the late 20th century. ***Clim. Dynam.***, 41, 2771-2744, doi:10.1007/s00382- 012-1607-6.

Trenberth, K. E., and D. J. Shea (2006), Atlantic hurricanes and natural variability in 2005, ***Geophys. Res. Lett.***, 33, L12704, doi:10.1029/2006GL026894.

Wang B., Q. H. Ding, X. H. Fu, I.-S. Kang, K. Jin, J. Shukla, and F. Doblas-Reyes, 2005: Fundamental challenge in simulation and prediction of summer monsoon rainfall. ***Geophys. Res. Lett.***, 32, L15711, doi:10.1029/2005GL022734.

Wang, B., J. Liu, H.-J. Kim, P.J. Webster, S.-Y. Yim, and B. Xiang, 2013: Northern Hemisphere summer monsoon intensified by mega-El Niño/southern oscillation and Atlantic multidecadal oscillation. ***PNAS***, 110 (14), 5347-5352, doi:10.1073/pnas.1219405110.

Wang, B., J. Liu, H.-J. Kim, P.J. Webster, and S.-Y. Yim, 2012: Recent Change of the Global Monsoon Precipitation (1979-2008). ***Clim. Dyn.***, 39 (5), 1123-1135

Wu, G., Y. Liu, T. Wang, et al., 2007: The Influence of the Mechanical andThermal Forcing of the Tibetan Plateau on the Asian Climate. ***J. Hydrometeorology*** 8: 770-789.

Wu, G., Y. Liu, B. He, Q. Bao, A.Duan& F.-F. Jin, 2012: Thermal Controls on the Asian Summer Monsoon. ***Sci. Rep.*** 2, 404; DOI:10.1038/srep00404.

Zhou T., R. Yu, H. Li, and B. Wang, 2008, Ocean Forcing to Changes in Global Monsoon Precipitation over the Recent Half-Century, ***Journal of Climate***, 21(15), 3833–3852

Zhou T., B. Wu, B. Wang, 2009, How Well Do Atmospheric General Circulation Models Capture the Leading Modes of the Interannual Variability of the Asian-Australian Monsoon? ***Journal of Climate***, 22, 1159-1173

Appendix: Description of the scientific objectives of four tasks of GMMIP

TASK-1: The global monsoons changes in the 20th century

The global monsoons have shown multi-decadal changes in the 20th century. Understanding the mechanisms of global monsoon changes and identifying the contributions of natural and anthropogenic forcing agents have been foci of the monsoon research community. **TASK-1** aims to reveal the role of forcing from the global oceans on monsoon precipitation change, and identify the relative contributions of natural and anthropogenic forcing (greenhouse gases and aerosols) by performing coupled, uncoupled, and partly coupled runs that cover the period from 1870 to 2013.

TASK-2: The role of Eurasian orography on the regional/global monsoons (Himalaya/Tibetan Plateau experiment)

Although monsoons are generally large-scale overturning circulations, apparent differences between characteristics of regional monsoons arise because of the different orography. This is most apparent for the Asia region, due to the existence of Tibetan – Iranian Plateau (TIP). The influence of the large-scale orography on the Asian summer monsoon includes both mechanical and thermal forcing. Various mechanisms have been suggested concerning the topographic effects; however, an overarching paradigm delineating the dominant factors determining these effects and the strength of impacts remains debated. **The goals of TASK-2** are to provide a benchmark of current model behavior in simulating the relationship of the monsoon to the Tibetan-Iranian Plateau (TIP, the highlands in 20-60^{o}N, 25-120^{o}E) so as to stimulate further research on the thermodynamical and dynamical effects of the TIP on the monsoon system. In particular the relative contributions of thermal and orographic mechanical forcing by the TIP to the Asian monsoon will be addressed. The task extends the studies from the TIP to other highlands including highlands in Africa, N. America and S. America.

TASK-3: Interannual variability of global monsoon precipitations

AGCM simulations with specified SST generally have low skill in simulating the summer precipitation over global monsoon domains, especially the Asian-western Pacific summer monsoon domain. This can be partly attributed to the exclusion of air-sea coupled processes. It is argued that in the real world the air-sea interaction in monsoon domains appears as “monsoon-driving-ocean”, but in an AMIP simulation, the interaction mechanism is “ocean-driving-monsoon” by construction (Wang et al. 2005). **The TASK-3 aims** to understand the air-sea interaction process in driving the interannual variability of global monsoons.

TASK-4: High resolution modeling of global monsoons

The monsoon rainbands are usually at a maximum width of 200 km. Climate models with low or moderate resolutions are generally unable to realistically reproduce the mean state and variability of monsoon precipitation. This is partly due to the model resolution. The **TASK-4 aims** to examine the performance of high-resolution models in reproducing both the mean state and year-to-year variability of global monsoons. High resolution rain-gauge observations and satellite precipitation products will be used to gauge models’ performances.

HighResMIP CMIP6 endorsement application

June 2015

Name of MIP

HighResMIP

Chairs

Rein Haarsma, KNMI, The Netherlands. (haarsma@knmi.nl)
Malcolm Roberts, Met. Office, UK. (malcolm.roberts@metoffice.gov.uk)

Suggested Members of the Scientific Steering Committee

Graeme Stephens, JPL, USA (graeme.stephens@jpl.nasa.gov)
Masahide Kimoto, Tokyo, University (kimoto@aori.u-tokyo.ac.jp)
Christiane Jablonowski, Univ. Michigan, USA (cjablono@umich.edu)
Lai-Yung (Ruby) Leung, PNNL, USA, Leung, (Ruby.Leung@pnnl.gov)

Websites

http://www.wcrp-climate.org/index.php/modelling-wgcm-mip-catalogue/modelling-wgcm-mips/429-wgcm-hiresmip

https://dev.knmi.nl/projects/highresmip/wiki

Goal of HighResMIP

For the first time, we want to assess 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 with constrained aerosol forcing.
Recent simulations with global high-resolution climate models have demonstrated the added value of enhanced resolution compared to the output from models in the CMIP3 and CMIP5 archive. They showed significant improvement in the simulation of aspects of the large scale circulation such as such as El Niño Southern Oscillation (ENSO) (Shaffrey et al 2009), Tropical Instability Waves (Roberts et al 2009), the Gulf Stream and its influence on the atmosphere (Chassignet and Marshall 2008; Kuwano-Yoshida et al 2010), the global water cycle (Demory et al. 2014), extra-tropical cyclones and storm tracks (Hodges et al. 2011) and Euro-Atlantic blocking (Jung et al 2012). In addition, the increased resolution enables more realistic simulation of small scale phenomena with potentially severe impacts such as tropical cyclones (Zhao et al. 2009), tropical-extratropical interactions (Haarsma et al. 2013) and polar lows. Other phenomena that are sensitive to increasing resolution are ocean mixing, sea-ice dynamics and monsoons. The improved simulation of climate also results in better representation of extreme events such as heat waves, droughts and floods.

The requirement for a multitude of multi-centennial simulations, including poorly constrained Earth System processes and feedbacks, has meant that model resolution within CMIP has progressed very slowly. In CMIP3 the typical resolution was 250km in the atmosphere and 1.5° in the ocean, while more than seven years later in CMIP5 this had only increased to 150km and 1° respectively. Until now high-resolution simulations have been performed at only a few research centers without overall coordination. Due to the large computer resources needed for these simulations, synergy will be gained if these runs are done in a coordinated way, which enables the construction of a multi-model ensemble (since ensemble size for each model will be limited) with common integration periods, forcing and boundary conditions. The CMIP3 and CMIP5 data bases provide outstanding examples of the success of this approach. The multi-model mean has proven often to be superior to individual models in seasonal and decadal forecasting. Moreover, significant scientific understanding has been gained from analyzing the inter-model spread and attempting to attribute to model formulation.

The primary goal of HighResMIP is to investigate the robustness across a multi-model ensemble of changes to the representation of climate processes as model horizontal resolution is increased (with, as far as possible, few if any other changes to the model at different resolutions). Specifically the top priority CMIP6 broad question for HighResMIP is “What are the origins and consequences of systematic model biases”, which will focus on understanding model error (applied to mean state, variability and teleconnections), via process-level assessment, rather than on climate sensitivity. This has motivated our choices in terms of proposed simulations, which emphasize sampling the recent past and the next few decades where climate variability is the more important factor than climate sensitivity. From the point of view of processes, a focus will be on the analysis of atmospheric eddies, including tropical and extra-tropical cyclones. Of particular value is the variability of hurricanes (and emerging process at high-resolution) in the North Atlantic basin.

HighResMIP will coordinate the efforts in the high-resolution modeling community. Joint analysis, based on process-based assessment and seeking to attribute model performance to emerging physical climate processes (without the complications of Earth System feedbacks) and sensitivity of model physics to model resolution, will further highlight the impact of enhanced resolution on the simulated climate. As a result of the widespread impact of resolution on the simulation of the climate, HighResMIP will contribute to all of the five grand challenges of the WCRP, and hence such analysis may begin to reveal at what resolution particular processes can be robustly represented.

HighResMIP will contribute an article to the GMD CMIP6 special issue as required by CMIP6. In addition HighResMIP will contribute to a synthesis paper in collaboration with the other modeling groups, focusing on the potential for cross-exploitation of simulations.

The European institutes in Table I are participating in the European H2020 project PRIMAVERA that coordinates the simulation and analyses of high-resolution runs. If HighResMIP is endorsed they will follow that protocol. Presently the following modelling centers plan to contribute Tier 1 simulations to HighResMIP:

- EC-Earth consortium, (KNMI, IC3, SMHI)
- Met. Office, UK
- CMCC, Italy
- GFDL, USA
- CNRM, France
- MRI, Japan
- MPI, Germany
- AWI, Germany
- CPTEC, Brazil
- IAP, China
- JAMSTEC, Japan
- NICAM, Japan
- BCC, China

In addition institutes that are not able to undertake the HighResMIP simulations currently due to limited computer resources, such as the ARC Centre of Excellence for Climate System Science (Australia) have expressed their strong interest in analyzing the HighResMIP simulations.

Institution	MO/NCAS/ NOCS	KNMI/SMHI/ IC3/CNR	CERFACS	MPI	CMCC	ECMWF	AWI
Model names	UM / NEMO	ECEarth / NEMO	Arpege / NEMO	ECHAM / MPIOM	CCESM / NEMO	IFS / NEMO	ECHAM/ FESOM
Atmospheric resolution	60-25	T239-T799	T359	T255	25km	T239-T799	T255
Oceanic resolution	¼-1/12^{o}	¼^{o}-1/12^{o}	¼-1/12^{o}	¼-1/10^{o}	¼	¼	¼ - 1/12 spatially variable

Table I: European institutes, together with the models and the resolutions that are committed to HighResMIP (note that the eddy resolving 1/10-1/12˚ ocean may be used for a small subset of simulations).

Proposed Experiments

The main experiments will be divided between Tier 1, 2 and 3.

The Tier 1 experiments will be historical AMIP runs. A few institutes have already performed high resolution AMIP runs and published their results. These runs will not impose such prohibitively large technical difficulties and it is feasible for a considerable number of institutes to deliver a coordinated and coherent set of experiments. Restricting the AMIP runs to the historical period make these runs also interesting for NWP centers and reduces the cost of the core simulations.

For the coupled experiments the situation is somewhat different. Although a few institutes already have carried out high resolution coupled simulations, there still remain issues with for instance biases and spin-up. Due to these issues and the large amount of computer resources needed, only a limited number of institutes will be able to afford these coupled simulations, and hence they will be done in Tier 2.

Future AMIP simulations will be done in Tier 3. This requires the construction of SST and sea-ice data sets that match the increased resolution of the HighResMIP models. Although the future period covers the entire present century, the simulations can for computational reasons be restricted to mid-century.

Standard CMIP6 resolution experiments
To evaluate the impact of increased resolution the experiments in Tier 1, 2 and 3will be repeated with the standard CMIP6 resolution. The experimental set-up and design of the standard resolution experiments will be exactly the same as for the high-resolution runs. This enables the use of HighResMIP simulations for sensitivity studies investigating the impact of resolution.

- **Tier 1**

AMIP runs

Resolution: *minimum 25-50 km at mid-latitudes for high resolution + a standard resolution configuration.*
This resolution is significantly higher than used in CMIP5. Multi-decadal to century integrations for this resolution are now feasible.

Period of integration: *Near Past 1950-2014*

This period is longer that the DECK AMIP run that starts in 1979. This is motivated primarily by work in many groups looking at climate variability over multi-decadal timescales with different phases of climate modes such as AMO, PDO, as well as improved sampling of ENSO teleconnections. The longer period will also improve the robustness of assessing the difference in variability between standard and higher resolution simulations, as well as being important for statistics of teleconnections (e.g. Rowell et al, 2013). The longer period of integration (as compared to ensembles of integrations over a shorter period) will also give a much more robust assessment of the ability of models to respond to known modes and their phases of variability, which is clearly a big issue for climate risk when we consider the effect of the global warming signal in addition to future natural variability.

Starting from 1870 though valuable to better compare with GMMIP, would not be affordable for many groups. In addition, the quality of observational/reanalysis datasets to the earlier period, to assess the modelled variability and processes, must be in question.

Forcing: Historical green house gases and aerosols, natural and anthropogenic. For SST and sea-ice a ¼ degree daily HadISST data set will be used (Rayner et al, in prep.)

Minimum number of integrations: *1*
Any manageable number is too low for a rigorous estimate of the internal variability. However, because the aim of the high-res protocol is to perform simulations at the highest possible resolution, the ensemble size has to be kept low. By using a strictly common protocol that is followed by many institutes, the effective multi-model ensemble will be much larger, enabling a much wider sampling than previously of the multi-model robustness of resolution impacts. In addition, if models can be proven to be portable, the ensemble size could be increased if other computer resources are available (discussions are already underway with the European PRACE supercomputing infrastructure). Some centers may be able to produce much larger ensembles, enabling a more robust estimate of internal variability.

- **Tier 2**

Coupled runs

The coupled runs will consist of pairs of both scenario (historic for the past) runs and, for comparison, control runs using fixed 1950s forcings. This will allow an evaluation of the model drift in addition to the climate change signal. The coupled simulations are also aimed to address questions of model bias in both mean state and variability similar to the AMIP simulations. There are many examples from previous studies (e.g. Scaife et al, 2011) where these biases become much more evident in the coupled context compared to the forced atmosphere simulations.

Resolution: *Atmosphere same as AMIP-runs. Ocean ~0.25 degree.*
This enables the ocean to have some variability (compared to non-eddy permitting models), particularly in the tropics, and has been shown to change the strength of atmosphere-ocean interactions (Kirtman et al, 2012).

Period of integration: 1950-2050
The past period is chosen to be able to compare with the Tier 1 AMIP simulations. The future end-date is based on a compromise between what is computationally affordable by a sufficient number of centers (~100 years of integration) and what is scientifically relevant. In addition, this period contains several phases where the surface warming rate is lower, and the multi-model multi-resolution ensemble may give some insight into this.

We again emphasize our interest in model error (bias, climate processes and variability) rather than climate sensitivity within these coupled simulations. Of course the number of ensemble members that will be possible, at least initially, in HighResMIP will not be sufficient to fully address internal variability, but it will form an important baseline set of simulations from which we can begin to learn, and should be useful for many of the other CMIP6 MIPs (e.g. DCCP, GMMIP, CORDEX, CFMIP).

Forcing: *CMIP6 scenario's*
CMIP6 scenarios that span the range from middle to high end scenarios. For the historical period all forcings, natural and anthropogenic, will be included.
To reduce the inter-model uncertainty the simulations should have the same aerosol forcing as far as possible. Due to the different approaches of aerosol modelling in different groups this is a challenging task. For optimal comparison between the models aerosol concentrations would be preferable and not emissions.

In discussion with RFMIP members there seem to be two upcoming developments with different timescales that could be used:

1. A method to prescribe anthropogenic optical depth and relative changes in cloud effective radius within the model radiation code, scaled according to emissions, in addition to natural aerosols as modelled or prescribed by each model natively. This should mean that all models see very similar aerosol forcing. This is a well advanced proposed protocol.
2. Time-varying dataset of mass mixing ratios for the main aerosol species, which models can then distribute over different modes and allows for other model parameterizations. This work is ongoing but may need datasets on emissions due late 2015-early 2016.

Other possibilities include:

a. Each modeling centre to use aerosol concentrations derived from their DECK simulation at standard resolution, the same concentrations to be used at both standard and high resolution
b. Models to use aerosol emissions as in the DECK.

However, the aerosol forcing choice will be a strong recommendation, rather than a specification, of the protocol. Similar as for the standard CMIP6 DECK, groups can choose to use aerosol emissions or prescribed concentrations

The full details of the forcing datasets and strategy proposed can be found from the WCRP website (http://www.wcrp-climate.org/index.php/modelling-wgcm-mip-catalogue/429-wgcm-hiresmip)..

Minimum number of integrations: *1 for each of control and historic forcings*
Ideally the ensemble number would be of order 3 simulations for each forcing, to help in evaluating model drift and enabling an improved sampling of internal variability, but this will quickly become very onerous on computing.

Coupling: *Minimal daily coupling between ocean and atmosphere. Preferably more frequent, 3hr or 1hr.*
Ocean-atmosphere interaction occurs on all time scales. With 3hr or 1hr the diurnal time scale can be resolved.

Initial state: Due to limited computer resources an equilibrated initial ocean state in the same way as the DECK experiments is not feasible. Discussions of methods for initializing the coupled models are ongoing. The current proposed methodology is to use the 1950 EN3 ocean analysis as the ocean initial condition, and then run the coupled model (with 1950's fixed forcing) until some criteria of convergence is satisfied (e.g. upper ocean drift or Top of Atmosphere radiation balance), at which point the historic run will be branched off, with the control 1950's run continuing. Hence it is the difference between these two simulations that is relevant, and this should remove the main part of the drift. Given the relatively short period into the future of the integrations, full convergence of the deep ocean is not essential for the main scientific goals of HighResMIP.

There are alternatives to this, including using the ocean state from the DECK historic simulation at standard resolution at 1950 – but these provide their own problems, such as how to initialize a consistent state with a higher resolution ocean. Interpolating the standard resolution ocean to higher resolution is rather unsatisfactory, since any biases due to the low resolution will be baked in.

- **Tier 3**

AMIP runs

Period of integration: 2015-2050 (2100)
As Tier 1 but extending the integration period to the future period (at least 2050, with an option to continue to 2100). For comparison with the coupled integrations in Tier 2 and due to limited computer resources the simulations can be restricted to 2050.

As regards time-slice, and having a stronger signal to noise – the discussion within the group will continue to decide what may be appropriate and most useful/comparable with the DECK experiments

Forcing:

- Same scenario and aerosol concentration as for the coupled runs in Tier 2.

- In consultation with the HadISST dataset development group, a method of extending the HadISST-based SST forcing from 2014-2050 is currently being developed. This will involve finding a period in the past with similar phases of the major climate modes as 2014, use an anomaly framework to add this to the SST forcing at 2014, and then retrend the dataset and impose the daily, ¼ degree variability as used in more recent HadISST work. The sea-ice forcing will then be based on this SST. This is a much simpler procedure than attempting to use an ensemble of coupled model outputs, leading to a simplified but realistic potential future scenario.

Although these simulations do not have an equivalent in the DECK, we consider that they have many uses, particularly given their main advantage over coupled simulations in that the biases are usually much less pronounced, and it is cheaper to run multiple ensembles and hence extract the signal-to-noise.

Various science questions include:

1. They may allow for a better estimation of the now occurring changes in extreme events (for instance related to the hydrological cycle), and will potentially be useful for informing regional climate adaptation and other users of climate output such as infrastructure investments that have a time horizon of 30 years or so.
2. There are several studies regarding detection and attribution of changes of extreme that would benefit from having experiments up to 2050 and for this shorter-term period the exact climate scenario chosen is not such a significant factor. Although the ensemble size of any one model will be small, it can be added to over time, and the multi-resolution multi-model ensemble can be a starting point.
3. The same thing applies to the time of emergence studies: many studies show time of emergence (ToE) now or in the next few decades (depending on the variable and regions of course) – e.g. Hawkins and Sutton (2012). There are also suggestions that ToE can be easier to determine if one looks at extreme/intense events rather than mean quantities. It seems reasonable to assume that having high-resolution simulations could help to achieve this for precipitation-related events.

Another potential use of these simulations is to give a baseline for near-term decadal predictions of the forced response only (using the best estimate of the SST forced response and the RCP/SSP radiative forcing). This can then be combined with coupled decadal predictions (or statistical modelling) which also include the ocean variability and its influence. See for instance Hoerling et al (2011) as a first attempt to do this with low-resolution models.

Connection with DECK

The DECK simulations will serve as an entry card for the CMIP6-Endorsed MIPS. The DECK experiments are

- AMIP simulations
- Pre-industrial control simulations

- 1%/yr increase in CO2 concentration
- Switch-on 4XCO2

Completion of the DECK simulations using the standard model resolution used in HighResMIP will serve as the entry card for HighResMIP. For Tier 2 the other three coupled simulations of DECK with the standard CMIP6 resolution will serve as an entry card. This applies also to the CMIP6 Historical simulation which consists of a historical simulation from 1850-2014 using specific forcings consistent with CMIP6. For the high-resolution simulations the DECK is too expensive in computer resources, but the comparison between the standard resolution simulations within HighResMIP and the DECK simulations will be informative in themselves. The premise is that the higher resolution model can be treated as a sensitivity study, particularly since it is not currently affordable by any group to run a full set of DECK simulations at a sufficiently high resolution to give a meaningful response. The relevance of HighResMIP is that the significant step in resolution can begin to clarify some of the outstanding climate science questions remaining from CMIP3 and CMIP5 exercises, and also to bridge the gap of what is scientifically interesting and what is useful for the society and decision makers on decadal timescales. It is assumed that these benefits will outweigh the lack of DECK simulations.

Connection with other MIPS

GMMIP for global monsoons.
There is known sensitivity to monsoon flow and rainfall with model resolution in the West African monsoon, Indian monsoon (particularly via monsoon depressions) and possibly East Asian monsoon. As stated in GMMIP the monsoon rainbands are usually at a maximum width of 200 km. Climate models with low or moderate resolutions are generally unable to realistically reproduce the mean state and variability of monsoon precipitation for the right reasons. This is partly due to the model resolution. The Tier 1 AMIP runs of HighResMIP will be used in Task-4 of GMMIP to examine the performance of high-resolution models in reproducing both the mean state and year-to-year variability of global monsoons.

SensMIP for parameter sensitivity
It is unclear how much the experimental design in SensMIP and HighResMIP overlap or complement each other. The multi-model high resolution ensemble could give one axis of uncertainty/variability from models, while a corresponding parameter sensitivity study would explore a different axis, but the limited number of parameters proposed to change in SensMIP may limit its use here.

CORDEX and **GGDEX** for downscaling
Collaboration with these will be sought. HighResMIP can provide boundary conditions for downscaling and provide the stimulus to cloud resolving simulations.

OMIP for ocean analysis and initial state
It will be investigated if OMIP can provide the equilibrated ocean initial conditions for the coupled runs and exchange diagnostic/analysis techniques to understand ocean circulation changes at different resolutions.

Grand Challenges

HighResMIP will address the grand challenges of the WCRP in the following way

Clouds, Circulation and Climate Sensitivity (Rank 1)
HighResMIP will address this Grand Challenge in many different ways. The sensitivity of increasing resolution on water vapor loading, cloud formation, circulation characteristics and climate sensitivity will be investigated.

To improve the robustness of our understanding, the multi-model ensemble at different resolutions, together with the longer period AMIP integrations, will allow us to:

- (i) link tropospheric circulation to changing patterns of SSTs, land-surface properties, and understanding the role of cloud processes in natural variability
- (ii) examine the extent and limits of our understanding of patterns of precipitation
- (iii) examine changes in model biases (such as humidity) with resolution, since there are some indications that these may be linked to climate sensitivity

Increasing resolution affects in particular small scale process such as the formation of clouds. Although the formation of clouds has still to be parameterized in the resolution of HighResMIP the dynamical constraints for the formation of clouds, such as the location and magnitude of upwards and downwards motion, as well as moisture availability, are sensitive to resolution. This also applies to the response of the circulation to cloud formation.

Cryosphere in a Changing Climate (Rank 5)
Because in the Tier 1 experiments the sea-ice distribution is prescribed the contribution to this grand challenge is limited. Its main impact will be on the distribution of snow fall and subsequent accumulation and melting of snowpack that affect land surface hydrology. For instance the occurrence of intense polar systems, such as deep polar-lows that are accompanied by abundant snowfall will be better represented with increasing resolution.
In the Tier 2 coupled simulations the historic simulation will affect the growth of sea-ice and the air-sea heat flux, processes that are strongly affected by small scale processes. Here we can study the effect of model resolution on Arctic sea-ice variability, and possible influences on mid-latitude circulation.

Understanding and Predicting Weather and Climate Extremes (Rank 3)
HighResMIP is strongly related to this grand challenge. Increasing resolution of climate models will bring us closer to the ultimate goal of seamless prediction of weather and climate. Extremes mostly occur and are driven by processes on small temporal and spatial scales that are not well resolved by standard CMIP6 climate models. Dynamical down scaling only partially resolves this due to the non-linear interaction between large and small spatial scales and the importance of representing global teleconnection patterns. We aim to improve our understanding of the interaction between global modes of variability (e.g. ENSO, NAO, PDO) and regional climate inter-decadal variability and extremes.

Regional Climate Information (Rank 4)
Regional climate information focuses on smaller scales and extreme events, which are relevant for stakeholders and adaptation strategies. This requires high resolution modeling to provide reliable information. Recent high resolution modeling studies (Di Luca et al. 2012; Bacmeister et al. 2013) and comparisons of CMIP3 and CMIP5 results (Watterson et al. 2014) have demonstrated the added value of increased resolution for regional climate information. Model outputs from HighResMIP could also be used by the regional climate modeling community for comparison of dynamical downscaling and global high resolution approaches and for further downscaling by cloud resolving regional models.

Sea-Level Rise and Regional Impacts (Rank 6)
For Tier 1 simulations there is no contribution to this grand challenge. For Tier 2 the contribution is limited although there is the potential for large contribution. If for instance the deep water formation and MOC response appears to be highly sensitive to resolution than there is a considerable impact on regional sea level rise. In addition resolving the topographic effect at high-resolution should have profound impacts on regional details about the sea level rise that are relevant for policy making and planning.

Changes in Water Availability (Rank 2)
HighResMIP is very relevant to this grand challenge. Resolution affects the hydrological cycle by modifying the land/sea partitioning of precipitation. Increasing resolution in general increases the moisture convergence over land (Demory et al. 2014) although regionally this can be reversed such for instance in Europe during the winter due to changes in the position of the storm track (Van Haren et al. 2014). In addition simulation of extreme precipitation events are highly sensitive to increasing resolution. How robust are these results across the multi-model ensemble? Can higher resolution models help to give insight into inconsistencies between global precipitation and energy balance datasets?

Biospheric forcings and feedbacks (Rank 7)
There is no direct link to this collaboration theme as the biosphere is not explicitly modelled. Because the response of the biosphere depends critically on the accurate simulation of the physical environment there is potential for spin-off studies, for instance by interpreting diagnostic information about vegetation production. Recycling of water is an important aspect of biospheric forcings and feedbacks, and the way that vegetation responds to drying depends on their role in recycling water - given the small scales of the involved processes this is strongly affected by model resolution.

GEWEX

HighResMIP fits in the GEWEX research focus of “Develop accurate global model formulation of the energy and water budget and demonstrate predictability of their variability and response to climate forcing”. Accurate modeling of the energy and water budget is sensitive to the adequate simulation of the energy conversions and phase transitions as well as the transport that occur on small spatial scales.

Overview of the proposed evaluation and analysis

The analysis will focus on the impact of increasing resolution on the simulation of the climate. The robustness of the impact of increasing resolution on the simulation of these phenomena among the different HighResMIP models will be investigated and their response to global warming assessed. One of the primary strengths of the simple experimental design for HighResMIP is that it enables a wide range of process-based analysis –simulation campaigns which included 1-2 models such as UPSCALE (Mizielinski et al. 2014) and Athena (Kinter et al. 2013) already have an extremely active number of analysis projects associated with them and insightful papers.

The increased resolution will enable a better simulation of regional climates. The analysis will therefore also have a focus on regional climate such as for instance Latin America.

The results of the analysis of HighResMIP will be compared with the CMIP6 DECK experiments. Their experimental design, data format and documentation will follow the DECK experiments as far as possible.

The storage and distribution of the high resolution model data is a challenging issue that requires further discussion within HighResMIP. In PRIMAVERA the JASMIN platform (STFC/CEDA, UK) will be used for data exchange and as a common analysis platform. Because the resolution of HighResMIP approaches the scales necessary for realistic simulation of weather, daily and sub-daily data will be stored to allow the investigation of weather phenomena including those related to monsoons and tropical climate.

HighResMIP output data will conform to all the CMIP requirements. However, given the data volumes from HighResMIP, it will not be possible to transmit and store all the required output data on an ESGF node (and indeed the volumes will mitigate against downloading multi-model data to individual centres for analysis), and hence a different methodology is needed to cope. Discussions with other international data centres are planned to further enable collaborative analysis. The European Copernicus Data Climate Store may also provide useful future avenues for data storage and sharing, which will be explored.

One useful approach may be to provide spatially and/or temporally coarsened model output on the ESGF, which would enable initial analysis compared to DECK simulations, and indicate which avenues of analysis may require full model resolution output, which remaining manageable volumes. It would then also be available for any automated assessment tools on the ESGF. This will be further discussed with the HighResMIP partners.

Proposed timing

Start of the experiments: Beginning of 2016
End of the experiments: No fixed date.

References

Bacmeister, J.T., M.F. Wehner, R.B. Neale, A. Gettelman, C.E. Hannay, P.H. Lauritzen, J.M. Caron, and J.E. Truesdale, 2014: Exploratory high-resolution climate simulations using the Community Atmosphere Model (CAM). Journal of Climate, 27, 3073–3099, DOI: 10.1175/JCLI-D-13-00387.1.

Bell, R.J., J. Strachan, P.L. Vidale, K.I. Hodges and M. Roberts, 2013: Response of tropical cyclones to idealized climate change experiments in a global high resolution coupled general circulation model. J. Climate, 26 (20), 7966-7980.

Chassignet, E.P., and D. P. Marshall, 2008: Gulf stream separation in numerical ocean models. In: Hecht, M., Hasumi, H. (eds.), Eddy-Resolving Ocean Modeling, AGU Monog. Ser., 39-62.

Demory, M.-E., et al. (2014). The role of horizontal resolution in simulating drivers of the global hydrological cycle. Clim. Dyn., 42, 2201-2225.

A Di Luca, R de Elía, R Laprise, 2014: Potential for added value in precipitation simulated by high-resolution nested Regional Climate Models and observations. Clim. Dyn., 38, 1229–1247. DOI 10.1007/s00382-011-1068-3

Haarsma, R.J., W. Hazeleger, C. Severijns, H. de Vries, A. Sterl, R. Bintanja, G.J. van Oldenborgh and

H.W. van den Brink, 2013: More hurricanes to hit Western Europe due to global warming. Geophys. Res. Lett. doi:10.1002/grl.50360.

Hawkins, E. and Sutton, R. (2012) Time of emergence of climate signals. Geophysical Research Letters, 39 (1). L01702. ISSN 0094-8276 doi: 10.1029/2011GL050087

Hertwig, E., J.-S. von Storch, D. Handorf, K. Dethloff, I. Fast and T. Krismer, 2014: Effect of horizontal resolution on ECHAM6-AMIP performance. Clin. Dyn. (in press).doi:10.1007/s00382-014-2396-x

Hoerling, M., J. Hurrell, A. Kumar, L. Terray, J. Eischeid, P. Pegion, T. Zhang, X. -W. Quan, and T. Y. Xu, 2011: On North American Decadal Climate for 2011-2020. J. Climate: Volume 24, Issue 16 (August 2011) pp. 4519-4528. doi: 10.1175/2011JCLI4137.1

Hodges, K. I., et al., 2011: A comparison of extratropical cyclones in recent re-analyses ERA-Interim, NASA MERRA, NCEP CFSR, JRA-25. J. Clim., 24, 4888-4906.

Jung, T., et al., 2012: High-resolution global climate simulation with the ECMWF model in project Athena: Experimental design, model climate, and seasonal forecast skill. J. Clim., 25, 3155-3172.

Kinter, III, J. L., B. Cash, D. Achuthavarier, J. Adams, E. Altshuler, P. Dirmeyer, B. Doty, B. Huang, E. K. Jin, L. Marx, J. Manganello, C. Stan, T. Wakefield, T. Palmer, M. Hamrud, T. Jung, M. Miller, P. Towers, N. Wedi, M. Satoh, H. Tomita, C. Kodama, T. Nasuno, K. Oouchi, Y. Yamada, H. Taniguchi, P. Andrews, T. Baer, M. Ezell, C. Halloy, D. John, B. Loftis, R. Mohr, and K. Wong, 2013: Revolutionizing Climate Modeling with Project Athena: A Multi-Institutional, International Collaboration. Bull. Amer. Meteor. Soc., 94, 231–245. doi: http://dx.doi.org/10.1175/BAMS-D-11-00043.1

Kirtman, B. P., et al, 2012: Impact of ocean model resolution on CCSM climate simulations. 39, 1303-1328. http://rd.springer.com/article/10.1007/s00382-012-1500-3.

Kuwano-Yoshida A, Minobe S, Xie S-P, 2010: Precipitation response to the Gulf Stream in an Atmospheric GCM. J. Clim 23, 3676–3698.

Mizielinski, M. S., et al, 2014: High resolution global climate modelling; the UPSCALE project, a large simulation campaign. Geosci. Model Dev., 7, 563-591.

Mizuta, R., Y. Adachi, S. Yukimoto, and S. Kusunoki, 2008: Estimation of the future distribution of sea surface temperature and sea ice using the CMIP3 multi-model ensemble mean, Tech. Rep. 56, 28 pp., Meteorol. Res. Inst., Tsukuba, Japan.

Roberts, M. J., et al., 2009: Impact of Resolution on the Tropical Pacific Circulation in a Matrix of Coupled Models. J. Clim., 22, 2541-2556.

Rowell, D. P. 2013: Simulating SST Teleconnections to Africa: What is the State of the Art?. J. Climate, 26, 5397–5418.doi: http://dx.doi.org/10.1175/JCLI-D-12-00761.1

Shaffrey, L. C., et al., 2009: U.K. HiGEM: the new U.K. High-Resolution Global Environment Model-

model description and basic evaluation. J. Clim., 22,1861-1896.

Scaife, A. A., D. Copsey, C. Gordon, C. Harris, T. Hinton, S.J. Keeley, A. O'Neill, M. Roberts and K. Williams 2011: Improved Atlantic Blocking in a Climate Model. Geophys. Res. Lett., 38, L23703, doi:10.1029/2011GL049573.

Van Haren, R., R.J. Haarsma, G.J. van Oldenborgh and W. Hazeleger, 2014: Resolution dependence of European precipitation in a state-of-the-art atmospheric general circulation model. J. Clim. DOI: 10.1175/JCLI-D-14-00279.1

Watterson, I.G., J. Bathols, and C. Heady, 2014: What Influences the Skill of Climate Models over the Continents?. Bull. Amer. Meteor. Soc., 95, 689–700. doi: http://dx.doi.org/10.1175/BAMS-D-12-00136.1

Zhao, M., et al., 2009: Simulations of Global Hurricane Climatology, Interannual Variability, and Response to Global Warming Using a 50km Resolution GCM. J. Climate, 33, 6653-6678.

Ice Sheet Model Intercomparison Project for CMIP6 (ISMIP6) Application for CMIP6-Endorsed MIPs

Revision Date: 10 June 2015

This MIP has two independent parts. The first uses only AOGCM diagnostic output to force standalone models of the Greenland and Antarctic ice sheets, while the second involves coupled AOGCM-ice sheet model and is focused on Greenland. Endorsement of the former does not imply a contribution to the latter.

Proposals from MIPs should include the following information:

** Preliminary information used to determine whether a MIP should be endorsed for CMIP6 or not.*
*** Information that must be provided later (and before the panel can determine which experiments, if any, will be incorporated in the official CMIP6 suite).*

- Name of MIP*

ISMIP6: Ice Sheet Model Intercomparison Project for CMIP6

- Co-chairs of MIP (including email-addresses)*

Eric Larour, NASA Jet Propulsion Laboratory, USA, eric.larour@jpl.nasa.gov
Sophie Nowicki, NASA Goddard Space Flight Center, USA, sophie.nowicki@nasa.gov
Tony Payne, University of Bristol, UK, a.j.payne@bristol.ac.uk

- Members of the Scientific Steering Committee*

Helene Seroussi, NASA Jet Propulsion Laboratory, USA,
Heiko Goelzer, Vrije Universiteit Brussel, BE,
Andrew Shepherd, University of Leeds, UK,
William Lipscomb, Los Alamos National Laboratory, USA,
Jonathan Gregory, University of Reading and Met Office Hadley Center, UK,
Ayako Abe-Ouchi, The University of Tokyo, JP

- Link to website (if available)*
  http://www.climate-cryosphere.org/activities/targeted/ismip6/about

- Goal of the MIP and a brief overview*

The primary goal of ISMIP6 is to improve projections of sea level rise via improved projections of the evolution of the Greenland and Antarctic ice sheets under a changing climate, along with a quantification of associated uncertainties (associated with both uncertainty in climate forcing and in the response of the ice sheets). As depicted in Figure 1, this goal requires an evaluation of AOGCM climate over and surrounding the ice sheets; analysis of simulated ice-sheet response from standalone models forced “offline” with CMIP AOGCM outputs and, where possible, with coupled ice sheet-AOGCM models; and experiments with standalone ice sheet models targeted at exploring the uncertainty associated with ice sheets physics, dynamics and numerical implementation. A secondary goal is to investigate the role of feedbacks between ice sheets and climate in order to gain insight into the impact of increased mass loss from the ice sheets on regional and global sea level, and of the implied ocean freshening on the coupled ocean-atmosphere circulation.

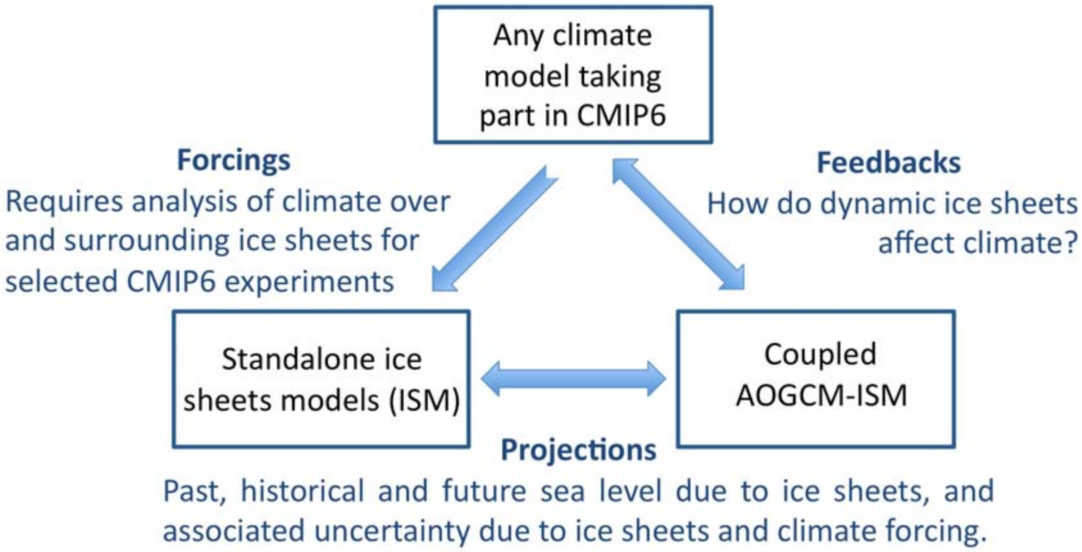


Figure 1: Overview of the ISMIP6 effort.

ISMIP6 is directly related to the WCRP Grand Challenges on ‘Changes in the Cryosphere’ and ‘Regional Sea-level Rise’. A white paper on the former identifies the need for “a focused effort on developing ice sheet models, with specific emphasis on the role of ice sheet dynamics on the rate of the sea-level rise”, which ISMIP6 is ideally placed to deliver by linking the improved process-based understanding achieved within the WCRP Climate and Cryosphere (CliC) project, and elsewhere, to projections of future ice-sheet mass budget. While a white paper on the latter identifies several open issues that strongly relate to our proposed activity, including the need to understand the ocean’s response to high latitude freshwater forcing and the impact of ice sheet dynamics. ISMIP6 is primarily focused on the CMIP6 scientific question “How does the Earth System respond to forcing?” and offers the exciting opportunity of widening the current CMIP definition of Earth System to include (for the first time) the ice sheets. The emphasis on standalone, ensemble modeling will also shed light on the question “How can we assess future climate changes given climate variability, predictability and uncertainties in scenario” for the mass budget of the ice sheets and its impact of global sea level.

ISMIP6 is a targeted activity of the CliC project. At the time of writing (31st May 2015), ISMIP6 has received an expression of interest from eleven modeling groups (CanESM, CESM, CNRM, EC-Earth, GFDL, GISS, INM, IPSL, MIROC-ESM, MPI-ESM, and UKESM). We will report on the results obtained from ISMIP6 via peer review publications and/or presentations, and modeling groups will be invited to become co-authors.

➢ References (if available)*

ISMIP6 is based on a long history of Ice Sheet Model Intercomparison Projects (ISMIP http://homepages.vub.ac.be/~phuybrec/ismip.html) and the more recent ice2sea (www.ice2sea.eu), Sea level Response to Ice Sheet Evolution (SeaRISE http://websrv.cs.umt.edu/isis/index.php/SeaRISE_Assessment), and COMBINE (https://www.combine-project.eu/) efforts. ISMIP6 brings together for the first time a consortium of international ice sheet models and coupled ice sheet-climate models to fully explore the sea level rise contribution from the Greenland and Antarctic ice sheets. Papers generated by these recent activities, that involved the ice-sheet modeling community, include:

- Bindschadler, R. et al. (2013) Ice-Sheet Model Sensitivity to Environmental Forcing and Their Use in Projecting Future Sea levels (The SeaRISE Project). Journal of Glaciology, 59 (214), 195-224.
- Edwards, T.L., et al. (2014) Effect of uncertainty in surface mass balance elevation feedback on projections of the future sea level contribution of the Greenland ice sheet. The Cryosphere 8(1), 195-208.
- Favier, L., et al. (2014) Retreat of Pine Island Glacier controlled by marine ice-sheet instability. Nature Clim. Change 4(2), 117-121.

- Nowicki et al. (2013) Insights into spatial sensitivities of ice mass response to environmental change from the SeaRISE ice sheet modeling project I: Antarctica. Journal of Geophysical Research- Earth Surface, 118 (2), 1002-1024.
- Nowicki et al. (2013) Insights into spatial sensitivities of ice mass response to environmental change from the SeaRISE ice sheet modeling project II: Greenland. Journal of Geophysical Research- Earth Surface, 118(2), 1025-1044.
- Pattyn, F., et al. (2013) Grounding-line migration in plan-view marine ice-sheet models: results of the ice2sea MISMIP3d intercomparison. Journal of Glaciology 59(215), 410-422.
- Rae, J., et al. (2012) Greenland ice sheet surface mass balance: evaluating simulations and making projections with regional climate models. The Cryosphere 9(6), 1275-1294.
- Shannon, S.R., et al. (2013) Enhanced basal lubrication and the contribution of the Greenland ice sheet to future sea level rise. Proceedings of the National Academy of Sciences 110(35), 14156-14161.
- Shepherd, A., et al. (2012) A Reconciled Estimate of Ice-Sheet Mass Balance. Science 338(6111), 1183-1189.

➢ An overview of the proposed experiments*

The overall framework for ISMIP6 is designed to deliver projections of the ice sheet contribution to sea level rise. Together with the CliC targeted activity GlacierMIP and projections of thermal expansion (that already sit within the CMIP framework), this will allow sea level to become part of the family of variables for which CMIP can provide routine IPCC-style projections. The proposed experiments will both use and augment the CMIP6-DECK, Historical, ScenarioMIP and PMIP experiments, as summarized in Table 1. ISMIP6 will use the standard CMIP AGCM and AOCGM experiments for analysis of the climate over and surrounding the ice sheets, and as forcing for the standalone ice sheet models (ISM) projections. Additional sensitivity experiments will be performed with the ISMs to investigate the uncertainty associated with these projections arising from ice sheet models. The key output will be an ensemble of past and future estimates of ice sheet contribution to sea level. To address the feedbacks introduced by interactive ice sheets, we propose that a small number of selected DECK experiments are repeated with coupled AOGCM-ISM, where the ice sheet is an interactive component of the AOGCM. Our assessment of the state of existing AOGCMs is that coupled models including an interactive Greenland ice sheet can realistically be expected for CMIP6, however including the Antarctic ice sheet remains challenging (because of the greater complexity of its response to climate forcing, and the issues associated with simulations of the Southern Ocean). It is for these reasons that ISMIP6 heavily relies on standalone ice sheet models driven offline by CMIP6 climate models for projections of sea level.

Existing CMIP exp. used by ISMIP6 (AGCM-AOGCM only, no dynamic ice sheet required)	Standalone ISMIP6 ice sheet model exp. (ISM only)
- AMIP simulation (*amip*) - CMIP6 Historical Simulation (*historical*) - Pre-Industrial Control (*piControl*) - 1% yr CO2 to quadrupling CO2 (*1pctCo2*) - Abrupt4xCO2 simulation (*abrupt4xCO2*) - ScenarioMIP SSP5-8.5 up to year 2300 (*ssp5-8.5*) - PMIP Last Interglacial (*lastInterglacial*)	- ISM control (*piControlforcedism*) ** - ISM for last few decades forced by *amip* (*amipforcedism*) - ISM for the historical period forced by *historical* (*historicalforcedism*) - ISM forced by *1pctCo2 (1pctCo2forcedism)* for quantification of feedback** - ISM for 21st century and up to 2300 sea level forced by ScenarioMIP *ssp5-8.5* (*ssp5-8.5forcedism*) * - ISM for Last Interglacial forced by PMIP *lastInterglacial* (*lastInterglacialforcedism*) - Other ISMIP6 specific experiments** to explore uncertainty due to ISM.
New ISMIP6 CMIP6 exp. (Coupled AOGCM-ISM)	
- Pre-Industrial Control (*piControlwithism*) ** - 1% yr CO2 to quadrupling CO2 (*1pctCo2withism*) ** - ScenarioMIP SSP5-8.5 up to year 2300 (*ssp5-8.5withism*)	

Table 1: Overview of experimental framework for ISMIP6 (further details on experimental design and motivation are explained in later sections). Name of experiments are indicated in italic. *These types of standalone ensemble ISM experiments were implemented in the European ice2sea and SeaRISE efforts for IPCC-AR5, but using forcing derived from AR4 (See www.ice2sea.eu and http://websrv.cs.umt.edu/isis/index.php/SeaRISE_Assessment) **These types

of experiments, where the ice sheet is an interactive component of the AOGCM, have been recently run as part as the European COMBINE effort (https://www.combine-project.eu/) by three modeling groups: IPSL, MIP-I, and DMI.

➢ An overview of the proposed evaluation/analysis of the CMIP DECK and CMIP6 experiments*

The primary goal of ISMIP6 is an analysis of the historical and future estimates of ice sheet contribution to sea level, and associated uncertainty, via evaluation of the ensemble simulations. For evaluating the feedbacks introduced by coupling dynamic ice sheets to AOGCM, we will compare the results of simulations of AOGCM with and without dynamic ice sheet models, and of ice sheet forced by offline with AOGCM and fully coupled to AOGCM.

This goal therefore also requires that three components of the Earth system are evaluated and analyzed by comparing to in situ, airborne and satellite observations:

1) The ice sheet dynamics. Flux gates will be defined along grounding lines at the coast, where estimated transports derived from observed surface velocities may be employed. The coupled system allows for an assessment of the total ice sheet contribution to sea level rise, which may be evaluated against GRACE data (available for 2003-present). Model performance may also be assessed with observed changes in surface elevation from ERS-1 and ERS-2 (1992-present) and ICESat (2003-2009), along with ice velocities, ground line and ice front locations.
2) The atmosphere and surface conditions over the ice sheets (surface radiative and turbulent fluxes, albedo, temperature, surface mass balance). This component may be divided into two parts: climate forcing that would be generated by an AOGCM and processes at the ice sheet surface (that may or may not be adequately simulated by an AOGCM but will be used in standalone ice-sheet models, such as SMB and albedo evolution). The evaluation of atmospheric state variables including temperature can make use of observations from established automatic weather station networks and surface radiation budget observatories at South Pole and Summit. Surface temperature and albedo may also be evaluated with remote sensing values from AVHRR (1982-present) and MODIS (2000-present). Simulated accumulation may be evaluated at in situ locations along the K-transect for Greenland, and with shallow ice cores distributed across both ice sheets. A comparison with regional climate model output and atmospheric reanalyses is also suggested as a quality test.
3) The ocean around the ice sheets (sea surface height, ocean, temperature, ocean induced melting rates, wind stress, hydrographic properties at the margins of the ice sheets to the extent available, sea-ice cover). This component may also be divided into two parts: ocean forcing that would be generated by an AOGCM and processes at the ice sheet boundary (that may or may not be captured within an AOGCM but will also be included in standalone ice-sheet models, such as ice-shelf melt). A key concern will be the validation of ocean thermal forcing of the ice sheets, which is likely to focus on evolving temperature at depth and, in particular, AOGCM simulation of the Southern Ocean.

➢ Proposed timing*

The analysis of atmospheric and oceanic climate over and surrounding the ice sheets from the CMIP5 archive will begin immediately in order to assess the quality and implied change in surface mass balance and temperatures. Analysis of the CMIP6 data would be ongoing and follow the simulation phase of CMIP6.

ISMIP6 started the design of the standalone ice sheet experiments during a workshop in July 2014, therefore further refining these experiments and data preparation would be completed by mid 2015. CliC is sponsoring an ISMIP6 workshop in summer 2015. The sea level projections and quantification of the uncertainty in sea level due to ice sheets would begin mid 2015, and continue in tandem with CMIP6. Analysis of the projection simulations and sensitivity experiments would be ongoing in order to identify the dominant sources of uncertainty.

The runs for the AOGCM-ISM simulations would occur towards the end of the CMIP6 cycle.

- For each proposed experiment to be included in CMIP6**
    - the experimental design;
    - the science question and/or gap being addressed with this experiment;
    - possible synergies with other MIPs;
    - potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, and (D) policy makers.

As summarized in Figure 1 and Table 1, the experimental design for ISMIP6, consist of three different types of modeling efforts: 1) standard AGCM-AOGCM experiments from the DECK, CMIP6 Historical and ScenarioMIP simulations (therefore all climate models participating in CMIP6 will be included in ISMIP6 without “extra work” from the climate modeling centers), 2) simulations with standalone ice sheet models, and 3) simulations with coupled AOGCM-ISMs when possible.

The detailed specifics for the ISMIP6 experiment design will be provided in a paper for the CMIP6 Special Issue. In summary, the following experimental design is proposed:

1) Use of selected standard AGCM & AOGCM CMIP6 experiments over and surrounding ice sheets:
    - *amip*: Allows the evaluation of AGCM climate over ice sheets, in particular surface mass balance (SMB: the combination of precipitation, evaporation and surface runoff).
    - Abrupt4xCO2: Allows the investigation of response timescale of AOGCM climate over and surrounding the ice sheets.
    - *historical*: Allows the evaluation of AOGCM climate over and surrounding the ice sheets for the CMIP6 historical period (1850-2014).
    - *ssp5-8.5:* standard ScenarioMIP SSP5-8.5 simulation starting from 2015 but continued to year 2300 if possible. The experiment would assess projected changes in SMB with fixed ice sheet extent and topography.
    - *lastInterglacial:* standard PMIP simulation for Last Interglacial.
    - *piControl* and *1pctCo2*: Will be used to assess the impacts of introducing dynamic ice sheets in AOGCM. For modeling groups taking part in the AOGCM-ISM experiments, the duration of the experiment need to be the same for both AOGCM and AOGCM-ISM simulations.

    Note: we would start with the existing CMIP5 output and repeat the analysis when CMIP6 output is available. The output from these experiments will be used to assess the uncertainty in sea level arising from climate forcing and to drive the standalone ice sheet models.
2) Standalone ice sheet models experiments:
    - *piControlforcedism*: ISM control, Constant forcing, needed to evaluate model drift.
    - *amipforcedism*: simulation for the last few decades to understand the well observed record of ice sheet changes. ISM would be driven by SMB anomalies obtained from the standard AMIP DECK simulations, and ice shelf basal melting or temperature anomalies from ocean models.
    - *historicalforcedism*: simulation for the historical period to understand the ice sheet contribution to 20th century GMSLR, forced by outputs obtained from the standard CMIP6 Historical simulation. The results of *amipforcedism* and *historicalforcedism* are likely to differ, and the comparison will provide some insight into the relative importance of biases, climate variability and climate change.
    - *1pctCo2forcedism*: simulation forced by 1% yr CO2 to quadrupling C02 obtained from DECK output: for comparison with the AOGCM-ISM experiment in order to evaluate ice sheet feedback.
    - *ssp5-8.5forcedism*: simulation for the 21st century (and maybe up to the 23rd century depending on ScenarioMIP) for the most realistic ice sheet contribution to sea level projections. ISM would be driven by SMB anomalies (with adjustments for ice sheet elevation change) and ice shelf mass balance or temperatures anomalies derived from the standard SSP5-8.5 ScenarioMIP simulation.

   - *lastInterglacialforcedism*: simulation of the Greenland ice sheet evolution for the Last Interglacial period (~135-115 kyr BP). Climatic forcing derived from PMIP Last Interglacial and other PMIP time slice and transient experiments (as available).
   - Additional ISM experiments would be designed to assess the uncertainty in sea level projections due to ice sheet models. These experiments would explore the model biases and uncertainties identified in the ice2sea and SeaRISE efforts, which include ice sheet initialization, poorly known basal conditions and subgrid-scale processes. In addition, ISMIP6 would investigate questions such as "How much excess oceanic heat flux is required to trigger marine ice sheet instability?" to shed light on the potential collapse of the Antarctic ice sheet.

   Note: Following the approach taken in the ice2sea and SeaRISE efforts, the anomalies derived from the DECK experiments would be added to the forcing used in the ISM control runs. The AGCM/AOGCM output is in most cases not suitable to directly force standalone ISMs, mainly due to differences in spatial resolution. A downscaling procedure (to be later specified) will be necessary to produce surface mass balance and for ice shelves basal mass balance, used to drive the ISMs. Modeling groups could decide to carry out the experiments for both the Greenland and Antarctic ice sheets, or to focus on one ice sheet. These experiments are targeted at the ice sheet community, and we envisage participation from 10-15 ice sheet models.

3) Coupled AOGCM-ISMs experiments (same set up as standard CMIP6 experiments but with evolving ice sheets models (ISM): GCM sends ISM an energy balance based SMB, and ISM sends GCM adjustments to land surface elevation and surface type)
   - *piControlwithism*: the pre-industrial control, where the aim is to produce a realistic non-drifting coupled state, and assess systematic model bias. The spin up may require the GCM and ISM to be asynchronously coupled until the system reaches quasi-equilibrium, which would be followed by a multi-hundred years run (500 yrs suggested), in order to capture unforced natural variability.
   - *1pctCO2withism*: the 1% per yr CO2 increase to quadrupling CO2 over 140 yrs and kept constant at 4xCO2 for an additional two to four centuries. This experiment, along with *piControlwithism*, are the core experiments that will be used for analysis of coupled ice sheet-climate system. Experiment would be compared to the standard DECK without ice sheets and to the standalone ISM forced by the standard DECK, in order to diagnose the strength of ice sheet-climate feedback and the associated uncertainty in projections resulting from excluding ice sheet models. Length of experiment *1pctCO2, 1pctCO2withism* and *1pctCO2forcedism* therefore needs to be the same for groups participating in this experiment. It is suggested the experiments are run for a minimum of 350 yrs and up to 500 yrs is encouraged, because results from COMBINE effort indicate that ice sheet model coupled runs start to clearly divert from the uncoupled runs after about 250-300 yrs of simulations.
   - *ssp5-8.5withism* scenario for analysis of coupled system and sea level projections from a coupled framework, which can be compared to the standalone ice sheet model projection. Experiment would cover the 21st century and preferably run out to the 23rd century. The set up would follow the set up for the standard SSP5-8.5, which may therefore first require the CMIP6 Historical simulation to be performed too with a coupled AOGCM-ISM setting.

   Note: We suggest that the pre-industrial control and 1% yr CO2 to quadrupling CO2 experiments are performed first, followed by ScenarioMIP SSP5-8.5. Modeling groups could decide to carry out the experiments for both the Greenland and Antarctic ice sheets, or to focus on one ice sheet. These experiments should only differ from the equivalent standard CMIP AOGCM setting in the manner in which the ice sheet is treated, so that the exploration of feedbacks is not affected by other changes. Feedbacks that we propose to explore include albedo-melt feedback, elevation-SMB feedback, precipitation-sea ice feedback, fresh water (runoff and icebergs calving and submarine melting)- ocean feedback, atmospheric circulation – ocean heat flux feedback (e.g. tip-jets, katabatic winds). This type of coupled experiments have been carried out by 3 modeling centers (IPSL, MPI-M, DMI) and soon MeteoFrance as part of the European COMBINE project. These coupled models (**IPSL**: IPSL – GRISLI; **MPI-M**: MPI-ESM – PISM and MPI-ESM – SICOPOLIS; **DMI**: EC-Earth – PISM, and soon CNRM-GRISLI) were described and evaluated in the COMBINE reports ‘Assessment of performance of AOGCMs coupled to Greenland and Antarctic models’ (http://www.combine-project.eu/fileadmin/user_upload/combine/dels/D4.3.pdf) and ‘Feedbacks of individual components: Cryosphere’ (http://www.combine-project.eu/fileadmin/user_upload/combine/dels/D7.7_v2.pdf, in the second part: ‘Impacts of including an interactive Greenland ice sheet in ESM). Efforts of including dynamic ice sheets into AOGCMs are also occurring with CESM, GFDL and ModelE for example, so it is expected that about 8-10 groups will be in a position to run such experiments for CMIP6.

The primary goal of these experiments is to improve sea level projections due to changes in the ice sheets, and assessing the uncertainty in these projections due to climate forcing versus that arising from ice sheet models. The secondary goal is to understand how ice sheets affect and are affected by climate. These experiments will thus shed light on the key science questions considered by CMIP6: “How does the Earth system respond to forcing?”, “What are the origin and consequences of systematic model biases”, and “How can we assess future climate change given uncertainty in scenarios?”. These goals directly contribute to the Cryosphere Grand Challenge and the Sea level Rise Grand Challenge of the Climate and Cryosphere (CliC) project and the World Climate Research Program. Finally, the ISMIP6 sea level projections will be relevant to the Impacts Adaptation and Vulnerability (IAV) community and policy makers.

Possible synergies with other MIPs include:

- High Resolution Model Intercomparison Project (HighResMIP). We will use the results from the high-resolution runs to quantify the impact of increased resolution in our standalone ice sheet suite, and to compare against the results from the DECK runs. Particular processes such as atmospheric blocking will be looked at to understand how well extreme melt-rate events are captured in our runs.
- Coordinated Regional Climate Downscaling Experiment (CORDEX). The CORDEX results will be used against the DECK runs to quantify additional sensitivities not captured in the low-resolution runs or HighResMIP runs. Biases and additional variability in the downscaled CORDEX results will be introduced in the offline ice sheet model runs.
- Land Surface, Snow and Moisture (LS3MIP). One of the objectives of LS3MIP is an evaluation of the current state of the snow cover representation in climate models, which impacts the surface mass balance over the ice sheets. These experiments with land-module may help towards understanding and quantifying the uncertainty in sea level due to surface forcing.
- Scenario Model Intercomparison Project (ScenarioMIP). We have an ongoing discussion with members of the ScenarioMIP steering committee to support an extension of the SSP5-8.5 beyond the planned 2015-2100 timeframe (ideally up to year 2300).
- Observations for Model Intercomparison (Obs4MIP). We would use the observations available on the current database to test how the inclusion of dynamic ice sheets affects the simulations. We would also suggest additional datasets that are pertinent to ice sheets and surface mass balance.
- Reanalysis for Model Intercomparison (Ana4MIP). We would use reanalysis in the assessment of the surface mass balance from AOGCM, and potentially as contemporary forcing for the ice sheets.
- Paleoclimate Modelling Intercomparison Project (PMIP). The ISMIP6 standalone ice sheet experiment lastInterglacialforcedism is a collaboration with PMIP, planned to use and augment PMIP experiments (time slices and transient, if possible). We plan to continue exchange with PMIP on the design of coupled climate-ice sheet simulations for different time periods.

These synergies with other MIPs illustrate potential collaborations within the climate community. It is hoped that other MIPs would be interested in using our simulations to investigate how changes in the ice sheets affect the component of the climate that is their expertise.

- If possible, a prioritization of the suggested experiments, including any rationale**

For the coupled AOGCM-ISM, we suggest that the pre-Industrial control and 1% yr CO2 to quadrupling CO2 experiments are performed first, followed by the SSP5-8.5 of ScenarioMIP. Our Tiers 1 experiments are thus: *piControlwithism* and *1pctCO2withism,* which allow for an easier evaluation of ice-climate feedback and have already been performed by many modeling groups. Our Tiers 2 experiment, *ssp5-8.5withism,* is however more relevant to our goal of sea-level rise projections that are in sync with the CMIP6 future climate, and SSP5-8.5 will be the focus of our sea level projection with standalone ice sheet models.

- All model output archived by CMIP6-Endorsed MIPs is expected to be made available under the same terms as CMIP output. Most modeling groups currently release their CMIP data for unrestricted use. If you object to open access to the output from your experiments, please explain the rationale.**

No objections.

- List of output and process diagnostics for the CMIP DECK/CMIP6 data request**
    - whether the variable should be collected for all CMIP6 experiments, or only some specified subset and whether the output is needed from the entire length of each experiment or some shorter period or periods;

For the standard climate simulations (with no dynamic ice sheets), the variables should be collected at a minimum for the CMIP6 AMIP, Historical Simulation, Pre-Industrial Control, 1% yr CO2 to quadrupling CO2, ScenarioMIP SSP5-8.5 and the PMIP Last Interglacial, for the entire length of the experiment.

For the simulations with climate models coupled to dynamic ice sheets, the variables would only be collected for the *piControlwithism*, *1pctCo2withism*, and *ssp5-8.5withism* experiments (see Table 1).

- whether the output might only be relevant if certain components or diagnostic tools are used interactively (e.g. interactive carbon cycle or atmospheric chemistry, or only if the COSP simulator has been installed);

The variables of priority 1 are always relevant and are variables that were variables of priority 1 in CMIP5. Some variables of lower priority are required for models that have either snow models over the ice sheets or interactive ice sheet models.

- whether this variable is of interest to downstream users (such as impacts researchers, WG2 users) or whether its principal purpose is for understanding and analysis of the climate system itself. Be as specific as possible in identifying why the variable is needed.

The principle purpose of the variable request is to understand and analyze the climate system, or force the standalone ice sheet models. Some variables (such as surface runoff or iceberg discharge) will affect sea level and are thus of interest to downstream users.

- whether the variables can be regridded to a common grid, or whether there is essential information that would be compromised by doing this;

The use of native grid may be preferable for variables that would be originating from the ice sheet grid. However, as indicated in the spreadsheet submitted to the WIP and below, the variables may be regridded to a common polar stereographic grid.

- the relative importance of the various variables requested (indicated by a tiered listing) is required if the data request is large.

The current CMIP5 CMOR tables Amon (Monthly Mean Atmospheric Fields), Omon (Monthly Mean Ocean Fields), and LImon (Monthly Mean Land Cryosphere Fields) already contains many of the output required to diagnose and intercompare the climate over glaciated land/ice sheets and to derive forcing for the ice sheets. However a few additional variables are needed to properly derive the forcings for ice sheets and to record outputs from the evolving ice sheets in the coupled AOGCM-ISMs experiments (such as ice elevation change). Table 2 and 3 list our assessment of the Amon, LImon, and Omon variables that we plan to use in ISMIP6, or that are missing. We have completed the CMIP6 data request forms by the deadlines set by CMIP and WIP. The submitted spreadsheet contains a tiered listing indicating the importance of the variable request.

Variable that are saved on the Amon and/or LImon Table which will be used by ISMIP6	Units	Existing CMOR variable name and CMIP5 location, and comments.	Tier
Temperature			
Near surface Air Temperature (2m)	K	tas in Amon	1
Surface Temperature	K	ts in Amon	1
Snow Internal Temperature	K	tsn in LImon	2
Temperature at the interface between ice sheet and snow/firn (NEW)	K	Similar to tsnint in OImon but over glaciated land. This would be the temperature used to force ice sheet models	2
SMB and its Components (at Upper Surface) SMB = snowfall flux – evaporation (sublimation) – Surface Melt + Refreezing where precipitation = snowfall flux + rainfall surface melt = Snow Melt + Ice Melt Surface runoff = available liquid water – refreezing = rainfall + surface melt - refreezing			
Surface Mass Balance flux (NEW, CF name exist)	kg m^{-2} s^{-1}		2
Precipitation	kg m^{-2} s^{-1}	pr in Amon	1
Snowfall Flux	kg m^{-2} s^{-1}	prsn in Amon	1
Rainfall Flux	kg m^{-2} s^{-1}	pr in Omon	2
Surface Snow and Ice Sublimation Flux	kg m^{-2} s^{-1}	sbl in LImon or sbl on Amon	2

Surface Snow and Ice Melt Flux (NEW, CF name exist)	kg m^{-2} s^{-1}		2
Snow Melt Flux	kg m^{-2} s^{-1}	snm in LImon or snm in Amon	3
Surface Ice Melt Flux (NEW)	kg m^{-2} s^{-1}	similar to tmelt in OImon	3
Surface Snow and ice refreezing flux (NEW, CF name exist)	kg m^{-2} s^{-1}		3
Surface Runoff (NEW, CF name exist)	kg m^{-2} s^{-1}	Similar to mrros in Lmon.	2
Total Runoff (NEW, CF name exist)	kg m^{-2} s^{-1}	Similar to mrro in Lmon.	2
Area fractions and land ice altitude, which may change in time when dynamic ice sheets are coupled			
Snow area fraction	%	snc in LImon	1
Land Ice area fraction	%	sftgif in fx	1
Grounded Ice area fraction (NEW)	%	Fraction of grid cell covered by grounded glaciated land (ice sheets or glacier, but NO ice shelves).	1
Land Ice Altitude	m	orog in fx. The altitude or surface elevation of the ice sheet, ice shelf or glacier.	1
Energy Fluxes			
Net latent heat flux over land ice	W m^{-2}	hfls in Amon	1
Sensible Heat flux over land ice	W m^{-2}	hfss in Amon	1
Downwelling Shortwave	W m^{-2}	rsds in Amon	1
Upward Shortwave over land ice	W m^{-2}	rsus in Amon	1
Downwelling Longwave	W m^{-2}	rlds in Amon	1
Upward Longwave over land ice	W m^{-2}	rlus in Amon	1
Albedo over land ice	1	similar to ialb in OImon	2

Table 2: Data in the LImon Table (Monthly Mean Land Cryosphere Fields) and/or Amon Table (Monthly Mean Atmospheric Fields) needed to capture the glaciated/ice sheet surface realm. Most of these variables already exist in

the CMIP5 tables and following the CMIP5 protocol, these fields are saved on the atmosphere grid and contain monthly output.

Variable that are saved on the Omon Table which will be used by ISMIP6	Units	Existing CMOR variable name and CMIP5 location, and comments/questions that need feedback.	Tier
Global Surface Height Above Geoid	m	zos in Omon	1
Global Average Thermosteric Sea Level Change	m	zostoga in Omon	1
Sea Water Potential Temperature	K	thetao in Omon	1
Sea Surface Temperature	K	tos in Omon	2
Sea Water Salinity	psu	so in Omon	1
Water flux into Sea Water from iceberg	kg m^{-2} s^{-1}	ficeberg in Omon computed as the iceberg melt water flux into the ocean divided by the area of the ocean portion of the grid cell.	2
Water flux into Sea Water from Ice Sheets (NEW)	kg m^{-2} s^{-1}	computed as the water flux into the ocean due to ice sheets (water runoff from the surface and base of the ice sheet, or melt water flux from base of ice shelf or vertical ice front)	3

Table 3: Data on the Omon Tables (Monthly Mean Ocean Fields) needed to capture the glaciated/ice sheet surface realm or for intercomparison of the model simulations. Most of these variables already exist in the CMIP5 Omon tables and following the CMIP5 protocol, these fields are saved on the ocean grid and contain monthly output.

For diagnosis and intercomparison of the dynamical ice sheet models within AOGCM (the coupled AOGCM-ISM), or for intercomparison of the standalone ice sheet models experiments, the variables in Table 4 would be saved on the dynamical ice sheet native grid or preferably on a regular (for example 1x1km or 5x5km) polar stereographic grid designed for the ice sheets (such as done in the SeaRISE effort). These variables would be recorded for the entire length of the experiments involving ice sheets.

Variables to be saved on ice sheet grid	Units	Comment	Tier
Ice Sheet Altitude (CF name exist)	m	The altitude or surface elevation of the ice sheet	1
Ice Sheet Thickness (CF name exist)	m	The mean thickness of ice sheet	1
Bedrock Altitude (CF name exist)	m	The bedrock topography.	1

Bedrock Geothermal Heat Flux	W m^{-2}	Upward Geothermal heat flux beneath the ice sheet	3
Land ice calving flux	kg m^{-2} s^{-1}		3
Land ice vertical front mass balance flux	kg m^{-2} s^{-1}	includes calving flux and melt on vertical ice front	2
Surface Mass Balance and its components	kg m^{-2} s^{-1}	see Table 2	1
Land ice basal melt flux	kg m^{-2} s^{-1}		1
Individual component of surface ice velocity (CF name exist)	m/yr		1
Individual component of basal ice velocity (CF name exist)	m/yr		1
U & V Vertical means ice velocity (CF name exist)	m/yr	The vertical mean land ice velocity is the average from the bedrock to the surface of the ice	3
Land ice basal stresses	Pa	Individual components of basal shear, normal, and longitudinal stresses	3
Surface Temperature of Ice Sheet	K		1
Basal Temperature of Ice Sheet	K		1
Area fraction	%	ice sheet and Grounded ice sheet	2
Scalar outputs / Integrated measures			
Ice Mass	Gt		2
Ice Mass not displacing sea water	Gt		2
Area covered by grounded ice	m^{2}		3
Area covered by floating ice	m^{2}		3

Table 4: Data to be saved on the ice sheet grid (sometimes monthly or yearly) to capture the dynamical ice sheet model realm. It would therefore be new CMOR tables.

- Any proposed contributions and recommendations for**
  - model diagnostics and performance metrics for model evaluation;
  - observations/reanalysis data products that could be used to evaluate the proposed experiments. Indicate whether these are available in the obs4MIPs/ana4MIPs database or if there are plans to include them;

- tools, code or scripts for model benchmarking and evaluation in open source languages (e.g., python, NCL, R).

Model evaluation over the ice sheets will include in situ, airborne, satellite data and reanalysis. Most of these cryospheric data products are not currently available in the obs4MIPs database and we will work closely with obs4MIPs to rectify this. The process is complicated by the need to evaluate both climate forcing over and around the ice sheets (i.e., AOGCM model output) and ice-sheet model response. We plan to have a workshop with data providers and modelers in 2015 to finalize the evaluation plan. The recent IMBIE project (Shepherd et al 2013) provides an excellent example of the ice-sheet observational community work together to provide a reconciled product suitable for testing ice-sheet models.

- Any proposed changes from CMIP5 in NetCDF metadata (controlled vocabularies), file names, and data archive (ESGF) search terms.**

- Explanation of any proposed changes (relative to CMIP5) that will be required in CF, CMOR, and/or ESGF.**

Some new standard CF names will be needed for ice sheet quantities, and there may be a need for ice sheet-grids to be handled in order to record the fields in Table 3, perhaps by CMOR.

Application for CMIP6-Endorsed MIPs: Land Surface, Snow and Soil Moisture (LS3MIP)

Date: 5 June 2015

Land Surface, Snow and Soil moisture MIP (LS3MIP)

- Co-chairs of MIP: Bart van den Hurk (hurkvd@knmi.nl), Gerhard Krinner (krinner@ujf-grenoble.fr), Sonia Seneviratne (sonia.seneviratne@ethz.ch), Chris Derksen (Chris.Derksen@ec.gc.ca), Taikan Oki (taikan@iis.u-tokyo.ac.jp) and Hyungjun Kim (hjkim@rainbow.iis.u-tokyo.ac.jp)
- Members of the Scientific Steering Committee: Martin Best, Paul Dirmeyer, Herve Douville, Richard Essery, Stefan Hagemann, Alex Hall, Randy Koster, Dave Lawrence, Twan van Noije, Helmut Rott, Andrew Slater, Matthew Sturm, Andrea Alessandri
- Endorsement: CliC and GEWEX
- Link to website: http://www.climate-cryosphere.org/activities/targeted/ls3mip. For sub-projects: http://hydro.iis.u-tokyo.ac.jp/GSWP3, http://www.iac.ethz.ch/GLACE-CMIP, and http://www.climate-cryosphere.org/activities/targeted/esm-snowmip

Goal of the MIP and brief overview

The goal of the LS3MIP experiment is to provide a comprehensive assessment of land surface, snow, and soil moisture-climate feedbacks, and to diagnose systematic biases in the land modules of current ESMs using constrained land-module only experiments. The solid and liquid water stored at the land surface has a large influence on the regional climate, its variability and its predictability, including effects on the energy and carbon cycles. Notably, snow and soil moisture affect surface radiation and flux partitioning properties, moisture storage and land surface memory. They both strongly affect the atmospheric conditions, in particular air temperature, but also large-scale circulation patterns and precipitation. However, models show divergent responses and representations of these feedbacks as well as systematic biases in the underlying processes. LS3MIP will provide the means to quantify the associated uncertainties and to better constrain climate change projections, of particular interest for highly vulnerable regions (densely populated regions, polar regions, agricultural areas, land ecosystems).

A short description of the role of snow and soil moisture in the climate system and of the rationale for the proposed experiments is provided hereafter.

Snow processes and snow-climate feedbacks

Snow cover is an essential component of the Earth System that interacts with the atmosphere and the surfaces it covers (land, ice, sea ice). It is also an important source of (positive) feedbacks within the climate system. A WCRP/CliC Initiative was proposed in 2013 for an ESM-SnowMIP intercomparison programme as a contribution to the WCRP Grand Challenge Cryosphere in a Changing Climate. This initiative builds on the evaluation of the current state of snow cover representation in climate models, which is being broadly addressed by observational and modeling groups across the snow community. It is a core element of the LS3MIP experiment.

It has been shown that CMIP5 models underestimate the observed spring snow cover trend in the Arctic (Derksen and Brown, 2012) and in the Northern Hemisphere (Brutel-Vuilmet et al., 2013). Snow-related climate feedbacks in the climate system arise primarily because of the well-known albedo feedback (e.g. Qu and Hall, 2007) that is also one of the main mechanisms leading to Arctic Amplification (e.g. Holland and Bitz, 2003). Snow-related biases in climate models may arise through this feedback, but also through the energy sink induced by snow melting in spring and through the strong thermal insulation effect of snow on the underlying soil. Koven et al. (2012) related strong biases in the simulated Northern Hemisphere permafrost extent in CMIP5 models to the representation of snow in these models.

Because of strong snow/atmosphere feedbacks, it is difficult to distinguish and quantify the various potential causes for disagreement in observed versus model snow trends. These causes include: the underestimation of the recent spring warming trend in CMIP5 models (e.g., Brutel-Vuilmet et al., 2013), weaknesses in their representation of snow processes, especially regarding the snow/albedo feedback (Qu and Hall, 2014), a positive pre-melt snow water equivalent (SWE) bias in CMIP5 models across the mid latitudes and the Arctic (Brown and Mote, 2009), increased deposition of light-absorbing impurities on snow which is not accounted for in most models (e.g., Dumont et al., 2014), or a combination of these with other unknown processes.

A better understanding of the links between snow cover and climate is critical to interpret the observed changes in recent years including links to variability in the atmospheric and ocean circulation, and the misrepresentation of polar amplification by climate models in the Arctic. It is a prerequisite for increasing the confidence in the projections of snow cover and its role in the subarctic (boreal) and Arctic climate. This understanding is also necessary for the long-term improvement of the representation of snow in climate models, which will also impact seasonal to interannual prediction of temperature, runoff and soil moisture.

The SnowMIP1 (Etchevers et al., 2002) and SnowMIP2 projects (Essery et al., 2009) evaluated the capacity of snow models of different complexity to simulate the snowpack evolution from local meteorological forcings. These projects were based on the evaluation of stand-alone simulations of snow models over a limited number of instrumented sites (see also http://www.wcrp-climate.org/index.php/modelling-wgcm-mip-catalogue/57-unifying-themes/modelling-wgcm/catalogue-of-model-intercomparison-projects/276-modelling-wgcm-catalogue-snowmip). However these pioneering projects did not explore snow-climate interactions, and were limited to the site scale.

LS3MIP will consider both stand-alone snow simulations at the global scale and snow outputs from climate simulations. Dedicated experiments will be designed for evaluating and understanding snow feedbacks within current climate models and assessing snow-related uncertainties in future projections. These will be a key action of the WCRP Grand Challenge Cryosphere in a Changing Climate coordinated by CliC/WCRP.

Soil moisture processes and soil moisture-climate feedbacks

Soil moisture modulates the energy and water balance at the land surface to a large extent (Koster et al., 2004; Seneviratne et al., 2010; van den Hurk et al., 2011). It interacts with vegetation, melting snow, ground water, boundary layer processes, atmospheric moisture, and is a key element for available fresh water resources, heat wave and drought propagation and soil erosion.

The modulating role of soil moisture is eminent at many relevant time scales: diurnal cycles of land surface fluxes, (sub-)seasonal predictability of droughts, floods, and hot extremes, annual cycles governing the water buffer in dry seasons, and shifts in the climatology in response to changing patterns of precipitation and evaporation (e.g. Betts 2004, Ek and Holtslag 2004, Santanello et al. 2009, Koster et al. 2010a,b, Douville et al. 2012, Mueller and Seneviratne 2012, Quesada et al. 2012, Dirmeyer et al. 2013, Miralles et al. 2014, Greve et al. 2014).

An important notion is the difficulty in generating reliable observations of soil moisture and land surface fluxes that can be used as boundary conditions for modeling and predictability studies. Satellite observations, in situ observations, offline model experiments and indirect estimates all have a potential to generate relevant information, but are largely inconsistent, covering different subdomains of the states, and suffer from methodological flaws. As a consequence, the pioneering work on deriving soil moisture related predictability and regional/global climate responses has been carried out using (ensembles of) modeling experiments. The following studies are particularly relevant in this respect.

The Global Soil Wetness Project (particularly phase 2, GSWP2; Dirmeyer et al., 2006) yielded a 10-year “climatology” (1986-1995) of all land surface states including soil moisture and surface fluxes based on an ensemble of offline land surface models, driven by pseudo-observed climatological forcings. Various follow-up projects to extend the period and applications of this product have taken place or are being planned. *For CMIP6 it is of utmost relevance to document the characteristics of the land surface component of the coupled models under observation-based constrained conditions, and document its main systematic biases.* A third edition of GSWP is being prepared (GSWP3; see http://hydro.iis.u-tokyo.ac.jp/GSWP3). Participation by a large subset of the land surface models used in the CMIP6 ensemble allows the generation of a well constrained CMIP6 climatology of land surface characteristics, and provides input to model evaluation and predictability studies. *Therefore, incorporating GSWP3 in the CMIP6 program can be seen as the LMIP of CMIP6, an analogy to AMIP or OMIP. The LMIP simulations will build upon the GSWP3 experiments and were identified, together with OMIP, as possible future DECK experiments at the recent WGCM-18 meeting. In CMIP6, these “proto-DECK” experiments are recommended for Tier1. They will allow an assessment of the representation of soil moisture and snow processes, as well as of other land surface processes (e.g. vegetation) and associated fluxes of water and energy in the CMIP6 land surface models.*

The Global Land Atmosphere Coupling Experiment (GLACE: Phases 1 and 2 on seasonal forecasting (Koster et al. 2004; 2010) and GLACE-CMIP5 on climate change projections (Seneviratne et al., 2013)) provided first assessments of the role of soil moisture for the climate system. The GLACE-1 analysis (Koster et al., 2004) pioneered the identification of regions where soil moisture has a significant impact on the local hydroclimate, based on an ensemble of idealized model simulations. At the seasonal time scale transitional wet-dry climate regions, mostly coinciding with monsoon regions, display an identifiable soil moisture-precipitation coupling. Expanding the GLACE framework at the climate time scale and for regional climate simulations in Europe, Seneviratne et al. (2006) illustrate changes in patterns of coupling strength between present and future climate conditions, showing a shift of the area of strong land-atmosphere interactions from the Mediterranean region to Central and Eastern Europe. More recently, the GLACE-CMIP5 multi-model experiment (Seneviratne et al., 2013) uses this expanded GLACE framework to investigate the role of soil moisture in modifying the regional temperature and precipitation response to a future climate forcing. *The experimental design of the GLACE-CMIP5 study, carried out with a limited CMIP5 ensemble with prescribed SSTs (AGCMs) and vegetation, is used as blueprint for the second set of proposed LS3MIP experiments, described in detail below. The new LS3MIP experiments will allow a full quantification of soil moisture-climate feedbacks in the CMIP6 models and provide reference diagnostics for the evaluation of the CMIP6 ESMs, which will be of key relevance for the application of constraints to reduce uncertainties in projections*.

In addition, LS3MIP will include an assessment of changes in land-based predictability in the CMIP6 models. These experiments build upon the GLACE2 predictability experiment (Koster et al., 2010a), in which the actual temperature and precipitation skill improvement of using observation constrained estimated soil moisture initializations is shown to be much lower than suggested by the coupling strength diagnostics. Limited quality of the initial states, limited predictability and poor representation of essential processes determining the propagation of information through the hydrological cycle in the models all play a role. *An update of the land surface related predictability in state of the art climate models will reveal essential information about the models' ability to represent the terrestrial hydrological processes, the inherent limitations to predictability, and possible shifts in patterns of predictability in response to climate change (Dirmeyer et al., 2013). This will be evaluated in the third branch of LS3MIP experiments.*

Both the LMIP (GSWP3) and the soil moisture-based LS3MIP experiments are key action items of the WCRP grand challenges on water availability and climate extremes, which are coordinated by the GEWEX project.

Objectives of LS3MIP

The Land Surface Snow and Soil moisture MIP (LS3MIP) will embrace a small number of multi-model experiments, encompassing simulations driven in offline mode (land-surface only), coupled to the atmosphere (driven by prescribed sea surface temperatures, SSTs), and embedded in fully coupled AOGCMs. The experiments are subdivided in two components, the first one addressing land systematic biases ("LMIP", building upon the GSWP3 experiment) and the second one addressing land feedbacks in an integrated framework ("LFMIP", building upon the ESMsnowMIP and GLACE-CMIP blueprints). The LS3MIP experiments address together the following objectives:

- an evaluation of the current state of land processes including surface fluxes, snow cover and soil moisture representation in CMIP6 DECK runs, revealing main *systematic biases and their dependencies* (LMIP-protoDECK)
- a *multi-model estimation* of the long-term terrestrial energy/water/carbon cycles, using the surface modules of CMIP6 models under observation constrained historical (land reanalysis) and projected future (impact assessment) conditions considering land use/land cover changes. (LMIP)
- an assessment of the role of snow and soil moisture feedbacks in the regional response to altered climate forcings, focusing on controls of climate extremes, water availability and high-latitude climate in historical and future scenario runs (addressing Arctic amplification and drought/heatwave characteristics) (LFMIP)
- an assessment of the contribution of land surface processes to the current and future *predictability* of regional temperature/precipitation patterns. (LFMIP)

These objectives respond to each of the three CMIP6 overarching questions: what are regional feedbacks and responses to climate change, what are the systematic biases in the current climate models, and what are the perspectives concerning the generation of predictions and scenarios.

Embedding of LS3MIP within WCRP and CMIP6

As illustrated in Figure 1, LS3MIP is addressing core research questions of the WCRP and is relevant for a large fraction of the WCRP activities. It is initiated by two out of four WCRP core projects (CliC and GEWEX) and directly related to three WCRP Grand Challenges (Cryosphere in a Changing Climate, Changes in Water Availability, and Climate Extremes). The LMIP experiment will provide best estimates of historical changes in snow and soil moisture on global scale, thus allowing the evaluation of changes in freshwater, agricultural drought, and streamflow extremes over continents. The LFMIP experiment is of high relevance for the assessment of key feedbacks and systematic biases of land surfaces processes in coupled mode, and is also addressing two of the main feedback loops over land: The snow-albedo-temperature feedback, involved in Arctic Amplification, and the soil moisture-temperature feedback leading to major changes in temperature extremes. Hence LS3MIP is directly addressing some of the main questions underlying the *Cryosphere in a Changing Climate* and *Changes in Water Availability* Grand Challenges, and will also provide essential insights on temperature and hydrological extremes for the *Climate Extremes* Grand Challenge. In addition, LS3MIP will also allow the exchange of data and knowledge across communities, as snow and soil moisture dynamics are often interrelated (e.g. (Hall et al. 2008) and contribute together to hydrological variability (e.g. Koster et al. 2010b). LS3MIP will thus constitute a core element within WCRP, binding together several communities that, in fact, address a common physical object: water on land, in its liquid or solid form.

LS3MIP within WCRP Core Projects and Grand Challenges

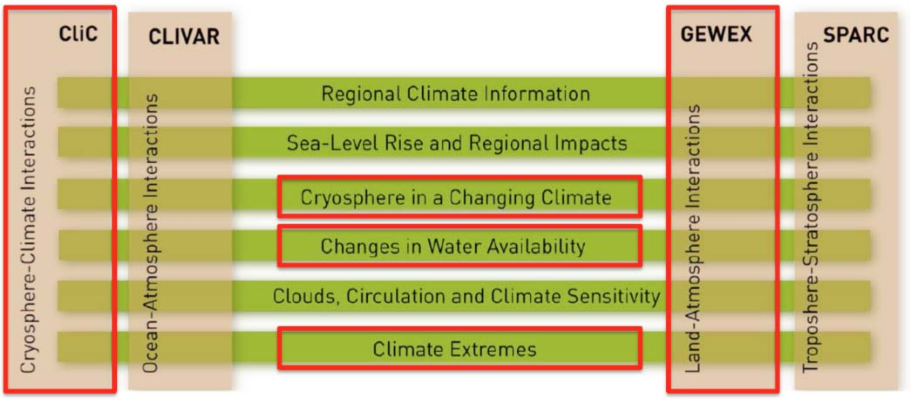


Figure 1: Relevance of LS3MIP for WCRP Core Projects and Grand Challenges

In addition, LS3MIP will provide relevant insights for other research communities within WCRP, such as estimates of freshwater inputs to the oceans (which are relevant for sea-level changes and regional impacts), the assessment of feedbacks shown to strongly modulate regional climate variability and thus relevant for regional climate information, as well as the investigation of land climate feedbacks on large-scale circulation patterns and cloud occurrence. This will thus also imply potential contributions to the other WCRP grand challenges and core projects.

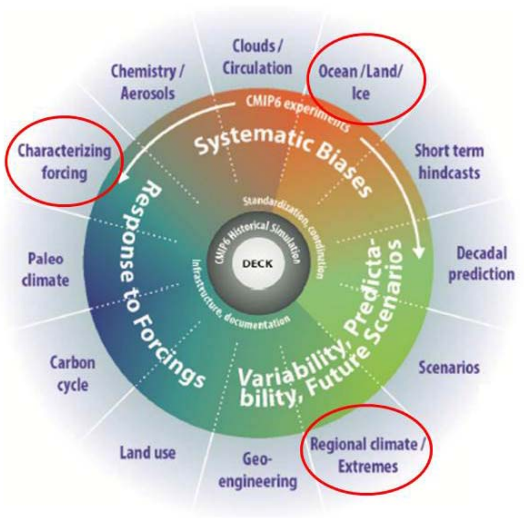


Figure 2: Embedding of LS3MIP within CMIP6

Figure 2 illustrates the embedding of LS3MIP within CMIP6. LS3MIP clearly fills a major gap, by allowing the consideration of land systematic biases and land feedbacks within the CMIP6 framework. In this context, LS3MIP can be seen as part of a larger "LandMIP" series of experiments fully addressing biases, uncertainties, feedbacks and forcings from the land surface (Figure 3), which are complementary to similar experiments for ocean or atmospheric processes. In particular, we note that while LS3MIP focuses on *systematic biases* in land surface processes (LMIP) and on *feedbacks* from the land surface processes on the climate system (LFMIP), the complementary LUMIP experiment (separate proposal) addresses the role of land surface *forcing* on the climate system. The role of vegetation and carbon stores in the climate system is a point of convergence between LUMIP and LS3MIP. In particular, the LMIP/GSWP3 experiment will serve as land-only reference experiments for both the LS3MIP and LUMIP experiments. In addition, there will also be links to the C4MIP experiment with respect to impacts of snow and soil moisture processes (in particular droughts) on terrestrial carbon exchanges and resulting feedbacks to the climate system.

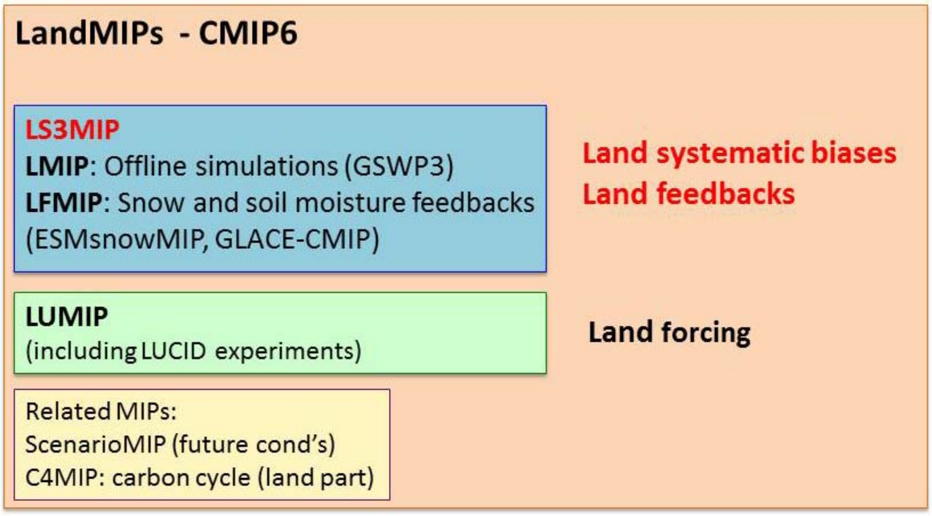


Figure 3: Overview of the embedding of LS3MIP in land-related MIPs. LS3MIP will allow the quantification of land systematic biases and feedbacks induced by snow and soil moisture processes, while LUMIP addresses land forcing on climate.

Overview of the proposed experiments

A number of complementary experiments are proposed as part of LS3MIP (see Figure 4 and Table 1):

(1) Offline land model experiment ("Land offline MIP", LMIP):

In the context of GSWP3 meteorological forcings are made available to drive land modules from climate models in an offline mode. Offline land simulations of land surface states and fluxes allow for the evaluation of trends and variability of snow, soil moisture and land surface fluxes, carbon stores and vegetation states, and climate change impacts. Ancillary data (e.g., land use/cover changes, surface parameters, CO_2 concentration) and documented protocols to spin-up and execute the experiments are currently being compiled. The LMIP meteorological forcing is developed and provided by the GEWEX-GLASS panel, which has a long-standing experience in the domain of large-scale land surface model evaluation. Nevertheless, some known and unknown issues inevitably remain. In areas affected by known issues with large-scale forcing data or in areas where conspicuous systematic large-scale biases common to many land surface models appear, these can be identified by a concomitant evaluation of large-scale and plot-scale meteorological data and simulations, the latter being forced by directly observed meteorological data. Our analysis will take advantage of these possibilities. Additional information on the provided meteorological forcing for the offline experiments is given in the Annex (page 14).

(1a) Land reanalysis: LMIP-Hist

One set of reference forcing data and a standard bias correction strategy will be provided to drive each land surface model for the historical (1850-2014) simulations. Although this historical experiment is not yet a formal member of the DECK simulations, the WGCM recognized the importance of these offline experiments for the process of model development and benchmarking. The subset (1979-2014) of this historical run, largely analogous to AMIP, constitute Tier 1 of LMIP and is proposed to become part of the DECK in future CMIP exercises. A future implementation into the DECK is foreseen and the LMIP simulations were therefore identified as proto-DECK experiments.

Complementary, additional site level validations using 1d time series of observational forcing variables from selected reference sites will be proposed in parallel to the CMIP6 framework.

(1b) Climate change impact assessment: LMIP-Fut

The future simulations (2015-2100) constitute Tier 2 of LMIP. In these simulations, the atmospheric output of at least 2 scenarios based on the ScenarioMIP (tentatively, SSP5-8.5(ref) and SSP1-2.6) will be exploited as forcing data with a statistical bias correction method[1] for constant and time varying variables. It focuses on climate change impact assessment (e.g., on water availability and climate extreme) and estimation of the sensitivity of land modules of CMIP6 GCMs to the projected future.

(2) Prescribed land surface states to assess the impact of snow and soil moisture feedbacks ("Land Feedback MIP", LFMIP):

---

[1] We envisage following the bias correction method described by Koven et al., *Proc. Nat. Acad. Sci.*, **112,** 3752-3757, doi:10.1073/pnas.1415123112, 2015.

Here the GLACE-CMIP5 protocol is followed, where apart from the CMIP6 DECK experiments a set of forced experiments is carried out, where land surface states are prescribed from an a priori defined database. In contrast to the earlier experiments coupled AOGCM simulations are anticipated, where the Historical (1980-2014) and future (2015-2100) simulations will be used as reference. For the future a single scenario from the ScenarioMIP will be selected at a later stage. The land surface states that are prescribed may vary across the participating models depending on the model structure, but at least include the water reservoirs (soil moisture, snow mass), but may be extended to other prognostic quantities related to vegetation or temperature.

The earlier experience with GLACE-related experiments has revealed a number of technical and scientific issues:

Technical aspects: because in most GCMs the land surface module is an integral part of the code describing the atmosphere, prescribing land surface dynamics requires a non-conventional technical interface, reading and replacing variables throughout the entire simulations. Many participants to LS3MIP have participated earlier in GLACE-type experiments, but for some the code adjustments will require a technical effort. In this interface some non-trivial choices have to be made on the selection of variables read in externally, resolving possible mismatches with other static or dynamic variables. A standardization of this procedure is, given the wide variety of land surface formulation, not feasible; the decision on how to implement this interface is up to the research teams, who are in the best position to optimize this interface. The disparity of possible implementations is adding to the uncertainty range generated by the model ensemble, similar to the degree to which implementation of land use, flux corrections or downscaling adds to this uncertainty range. Understanding of the contribution of the protocol to prescribe land surface variables is helped by a careful documentation of the way the modelling groups have implemented this interface. The LS3MIP leadership will organize a detailed discussion of the forcing techniques among the participating modeling groups in order to favor as much as possible a homogenous implementation, and it will ensure careful and coordinated checks of the prescribed land-surface variables in preliminary and final LS3MIP runs.

Water and energy conservation: similar to AMIP, nudged, and data assimilation experiments, the prescribed land surface experiments do not conserve water and energy. A systematic generation or destruction of water or energy can even in emerge as a result of asymmetric land surface responses to dry and to wet conditions, if e.g. surface evaporation or runoff depend strongly non-linearly to soil moisture or snow states. Also, unrepresented processes (such as water extraction for irrigation or exchange with the groundwater) may lead to imbalances in the budget. This systematic alteration of the water and energy balance may interact with the projected climate change signal, where altered climatological soil conditions can contribute to the climate change induced temperature or precipitation signal. In the analyses of the experiments this asymmetry and lack of energy/water balance closure will be examined and put in context of the climatological energy and water balance and its climatic trends[2].

---

[2] In case this question spurs sufficient interest among the engaged modeling groups, additional experiments enforcing water conservation (at least during the snow-free season) can be envisaged by imposing that any water added to or withdrawn from the soil is taken from (added to) runoff or groundwater storage, and if this

Earlier GLACE-type experiments revealed that often the issue of water conversation is reduced when prescribed soil water conditions are taken as the median rather than the mean of a sample over which a climatological mean is calculated. Therefore, we will use median soil water and snow conditions in LS3MIP. We further note that many of our simulations are carried out in the atmosphere-only mode in which water and energy conservation is not respected anyway.

(2a) Core experiments: 2 experiments are considered to be "core":

- prescribed climatology derived from "present climate" conditions (e.g. 1980-2014), aiming at diagnosing the role of land-atmosphere feedback at the climate time scales;
- prescribed climatology using a transient 30-yr running mean, where a comparison to the standard CMIP6 runs allows diagnosing shifts in the regions of strong land-atmosphere coupling, and shifts in potential predictability related to land surface states.

Both simulations cover the historical period and extend to 2100, based on a forcing scenario to be identified at a later stage. In both cases (and in corresponding cases below), the coupled model's climatological median soil water and snow water equivalent for will be used instead of the climatological mean (see above). We note that the "prescribed climatology derived from present climate conditions" experiments can be seen as an idealized future irrigation experiment.

Output in high temporal resolution (daily, as well as sub-daily for some fields and time slices) is planned in order to address the role of land surface-climate feedbacks (including snow and soil moisture feedbacks) on climate extremes on land. These outputs may be generated for shorter time slices only.

A single member of each of these core simulations is considered to be part of the Tier 1 simulations, but multi-member experiments are encouraged (and included in a Tier 2 set of simulations).

(2b) As (2a) for AGCM simulations

The AOGCM simulations from (2a) are duplicated with a prescribed SST configuration (AGCM), and also these simulations are included in the Tier 2 set of LS3MIP experiments.

(2c) Separate effects of soil moisture and snow, and role of additional land parameters and variables

Additional experiments in which only snow, snow albedo or soil moisture is prescribed will be conducted to assess the respective feedbacks in isolation, and have control on possible interactions between snow cover and soil moisture content. At a later stage, also vegetation parameters and variables (e.g. leaf area index) could be considered. These experiments are all part of the Tier 2 batch of LS3MIP.

(2d): As (2a) for fixed land use conditions

In conjunction with the Land Use MIP (LUMIP) a repetition of experiment (2a) under unchanging land cover and land use conditions is planned. This experiment highlights the role of soil moisture in

---

is not possible this additional water is not provided. This would correspond to a maximum irrigation experiment that could serve as a sensitivity experiment to investigate this issue in more detail.

modulating the climate response to land cover and land use. It is a tier 2 set of experiments in LS3MIP.

Apart from the above experiments, particular sensitivity experiments are proposed to isolate the role of individual processes (such as prescribed albedo to address snow-related feedbacks, or vegetation parameters addressing carbon/water interactions). These all will be Tier 2 experiments, and be designed throughout the runtime of LS3MIP.

(3) Prescribed land surface states derived from pseudo-observations (LFMIP-predictability)

The use of experimental batch (1) (offline land models) to initialize the AOGCM experiments (batch 2) allows a set of predictability experiments in line with the GLACE2 set-up. Here historical runs from 1980 to 2014 are proposed in AOGCM mode, with a prescribed series of 'reconstructed' land surface states, either derived from the offline simulations or derived from various observational data sources (such as for SWE or snow albedo, using satellites, reanalysis and land surface model outputs). The predictability assessments include the evaluation of the contribution of snow cover melting and its related feedbacks to the underestimations of recent boreal polar warming by climate models.

Table 1 and Figure 4 summarize the experimental overview, where experiments focusing on specific processes and the LUMIP configuration (2c and 2d) are not included in this inventory.

Experiment Name	**Tier**	**Experiment Description / Design**	**Configuration**	**Start**	**End**	**# Years of Simulation**	**# Ens**	**# Total Years**	**Science Question and/or Gap Being Addressed with this**	**Possible Synergyes with other MIPs**	**Run Schedule**
LMIP-H	1	Land only simulations	LND	1850	2014	165	2	**330**	Land reanalysis	LUMIP, C4MIP, CMIP6 historical	Jan.-Jun., 2016
LMIP-F	2	Land only simulations	LND	2015	2100	86	4	**344**	Climate trend analysis	LUMIP, C4MIP, ScenarioMIP	Jan.-Jun., 2016
LFMIP-CAO1	1	Prescribed land conditions 1980-2014 climate	LND-ATM-OC	1980	2100	121	1	**121**	diagnose land-climate feedback including ocean response	ScenarioMIP	After DECK (2017?)
LFMIP-CAO4	2	Prescribed land conditions 1980-2014 climate	LND-ATM-OC	1980	2100	121	4	**484**	diagnose land-climate feedback including ocean response	ScenarioMIP	
LFMIP-CA5	2	Prescribed land conditions 1980-2014 climate; SSTs prescribed	LND-ATM	1980	2100	121	5	**605**	diagnose land-climate feedback over land	ScenarioMIP	
LFMIP-RAO1	1	Prescribed land conditions 30yr running mean	LND-ATM-OC	1980	2100	121	1	**121**	diagnose land-climate feedback including ocean response	ScenarioMIP	
LFMIP-RAO4	2	Prescribed land conditions 30yr running mean	LND-ATM-OC	1980	2100	121	4	**484**	diagnose land-climate feedback including ocean response	ScenarioMIP	
LFMIP-RA5	2	Prescribed land conditions 30yr running mean; SSTs prescribed	LND-ATM	1980	2100	121	5	**605**	diagnose land-climate feedback over land	ScenarioMIP	
LFMIP-HP10	2	Initialized pseudo-observations land	LND-ATM-OC	1980	2014	35	10	**350**	land-related seasonal predictability	CMIP6 historical	After DECK (2017?)

Table 1: Summary of LS3MIP experiments. Details on separate sensitivity studies and selected scenarios have not been included.

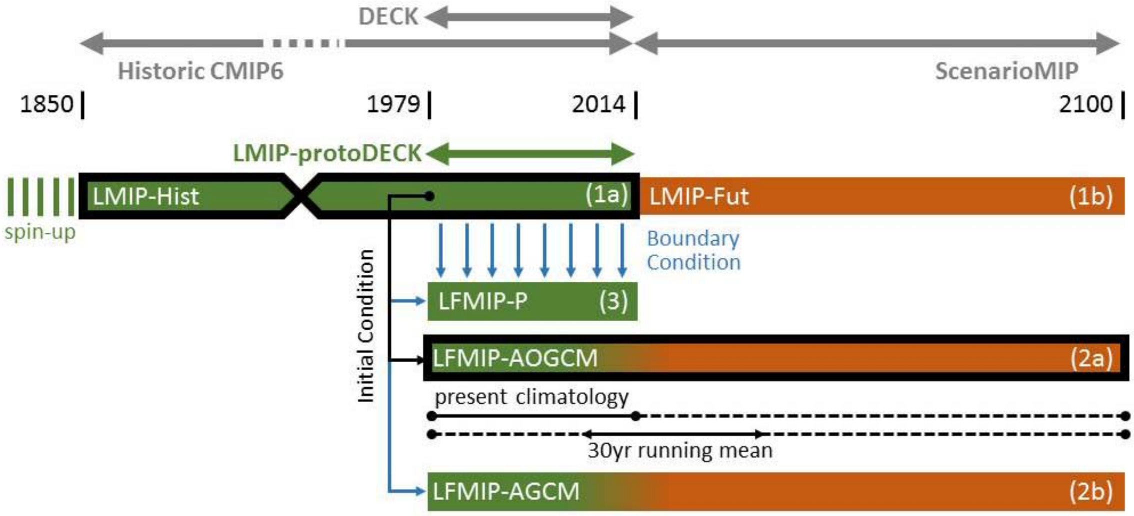


Figure 4: Schematic diagram for the experiment structure of LS3MIP (black-outline for the Tier 1 experiment).

Overview of the proposed evaluation/analysis of the CMIP DECK and CMIP6 experiments

LS3MIP brings together climate modelers, snow and soil moisture model specialists and experts in local and remotely sensed data of soil moisture and snow properties, mass and extent. This diversity is reflected in the composition of the steering group of LS3MIP and ensures that the experiment setups, model evaluations and analyses/interpretations of the results are pertinent.

Analyses for snow

Concerning the analysis of climate model runs, large-scale datasets of snow mass (SWE) and snow cover extent (SCE) are particularly relevant for the analysis of the historical simulations in the LS3MIP framework (i.e. the AMIP runs and the historical coupled run). These large-scale, high-quality snow data are available through close links to the Satellite Snow Product Intercomparison and Evaluation Experiment (SnowPEX, http://calvalportal.ceos.org/projects/snowpex), via the composition of the steering group. The quality of the representation of these fundamental snow-related variables in the historical simulations (coupled and AMIP) will be evaluated against these datasets. Output from the historical simulations are required to update analyses of the agreement between observations and historical simulations, and determine new projections of the variability and trends in terrestrial snow cover extent and mass (this was examined with CMIP3 and CMIP5 output in studies such as Brown and Mote (2009); Derksen and Brown (2012); Brutel-Vuilmet et al. 2013). These analyses, besides their genuine interest, can also provide clues to the interpretation of general model deficiencies in the representation of boreal and polar climates. The representation of albedo over snow-covered areas in DECK simulations will be analyzed. Multiple satellite-derived datasets are available for the evaluation of simulated albedo, including 16-day MODIS data (2001-present; http://modis-atmos.gsfc.nasa.gov/ALBEDO/) and the recently updated twice-daily APP-x product (1982-2011; http://stratus.ssec.wisc.edu/products/appx/appx.html). Specific attention will be paid to the role of the models' representation of snow cover fraction in forested and mountainous areas. The DECK simulations will be used to update analyses of observed and simulated snow-albedo feedback, an important diagnostic in determining climate sensitivity to snow cover (Qu and Hall, 2014; Fletcher et al. 2012).

The LS3MIP will be analyzed in concert with the control runs to quantify various climatic effects of snow, including very accurate estimates of snow albedo feedback. For example, the prescribed albedo experiments (simulation set 2c) do not allow the optical properties of vegetation to change in snow-covered areas as the climate warms. However, the prescribed SWE experiments do allow for this effect. The surface albedo change in the Prescribed SWE experiments can be compared to the overall albedo change in the control experiments to quantify the degree to which the surface albedo changes in snow-covered areas are due to vegetation changes, rather than snow changes. These estimates can be used to confirm that snow albedo feedback effects diagnosed from the Prescribed Albedo experiments are not misleading due to vegetation effects. Similarly, the Prescribed albedo experiments contain changes in soil moisture and hydrology due to melting snow. These can be compared to the control experiments to ascertain the degree to which snowmelt influences hydrology independently of its substantial influence on surface absorbed solar radiation. In this way, one can assess the degree to which the Prescribed SWE experiments produce snow effects unrelated to snow albedo feedback.

The geographical focus of the first stage of this project is on the continental snow cover of both hemispheres, both in ice-free areas (Northern Eurasia and North America) and on the large ice sheets (Greenland and Antarctica). In later stages of LS3MIP, the effect of snow on sea ice will be analyzed. Major scientific questions concerning snow on sea ice are related to strong recent trends of Arctic sea ice decline and the potential amplifying effect of earlier snow melt. These questions can be tackled by AGCM runs with a dynamic atmospheric nudging to eliminate biases related to misrepresentation of NH circulation trends (AO, NAO). Some of the modeling groups that have declared interest in participating in the snow-related part of LS3MIP (ESM-SnowMIP) are currently carrying out “proof of concept” simulations using prescribed snow mass (SWE) with an AMIP-type DECK control run; note that similar experiments have already been carried out (e.g., Lawrence and Slater, 2009; Alexander et al., 2011), demonstrating the feasibility and scientific interest of the proposed experiments.

Analyses for soil moisture

The analyses will focus on 1) systematic biases in offline land simulations (LMIP/GSWP3 simulations) and on 2) the role of soil moisture – climate feedbacks for past and projected changes in land climate conditions.

In the case of systematic land biases, the LMIP/GSWP3 simulations will be evaluated with observations available over the historical time period (e.g. for runoff, storage anomalies, vegetation activity) to assess their degree of realism and typical biases compared to measurements. Uncertainties of current land surface models in the representation of historical variations in land water availability/droughts (due to model parameterizations and/or atmospheric forcings, Sheffield et al. 2012, Trenberth et al. 2013, Greve et al. 2014) as well as systematic biases in water, energy and carbon exchanges between the land and the atmosphere (e.g. Mueller and Seneviratne 2014) will be assessed. These assessments will be used for the evaluation of the offline simulations for future land conditions as well as the coupled experiments.

In the case of soil moisture-climate feedbacks, the focus will be set on the following topics:

1. The quantification of the impact of soil moisture variability for climate variability (trends, decadal variability, interannual anomalies, extremes) on land and its interaction with large-scale drivers (large-scale modes of variability, ocean-climate interactions)
2. The attribution of model disagreement in land temperature, precipitation, runoff vegetation activity, carbon sink to the representation of soil moisture, related processes (plant transpiration and photosynthesis) and feedbacks to the atmosphere
3. The derivation of emergent constraints to reduce uncertainties in projections of mean climate and extremes (hot temperatures, droughts, floods) using observations characterizing the identified soil moisture-climate feedbacks
4. The regional assessment of the relationship between bias in modelled soil moisture/land surface representation and climate response

5. A robust estimate on the geographical patterns of “hot spots” of changes in soil moisture dynamics and their impact on occurrence of droughts, heat waves, irrigation limitations or river discharge anomalies.
6. The assessment of the role of soil moisture for subseasonal to seasonal predictability over land in both present and future climate.

Proposed timing

The proposed experiments are continuous model runs duplicating the Historical and ScenarioMIP simulations. AMIP mode runs are foreseen as a Tier 2 set of experiments. The experimental setup requires a reasonable amount of additional coding for reading and prescribing land surface characteristics, while many groups already participated in one of the earlier experiments. It is anticipated that the feedback simulations of LS3MIP are carried out after the first set of core CMIP6 experiments (i.e. DECK and historical runs). Stand-alone simulations with the ESMs’ land surface modules in uncoupled mode are currently planned in the context of GSWP3, and will initiate early 2016. The evaluation of land surface processes in CMIP6 Historical Simulation experiments will start as soon as historical runs are available.

A 6 month preliminary period will be dedicated to a wide consultation of the climate modeling community aiming at finalizing the detailed experiments design.

Group commitment

The following ESM groups and contact persons (represented in Scientific Steering Committee) will participate in LS3MIP (according to inventory of CMIP 25 March 2015):

- ACCESS (Andy Pitman)
- BCC
- CanESM
- CAS-ESM
- CESM (David Lawrence)
- CESS-THU
- CNRM (Hervé Douville)
- EC-Earth (Andrea Alessandri)
- FGOALS
- IPSL (Gerhard Krinner)
- MIROC6 and MIROC-ESM (Hyungjun Kim)
- MPI (Stefan Hagemann)
- MRI-ESM
- UKESM (Martin Best)

A list of potentially interested groups includes:

- BNU
- CMCC
- GFDL (Kirsten Findell, tbc)
- GISS
- NorESM

Annex: the offline forcing

The atmospheric boundary condition for LMIP is based on 20th Century Reanalysis (20CR) (Compo et al. 2011) assimilating only surface pressure and using observed monthly sea-surface temperature and sea-ice distribution. The ensemble uncertainty in the synoptic variability of 20CR has been evaluated, and this uncertainty varies with the time-changing observation network. High correlations for geopotential height (500 hPa) and air temperature (850 hPa) with an independent long record (1905-2006) of upper-air data are found. Although observation density varies over time and location, the long-term radiosonde based correlation is similar to the state-of-the-art three days forecast skill.

A spectral nudging dynamical downscaling technique effectively retains synoptic features in the higher spatial resolution. Additional bias corrections using observations introduce considerable improvements. As further efforts putting into the dynamical downscaling, vertical damping (Hong and Chang 2012) and single ensemble member correction (Yoshimura and Kanamitsu 2013) have been applied. Figure A1 shows the performance of the forcing data generation framework in each step relative to the original reanalysis. Phase and amplitude of daily variability of 20 globally distributed in-situ observation sites in FLUXNET are used as a reference. Although the intrinsic heterogeneity leads to significant differences comparing to in-situ observations for precipitation, many variables (e.g., long- and short-wave downward radiation and air temperature) show variability similar to the observations.

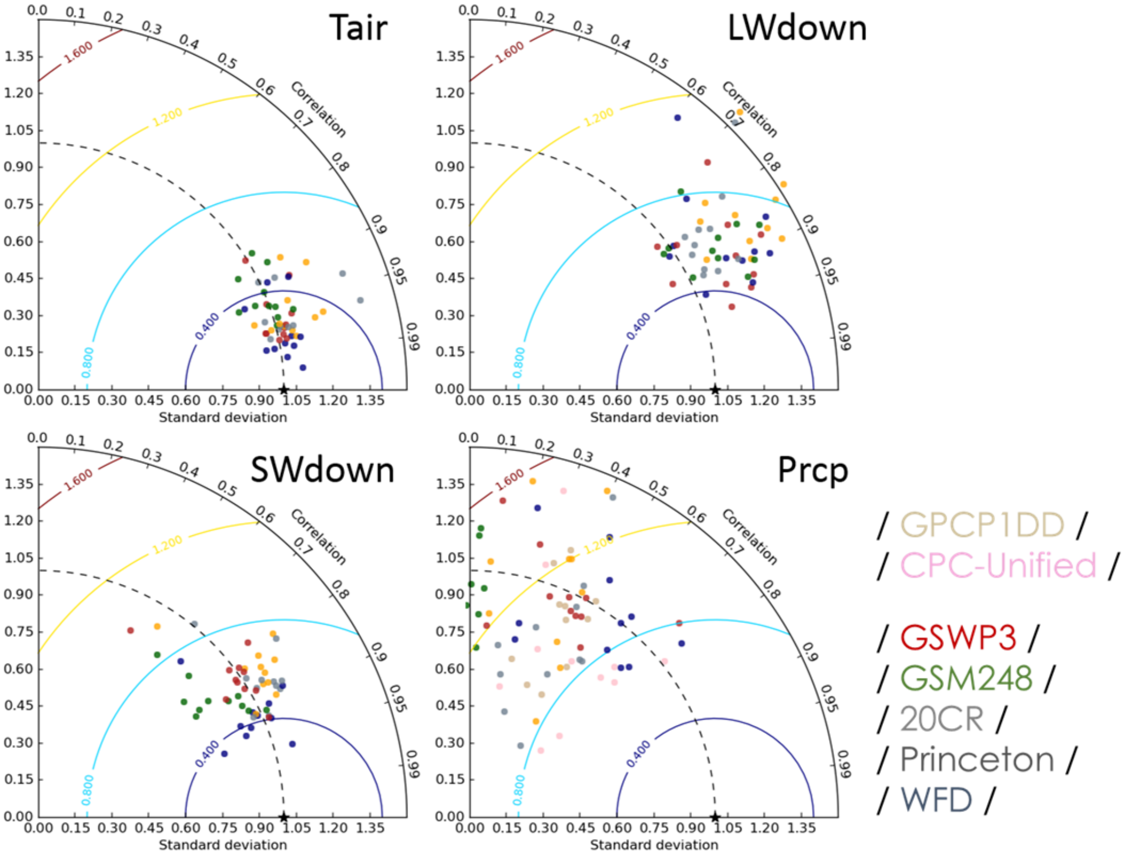


Figure A1: Taylor diagram for atmospheric boundary variables comparing to FLUXNET sites

References

Alexander, M.A., R. Thomas, C. Deser, and D. Lawrence, 2009: The atmospheric response to projected terrestrial snow changes in the late 21st century. *J. Clim.*, **23**, 6430-6437, doi: 10.1175/2010JCLI3899.1.

Betts, A.K., 2004. Understanding hydrometeorology using global models. B. Am. Meteorol. Soc. 85 (11), 1673–1688.

Brown, R. D., and P. W. Mote, 2009: The Response of Northern Hemisphere Snow Cover to a Changing Climate*. *J. Clim.*, **22**, 2124–2145, doi:10.1175/2008JCLI2665.1.

Brutel-Vuilmet, C., M. Ménégoz, and G. Krinner, 2013: An analysis of present and future seasonal Northern Hemisphere land snow cover simulated by CMIP5 coupled climate models. *The Cryosphere*, **7**, 67–80, doi:10.5194/tc-7-67-2013.

Compo, G. P., and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Q. J. R. Meteorol. Soc.*, **137**, 1–28, doi:10.1002/qj.776.

Derksen, C., and R. Brown, 2012: Spring snow cover extent reductions in the 2008–2012 period exceeding climate model projections. *Geophys. Res. Lett.*, **39**, L19504, doi:10.1029/2012GL053387.

Dirmeyer, P. A., X. Gao, M. Zhao, Z. Guo, T. Oki, and N. Hanasaki, 2006: GSWP-2: Multimodel Analysis and Implications for Our Perception of the Land Surface. *Bull. Am. Meteorol. Soc.*, **87**, 1381–1397, doi:10.1175/BAMS-87-10-1381.

——, S. Kumar, M. J. Fennessy, E. L. Altshuler, T. DelSole, Z. Guo, B. A. Cash, and D. Straus, 2013: Model Estimates of Land-Driven Predictability in a Changing Climate from CCSM4. *J. Clim.*, **26**, 8495–8512, doi:10.1175/JCLI-D-13-00029.1.

Dumont, M., and Coauthors, 2014: Contribution of light-absorbing impurities in snow to Greenland/'s darkening since 2009. *Nat. Geosci*, **7**, 509–512.

Fletcher, C. G., S. C. Hardiman, P. J. Kushner, and J. Cohen, 2009: The Dynamical Response to Snow Cover Perturbations in a Large Ensemble of Atmospheric GCM Integrations. *J. Clim.*, **22**, 1208–1222, doi:10.1175/2008JCLI2505.1.

——, H. Zhao, P. J. Kushner, and R. Fernandes, 2012: Using models and satellite observations to evaluate the strength of snow albedo feedback. *J. Geophys. Res. Atmospheres*, **117**, D11117, doi:10.1029/2012JD017724.

Hall, A., X. Qu, and J. D. Neelin, 2008: Improving predictions of summer climate change in the United States. *Geophys. Res. Lett.*, **35**, L01702, doi:10.1029/2007GL032012.

Holland, M. M., and C. M. Bitz, 2003: Polar amplification of climate change in coupled models. *Clim. Dyn.*, **21**, 221–232, doi:10.1007/s00382-003-0332-6.

Hong, S.-Y., and E.-C. Chang, 2012: Spectral nudging sensitivity experiments in a regional climate model. *Asia-Pac. J. Atmospheric Sci.*, **48**, 345–355, doi:10.1007/s13143-012-0033-3.

Van den Hurk, B., M. Best, P. Dirmeyer, A. Pitman, J. Polcher, and J. Santanello, 2011: Acceleration of Land Surface Model Development over a Decade of Glass. *Bull. Am. Meteorol. Soc.*, **92**, 1593–1600, doi:10.1175/BAMS-D-11-00007.1.

Koster, R. D., and Coauthors, 2004: Regions of Strong Coupling Between Soil Moisture and Precipitation. *Science*, **305**, 1138–1140, doi:10.1126/science.1100217.

Koster, R. D., and Coauthors, 2010: Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment. *Geophys. Res. Lett.*, **37**, L02402, doi:10.1029/2009GL041677.

Koven, C. D., W. J. Riley, and A. Stern, 2012: Analysis of Permafrost Thermal Dynamics and Response to Climate Change in the CMIP5 Earth System Models. *J. Clim.*, **26**, 1877–1900, doi:10.1175/JCLI-D-12-00228.1.

Qu, X., and A. Hall, 2007: What Controls the Strength of Snow-Albedo Feedback? *J. Clim.*, **20**, 3971–3981, doi:10.1175/JCLI4186.1.

——, and ——, 2014: On the persistent spread in snow-albedo feedback. *Clim. Dyn.*, **42**, 69–81, doi:10.1007/s00382-013-1774-0.

Seneviratne, S. I., D. Luthi, M. Litschi, and C. Schar, 2006: Land-atmosphere coupling and climate change in Europe. *Nature*, **443**, 205–209, doi:10.1038/nature05095.

——, T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling, 2010: Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Sci. Rev.*, **99**, 125–161, doi:10.1016/j.earscirev.2010.02.004.

——, and Coauthors, 2013: Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment. *Geophys. Res. Lett.*, **40**, 2013GL057153, doi:10.1002/grl.50956.

Yoshimura, K., and M. Kanamitsu, 2013: Incremental Correction for the Dynamical Downscaling of Ensemble Mean Atmospheric Fields. *Mon. Weather Rev.*, **141**, 3087–3101, doi:10.1175/MWR-D-12-00271.1.

Application for CMIP6-Endorsed MIPs:
Land-Use Model Intercomparison Project (LUMIP)

Date: March 31, 2015

- **Name of MIP***

Land-Use Model Intercomparison Project (LUMIP)

- **Co-chairs of MIP (including email-addresses)***

George Hurtt (george.hurtt@umd.edu)
David Lawrence (dlawren@ucar.edu)

- **Members of the Scientific Steering Committee***

Almut Arneth (almut.arneth@kit.edu)
Victor Brovkin (victor.brovkin@mpimet.mpg.de)
Kate Calvin (katherine.calvin@pnnl.gov)
Andrew Jones (adjones@lbl.gov)
Chris Jones (chris.d.jones@metoffice.gov.uk)
Peter Lawrence (lawrence@ucar.edu)
Nathalie de Noblet-Ducoudré (nathalie.de-noblet@lsce.ipsl.fr)
Julia Pongratz (julia.pongratz@mpimet.mpg.de)
Sonia Seneviratne (sonia.seneviratne@ethz.ch)
Elena Shevliakova (elena@princeton.edu)

- **Link to website (if available)***

https://www2.cgd.ucar.edu/research/mips/lumip

- **Goal of the MIP and a brief overview***

Human land-use activities have resulted in large changes to the biogeochemical and biophysical properties of the Earth surface, with resulting implications for climate. In the future, land-use activities are likely to expand and/or intensify further to meet growing demands for food, fiber, and energy. CMIP5 achieved a qualitative scientific advance in studying the effects of land-use on climate, for the first time explicitly accounting for the effects of global gridded land-use changes (past-future) in coupled carbon-climate model projections. Enabling this advance, the first consistent gridded land-use dataset (past-future) was developed, linking historical land-use data, to future projections from Integrated Assessment Models, in a standard format required by climate models. Results indicate that the effects of land-use on climate, while uncertain, are sufficiently large and complex to warrant an expanded activity focused on land-use for CMIP6. Land-use change is an essential forcing of the Earth System, and as such LUMIP is directly relevant and necessary for CMIP6 Question 1: "How does the Earth System respond to forcing?" LUMIP will also play a strong role in addressing the WCRP Grand Challenges, particularly with respect to the "AIMES theme for collaboration: biospheric forcings and feedbacks". Due to the broad range of effects of land-use change and the major activities proposed, LUMIP is also of cross-cutting relevance to CMIP6 science questions 2 and 3, and to many of the WCRP Grand Challenges including Climate Extremes, Regional Climate Information, and Water Availability.

The goal of LUMIP is to take the next steps, and enable, coordinate, and ultimately address the most important science questions related to the effects of land-use on climate. The primary science questions of LUMIP are:

- What are the effects of land use and land-use change on climate and biogeochemical cycling (past-future)?
- Are there regional land management strategies with promise to help mitigate and/or adapt to climate change?
- What are the effects of climate change on land-use and land-use change? [*]

In addressing these questions, LUMIP will also address a range of more detailed science questions to get at process level attribution, uncertainty, data requirements, and other related issues in more depth and sophistication for the community than possible to date. Of particular focus will be the separation and quantification of the effects on climate from fossil fuel emissions and land-use change, biogeochemical from biogeophysical effects, the unique impact of land cover change versus land management change, and modulation of land use impact on climate by land-atmosphere coupling strength.

Three major sets of science activities are envisioned. First, a set of metrics and diagnostic protocol will be developed to quantify model performance, and related sensitivities, with respect to land use. As part of this activity, benchmarking data products will be identified to help constrain models. These metrics will be incorporated into the International Land Model Benchmarking (ILAMB) system. This benchmarking/metrics emphasis in LUMIP dovetails with expanding emphasis in CMIP on metrics.

Second, data standardization efforts will build off the lessons learned and protocols in CMIP5, and work with new historical data, present data, IAMS, and ESMs to produce an enhanced standardized land-use data for CMIP6 model experiments passing the maximum amount of common information between these relevant domains. New output data standardization will also enrich and improve analysis of model experiment results. Particular emphasis is on promoting the archival of subgrid land information in CMIP6. In most land models, physical, ecological, and biogeochemical land state and surface flux variables are calculated separately for several different land surface type or land management 'tiles' (e.g., natural and secondary vegetation, crops, pasture, urban, lake, glacier). Frequently, including in the CMIP5 archive, the tile-specific quantities are averaged and only grid-cell mean values are reported. Consequently, a large amount of valuable information is lost with respect to how each surface type responds to climate change and/or direct anthropogenic modifications. LUMIP is developing a proposal outlining the need and protocol for archival for selected key variables on multiple land tiles (see Appendix A for draft proposal).

Third, an efficient model experiment design including both idealized and scenario-based cases has been developed to isolate and quantify land-use effects. These experiments, described in greater detail below, include both idealized and realistic scenario simulations with and without transient land use. The experimental protocol enables integrated analysis of coupled and offline land models (forced with observed meteorology) which will support understanding and assessment of the forced response and climate feedbacks associated with land-use and the relationship of these responses to land and atmosphere model biases.

---

[*] Note that experiments to address this question are not included in this proposal because our understanding is that very few Earth System Models have the capability to address this question yet. We maintain this question within LUMIP because it is a high priority land use change science question that LUMIP will promote through individual model efforts until enough models have the capability to do two-way climate-land use interactions.

LUMIP priorities and model experiments have been developed in close consultation with several existing model intercomparison activities and research programs that focus on the role of land use in climate including LUCID, GSWP3, LUC4C, TRENDY, and AgMIP. In addition, discussions have begun and are ongoing with other proposed CMIP MIPs to ensure that our proposed experiments are complementary and not duplicative. These proposed MIPs include ScenarioMIP, AerChemMIP, C4MIP, LS3MIP, DAMIP, and RFMIP.

- **An overview of the proposed experiments***

LUMIP proposes a two phase, tiered, model experiment plan. Phase one, which can start soon, will feature idealized coupled and land-only model experiments designed to improve process understanding and assess how models represent the impact of changes in land use on climate, as well as to quantify model sensitivity to potential land cover and land use changes. Phase two experiments will be based on historical land use and realistic scenarios identified by ScenarioMIP. Phase two experiments are designed to isolate the role of historical and projected future land-use changes on climate. As there are more possible experiments than are achievable with available resources by all groups, experiments are tiered in order of importance.

Details of the model experiments are included below. The total request includes (all at standard resolution):

Tier 1: 520 years GCM/ESM; 165 years LND-only
Tier 2: 330 years GCM/ESM; 1650 years LND-only
Tier 3: 280 years GCM/ESM; 120 years LND-only

Overview of Phase 1 experiments

Phase 1 consists of two sets of experiments (see Table 1). The first set of experiments are idealized deforestation experiments that enable analysis of the biogeophysical and biogeochemical response to land cover change and the associated changes in climate in a controlled and consistent set of simulations. The idealized deforestation experiment in which as specified area of forest is removed each year for 50 years is new to the land use change modeling community and is designed to be somewhat analogous/complementary to the 1% CO_2 simulations in the DECK. This idealized deforestation experiment has the advantage that it will be easier to ensure conformity across models in terms of the land cover change (differences in the representation of realistic land cover changes across different models is a problem that has plagued prior land cover change model intercomparison projects, e.g. LUCID). Two modeling centers are conducting test idealized deforestation simulations. Additional regional deforestation simulations (Tier 3) are being planned within the LUCID/LUC4C projects. Protocols for these regional experiments will be developed by LUCID and LUC4C.

The second set of Phase 1 experiments are a series of offline land-only simulations, which will build on the LMIP simulation proposed in LS3MIP. This series of experiments is designed to assess how the specification of land cover change and increasingly comprehensive treatment of land management affects the carbon, water, and energy cycle response to land use change. Only a limited number of models will be able to perform all the experiments, but the experimental design will allow for multiple levels of participation, according to each model's capabilities. This set of experiments utilizes state-of-the-art model developments anticipated across several contacted modeling centers and will contribute to the setting of priorities for land use for future CMIPs. Test experiments are planned for 2015 to finalize the experimental design.

It is critical to acknowledge that all observed historic forcing datasets are subject to considerable errors and uncertainty and that the weather and climate trends represented in these datasets may not accurately reflect reality, especially in remote regions where very limited data went into either the underlying reanalysis or the gridded products. These limitations pose a challenge when comparing the model outputs (like latent heat flux, for example) to observed estimates because errors may actually be a function of errors in the forcing dataset rather than the model. The GSWP3 dataset is being put through a thorough analysis and its strengths and limitations, especially with respect to trends, will be documented in one or more papers. LUMIP, in conjunction with LS3MIP, are considering whether or not there is value in requesting runs with an additional forcing dataset (several others century-scale are available) to enable an assessment of forcing uncertainty. Additional offline runs could be an excessive burden on modeling centers so if this is included, it would be in Tier 3, aimed at groups that are interested in pursuing the forcing uncertainty question. Irrespective of the uncertainties, there is strong value in the inclusion of land-only experiments in which all models are forced with the same historical climate since it enables much cleaner comparison across models of the simulated land response to climate change and/or land use change.

Table 1: Phase 1 experiments.

Process understanding	Idealized experiments designed to assess biogeophysical role of land cover change on climate	
idealized_deforest	Idealized deforestation experiment, 20 million km^2 forest removed linearly over a period of 50 years, with an additional 20 years of constant forest cover (Tier 1)	1850-1920
reg_deforest_LND, **reg_deforest_ATM,** **reg_deforest_GCM**	Land, atm, AOGCM simulations with some set of tropical, boreal, or temperate deforestation (defined by LUC4C/LUCID) (Tier 3)	1980-2010

Land cover versus land management change landcover_mange_LND (Tier 2)	Assess relative impact of land cover and incrementally more comprehensive land management change on fluxes of water, energy, and carbon; forced with historical observed climate and projected climate anomalies (1850 to 2014)
LND_LULCC_AM	All LULCC and All Management (AM) features for each particular model turned on; 1700 start; transient CO$_2$, N-dep, aerosol dep, etc.; This run is same as **LMIP-Hist (LS3MIP)** if GCM runs include all management capabilities
LND_LULCC1850	LND_LULCC_AM with land use change starting at 1850 (testing impact of pre-1850 land use)
LND_noLULCC	LND_no land cover change (Same as Tier 1, LND_noLULCC_hist)
LND_grasscrop	LCC with ‘grassland’ crop/pasture; no land management
LND_gross_vs_net	LND_grasscrop except with net transitions instead of gross
LND_fire	LND_grasscrop with human fire management
LND_woodharv	LND_grasscrop or LND_fire with wood harvest
LND_pasture	LND_grasscrop but with grazing on pastureland
LND_crop	LND_grasscrop but with crop area utilizing prognostic crop model
LND_crop-irrig	LND_crop with realistic transient irrigated area
LND_crop-irrig-fert	LND_crop-irrig with realistic transient fertilization

* It is still being evaluated whether additive or subtractive scheme is preferred for these land only offline experiments.

Overview of Phase 2 experiments

The Phase 2 experiments build off of the CMIP6 Historical and historical LMIP simulations as well as the ScenarioMIP simulations. They will include land-only and coupled historical and future simulations with land use held constant or modified to an alternative land use scenario (Table 2). These simulations will be used to assess the role of land use on climate from the perspective of both the biogeophysical and biogeochemical impacts and will be of interest to the Detection and Attribution MIP.

Scientific advantages for particular model configuration:

- Concentration-driven simulations allow focus on biogeophysical impacts on climate and help establish when/where land management could be used as a regional climate mitigation tool.

- Emission-driven simulations allow assessment of the full feedback onto climate and assess whether or not IAM predictions about land use and land use change carbon fluxes are consistent with ESM modeled land use emissions.

- Including experiments in low and medium/high radiative forcing scenarios in concentration-driven scenarios allows examination of how the impact of land use change differs at different levels of climate change and at different levels of CO_2 fertilization.

Consequently, for the projection period, LUMIP includes an additional simulation for both a high and a low radiative forcing scenario with land use from a different SSP-RCP configuration with strongly different land use but with all other forcings remaining the same as in the original ScenarioMIP simulation (e.g., ScenarioMIP SSP3-7 includes strong deforestation; LUMIP experiment will be SSP3-7 but with SSP1-2.6 land use, which is afforestation scenario). Note that these simulations should be considered sensitivity simulations since they will include a set of forcings that are inconsistent with each other (e.g., land use from SSP1-2.6 in a simulation that in all other respects is equivalent to SSP3-7). See figure 1 for further details of the proposed design.

Table 2: Phase 2 experiments.

Land use change impact on land to atmosphere fluxes of water, energy, carbon		
noLULCC_hist_LND (Tier 1)	Same as LMIP-Hist (LS3MIP) except with land use and land cover held constant at 1850, all human impact removed	1850-2014
Land use change impact on past and future climate (Tier 1)		
noLULCC_hist (Tier 1)	Same as **historical CMIP6** except with land cover/use held constant at 1850, concentration-driven (for DAMIP); two additional ensemble members requested in Tier 2	1850-2014
SSP3-7_SSP1-2.6landuse (Tier 1)	Additional land use policy sensitivity simulation for high radiative forcing scenario, keep all forcings the same as ScenarioMIP SSP3-7 (deforestation scenario), but replace land use from SSP1-2.6 (afforestation) scenario;	2015-2100

	concentration-driven; 2 additional ensemble members requested (Tier 3)	
SSP1-2.6_SSP3-7landuse (Tier 1)	Additional land use policy sensitivity simulation for low radiative forcing scenario, keep all forcings the same as ScenarioMIP SSP1-2.6 (afforestation scenario), but replace land use from SSP3-7 (afforestation) scenario; concentration-driven	2015-2100
esmSSP5-8.5_SSP1-2.6landuse	Additional land use polict sensitivity simulation for high radiative forcing scenario, keep all forcings the same as in C4MIP esmssp5-8.5 scenario except use SSP1-2.6 land use; emission driven	

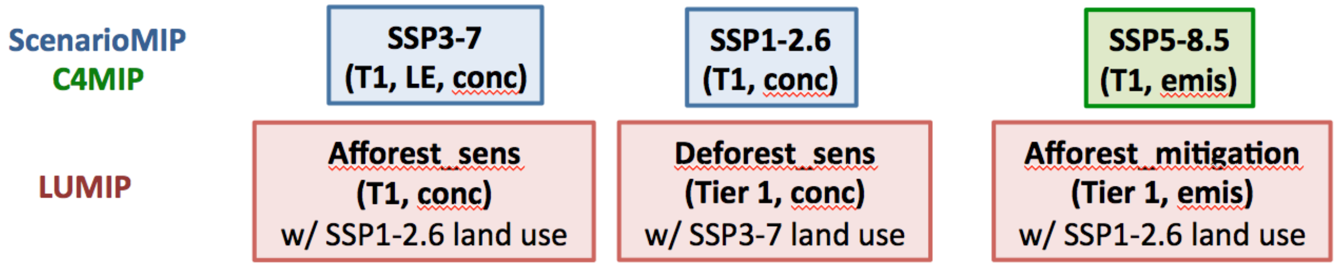


Figure 1: Example set of realistic and sensitivity studies designed to assess potential impact of strongly different land use trajectories on climate outcome.

➢ An overview of the proposed evaluation/analysis of the CMIP DECK and CMIP6 experiments*

The goal is to establish a useful set of model diagnostics that enable a systematic assessment of land use-climate feedbacks and improved attribution of the roles of both land and atmosphere in terms of generating these feedbacks. The need for more systematic assessment of the terrestrial and atmospheric response to land cover change is one of the major conclusions of the LUCID study. Boisier et al. (2012) and de Noblet-Ducoudré et al. (2012) argue that the different land use-climate relationships displayed across the LUCID models highlights the need to improve diagnostics for land surface model evaluation. These analyses need to assess how land surface models respond to a land-cover perturbation in uncoupled (off-line) simulations as well as coupling between land and atmosphere components. One axis of analysis that has previously not been investigated in great detail is how a particular model's regional land-atmosphere coupling strength signature affects how the model simulates the impact of land use change on climate. Here, LUMIP will interface with LS3MIP to investigate the cross-relationship between land-atmosphere coupling strength and land-use change impacts on weather and climate.

In addition, LUMIP will promote the development of biogeophysical and biogeochemical metrics of land use change, based on observations, that will help constrain model dynamics and dovetails with expanding emphasis in CMIP on metrics. Any useful metrics will be integrated into the International Land Model Benchmarking (ILAMB) package that is currently under development. The availability of both land-only and coupled historic simulations enables a much more systematic assessment of the roles of land and atmosphere in the simulated response to land use change.

LUMIP also proposes to develop a set of metrics that quantify a model response to land use across a range of spatial scales and temporal scales that can then be used to quantitatively compare model response across different models, regions, and land management scenarios. For a given variable, say surface air temperature, the diagnostic calculations will be completed for a pair of simulations (offline or coupled) with and without land use change. Across a range of spatial scales, spanning from a single grid cell up to regional (5^{o} by 5^{o} and 10^{o} by 10^{o}) to continental to global, seasonal mean differences between control and land use change simulations will be examined. Differences will be expressed both in terms of seasonal mean differences (and their statistical significance based on student-t tests) and in terms of signal to noise (where ‘noise’ refers to the natural interannual climate variability simulated in the model). Effects on extremes (e.g. Davin et al. 2014) will receive particular attention.

Analysis could focus on critical regions, such as the intensive agricultural region in the central United States and the deforestation region in the Amazon, telescoping out from point to continental scale for each region. Five primary variables will be considered (net radiation, evapotranspiration, temperature, precipitation, and land carbon stocks) that together characterize the biogeophysical and biogeochemical impacts of land use on climate. The first two variables, net radiation and evapotranspiration (ET) define the surface biogeophysical response to land use change and will be evaluated in both offline and coupled model contexts. The temperature and precipitation response to biogeophysical changes in net radiation and ET will be evaluated in land-atmosphere simulations only. Land carbon stocks can be evaluated in offline and coupled simulations.

There are several axes of analysis that can be performed within this framework that are relevant to assessing land use-climate effects relative to natural variability and greenhouse gas-induced climate change. For instance, by varying the number of years and/or the number of ensemble members included in our analysis, one can establish over what time/spatial scale a land use change signal can be detected. One can also investigate the relative difficulty in isolating a land use-climate signal in transient climate simulations with anthropogenic greenhouse gas forcing versus, for example, timeslice atmosphere-land simulations.

- **Proposed timing***

The initial plans for LUMIP have been developed through conference calls and especially during a series of meetings during the summer of 2014.

2013 August 5-9: Initial concept, Aspen
2013 October 3: Presentation of Initial concept, WGCM Meeting
2014 Spring: Workshop 1, GLP Meeting
2014 July 17-18: GEWEX – Biogeophysics
2014 July 21-22: Hamburg – Biogeochemistry
2014 July 28-Aug 1: EMF Snowmass Meeting
2014 August 4-8: AGCI Aspen Joint-MIP Workshop
2014 September 1: Testing of idealized model experiments
2014 September 15: Initial proposal due to CMIP6 Panel
2014 October 8-10: Presentation of revised proposal, WGCM Meeting
2015 January: New prototype land use data/data format released to modeling groups
2015 March 31, final proposal due to CMIP6 Panel
2015 Model I/O and testing with new prototype land use data
2015 October, LandMIP meeting in Zurich to finalize protocol details and begin analysis coordination
2015 Fall, prepare and submit two GMD papers for LUMIP: one on land use dataset and one on experimental protocol and subgrid archiving
2016 January: Initiate multi-model idealized Phase 1 experiments
2016 January: Final land use data made available* (*pending final scenario selection)
2016 Summer: Phase 1 experiments delivered to ESGF
2016 Phase 1 experiments analysis and papers
Starting mid-2016: Phase 2 GCM/ESM realistic experiments, contingent on ScenarioMIP schedule
2018-2019: Model analysis and synthesis, LUMIP SSC will coordinate analysis and submit at least one overview paper to which all modeling center representatives will be invited to join

➢ Selected Key References *

Note that this list of references is representative only. Many additional references on land use and land use change impact on climate are available.

Boisier, J. P., de Noblet-Ducoudré, N., & Ciais, P. (2013). Inferring past land use-induced changes in surface albedo from satellite observations: a useful tool to evaluate model simulations. Biogeosciences, 10, 1501-1516.

Boysen, L. R., Brovkin, V., Arora, V. K., Cadule, P., de Noblet-Ducoudré, N., Kato, E., ... & Gayler, V. (2014). Global and regional effects of land-use change on climate in 21st century simulations with interactive carbon cycle. Earth System Dynamics Discussion, 5, 443-472.

Brovkin, V., Boysen, L., Arora, V. K., Boisier, J. P., Cadule, P., Chini, L., ... & Weiss, M. (2013). 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), 6859-6881.

Davin, E.L., S.I. Seneviratne, P. Ciais, A. Olioso, and T. Wang, 2014: Preferential cooling of hot extremes from cropland albedo management. *Proc. Natl Acad. Sci.*. Published ahead of print June 23, 2014, doi:10.1073/pnas.1317323111.

Hurtt, G. C., Chini, L. P., Frolking, S., Betts, R. A., Feddema, J., Fischer, G., ... & Wang, Y. P. (2011). 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), 117-161.

Kumar, S., Dirmeyer, P. A., Merwade, V., DelSole, T., Adams, J. M., & Niyogi, D. (2013). Land use/cover change impacts in CMIP5 climate simulations: A new methodology and 21st century challenges. Journal of Geophysical Research: Atmospheres, 118(12), 6337-6353.

Lawrence, P. J., Feddema, J. J., Bonan, G. B., Meehl, G. A., O'Neill, B. C., Oleson, K. W., ... & Thornton, P. E. (2012). Simulating the biogeochemical and biogeophysical impacts of transient land cover change and wood harvest in the Community Climate System Model (CCSM4) from 1850 to 2100. Journal of Climate, 25(9), 3071-3095.

de Noblet-Ducoudré, N., Boisier, J. P., Pitman, A., Bonan, G. B., Brovkin, V., Cruz, F., ... & Voldoire, A. (2012). Determining robust impacts of land-use-induced land cover changes on surface climate over North America and Eurasia: results from the first set of LUCID experiments. Journal of Climate, 25(9), 3261-3281.

Pitman, A. J., de Noblet□Ducoudré, N., Cruz, F. T., Davin, E. L., Bonan, G. B., Brovkin, V., ... & Voldoire, A. (2009). Uncertainties in climate responses to past land cover change: First results from the LUCID intercomparison study. Geophysical Research Letters, 36(14).

Pongratz, J., Reick, C. H., Raddatz, T., Caldeira, K., & Claussen, M. (2011). Past land use decisions have increased mitigation potential of reforestation. Geophysical Research Letters, 38(15).

➢ Appendix A: Sub-grid archiving of selected land output for CMIP6

➢

Co Task Leads: Elena Shevliakova and David Lawrence

Note: A detailed protocol for sub-grid archiving, including instructions about how models should deal with ‘edge cases’ (e.g., what if a model doesn’t represent one of the land use tiles) is being developed and will be circulated among the modeling centers. A final protocol will be included in the GMD special issue paper and will also be posted on the LUMIP website.

1. Motivation

The majority of CMIP5-class climate models and Earth system models (ESMs) represent land sub-grid spatial heterogeneity by splitting each land grid into sections (i.e. tiles or units) with similar ecological, biogeochemical, and hydrological characteristics. Current land components capture two kinds of sub-grid heterogeneity: 1) hydrological - land surfaces covered by liquid or frozen water (e.g. lakes, wetlands, glaciers) or not (e.g. bare and vegetated surfaces) and 2) land-use and land management induced (e.g. cropland, pastures, urban, natural and secondary, i.e., harvested forests, plantations, abandoned land). Sub-grid tiling applies to both above- and below-ground sections of the land components. Physical, ecological, and biogeochemical land state variables and surface fluxes are calculated separately for each tile. However, frequently, including in the CMIP5 archive, the tile-specific variables were averaged and the grid-cell mean values were reported. Consequently, a large amount of valuable information was lost with respect to how each surface type with different hydrological and land-use properties responds to climate change and/or direct anthropogenic modifications.

In order to better characterize surface climate, its variability and change, we propose to expand the CMOR data convention in order to capture horizontal land sub-grid heterogeneity. In addition to the land-grid cell values, we propose to request a subset of selected variables on multiple land tiles. This reporting and archiving modification will significantly expand the utility of Earth System Model output for scientific analysis and climate change impacts studies.

Each land model has a unique tiling scheme so the archiving protocol needs to be general enough to work for the range of existing model structures.

2. Proposed sub-grid reporting

2.1 Types of tiles

In the context of CMIP6 we propose to report tile-specific information *for a subset of 4 categories* to capture land-use induced surface heterogeneity: (1) Natural and Secondary land types (including bare ground and vegetated wetlands), (2) pasture-land, (3) croplands, and (4) urban. The remaining tiles, such as lakes, rivers and glaciers, will be excluded from the reported tile-specific values. The proposed set of land-use tile reporting units closely corresponds to land-use units to be used in the CMIP6 historical land-use reconstructions and future scenarios. Primary (i.e., natural vegetation never affected by LULCC activity) and secondary vegetation (i.e., natural vegetation that has previously been harvested or establishes on abandoned agricultural lands) are combined because most land models do not yet distinguish between these two land types.

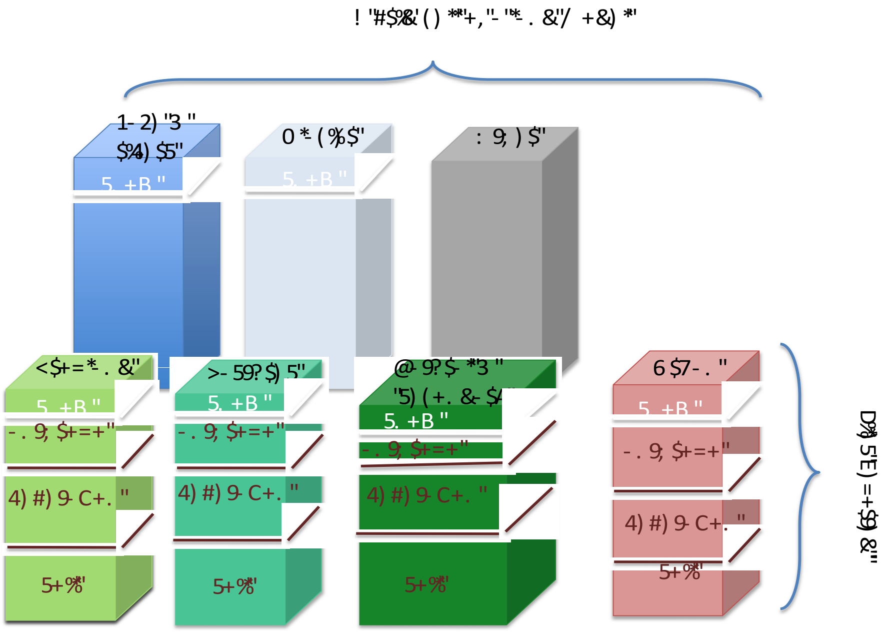


Figure 1.Proposed reporting structure for land-tiles

Tile type	***Tile Suffix***	***Comment***
Urban and rural settlement	*“urb”*	
Primary and secondary lands	*“psl”*	*Forest, grasslands, etc or bare*
Croplands	*“crp”*	
Managed pasturelands	*“pst”*	

For selected key variables, data is requested for each tile separately, in addition to the grid cell mean. The tiles containing biogeochemical information will report up to four stocks of biogeochemical tracers (e.g. Carbon, Nitrogen) – vegetation, soil, litter, and anthropogenic storage. The latter is used in a subset of land models and reflects the fact that some harvested carbon is not released into the atmosphere immediately, but rather with some time-delay from a year to century (e.g., wood products, food)

2.2 Variables reported by tile

Subgrid tile variables should be submitted according to the following structure, using Leaf Area Index (LAI) as an example:

laiLut(lon, lat, time, landusetype4)

landusetype4 dimension with the order spl, crp, pst, urb where "spl" = secondary and primary land, "crp" = cropland, "pst" = pastureland, and "urb" = urban.

Notes for tiled variable reporting: Sum of fractional area for urb+spl+crp+pst may not add up to 1 for grid cells with lakes or glaciers. If a model does not represent one of the requested tiles, then it should be reported as missing value. In cases where more than one land use tile shares information (e.g., if pastureland and cropland share same soil column), then duplicate information should be provided for both tiles. Further details on tile reporting will be included in the LUMIP GMD paper.

A limited set of variables are requested (see LUMIP data request spreadsheet) including variables that describe (a) the subgrid structure and how it evolves through time, (b) biogeochemical fluxes, (c) biogeophysical variables, and (d) LULCC carbon transfers, and (e) carbon stocks on land use tiles.

We recognize that models have very different implementation of LU processes and may only be able to report a subset of variables.

Ocean Model Inter-comparison Project (OMIP): Application for CMIP6-Endorsed MIPs

13 March 2015

Name of MIP: Ocean Model Inter-comparison Project (OMIP)

Co-chairs of OMIP

Gokhan Danabasoglu, NCAR, USA (gokhan@ucar.edu)
Stephen M. Griffies, NOAA/GFDL, USA (stephen.griffies@noaa.gov)
James Orr, IPSL, France (james.orr@lsce.ipsl.fr)

Scientific Steering Committee

Physical processes (CLIVAR OMDP and collaborators)

1. Claus Boning (Germany)
2. Eric Chassignet (USA)
3. Enrique Curchitser (USA)
4. Gokhan Danabasoglu (USA)
5. Helge Drange (Norway)
6. Stephen Griffies (USA)
7. David Holland (USA)
8. Yoshiki Komuro (Japan)
9. William Large (USA)
10. Simon Marsland (Australia)
11. Simona Masina (Italy)
12. George Nurser (UK)
13. Anna Pirani (CLIVAR ICPO)
14. Anne-Marie Treguier (France)
15. Mike Winton (USA)
16. Stephen Yeager (USA)

Chemical and biogeochemical processes

1. James Orr (France)
2. Laurent Bopp (France)
3. Scott Doney (USA)
4. John Dunne (USA)
5. Fortunat Joos (Switzerland)
6. Galen McKinley (USA)
7. Andreas Oschlies (Germany)
8. Toste Tanhua (Germany)
9. Keith Lindsay (USA)

Websites

OMIP merges the CORE-II and OCMIP projects:

1. Coordinated Ocean-ice Reference Experiments phase II (CORE-II):
   http://www.clivar.org/clivar-panels/omdp/core-2
2. Ocean Carbon Model Inter-comparison Project OCMIP2:
   http://ocmip5.ipsl.jussieu.fr/OCMIP/phase2/simulations/CFC/HOWTO-CFC.html
3. Ocean Carbon Model Inter-comparison Project OCMIP3:
   http://ocmip5.ipsl.jussieu.fr/OCMIP/phase3/simulations/NOCES/HOWTO-NOCES.html

Coordinating CMIP6 ocean analyses & providing an ocean/sea-ice comparison

OMIP provides a framework for evaluating, understanding, and improving ocean/sea-ice/tracer/biogeochemical components of climate and earth system models contributing to CMIP6. OMIP addresses this goal in two related manners.

1. **Ocean Diagnostics**: OMIP coordinates CMIP6 ocean diagnostics (including ocean physics, inert chemical tracers, and biogeochemistry) for *all* CMIP6 simulations that include an ocean component. Such CMIP6 simulations using an ocean component include, but are not limited to, the CMIP6/DECK, CMIP6/historical, CMIP6/OMIP ocean/sea ice, CMIP6/C4MIP, and CMIP6/FAFMIP perturbation simulations.
2. **Global Ocean/sea-ice Simulation**: OMIP provides a protocol for two global ocean/sea-ice simulations forced with common atmospheric data sets (one Tier 1 and one Tier 2). The OMIP ocean/sea-ice simulations include physical, inert chemical, and biogeochemical tracer components.

OMIP provides a framework for the following types of studies:

1. To investigate physical, chemical, and biogeochemical mechanisms that drive seasonal, inter-annual, and decadal variability;
2. To attribute ocean-climate variations to boundary forced versus natural;
3. To evaluate robustness of mechanisms across models and forcing data sets;
4. To bridge observations and modeling by complementing ocean reanalysis from data assimilation;
5. To provide consistent ocean and sea-ice states useful for initialization of climate (e.g., decadal) predictions.

Merging OMIP and OCMIP

In response to the WGCM comments provided 18 Feb 2015, this updated OMIP proposal merges with the previously separate OCMIP (Ocean Carbon Model Inter-comparison Project). There is no longer a separate OCMIP proposal for CMIP6. The merger of ocean physical, chemical, and biogeochemical efforts into a single project, OMIP, allows for efficient communication across ocean physical, chemical, and biogeochemical groups participating in CMIP6. Hence, the name "OMIP" used in this document refers to the ocean physical, chemical, and biogeochemical aspects of CMIP6 simulations.

I. OMIP: Diagnostic analysis of CMIP6 ocean components

The first role for CMIP6/OMIP is to coordinate diagnostic analysis for all CMIP6 experiments that involve an ocean component. As part of this role, CLIVAR OMDP and collaborators produced a CMIP6 ocean model diagnostic document (Griffies et al. 2015), which built from the earlier CMIP5 document (Griffies et al. 2009). The Griffies et al (2015) document (submitted in this proposal as *CMIP6_OMIP_physics_diagnostics.pdf*) presents recommendations and scientific justifications for sampling physical ocean fields included as part of *all* CMIP6 simulations that include an ocean component.

In addition, we have produced three diagnostic spreadsheets, also submitted with this OMIP proposal. These spreadsheets list ocean diagnostics to be saved from CMIP6 experiments that include an ocean component. Each spreadsheet started from the most recent CMIP5 version for the ocean. Entries to the OMIP spreadsheets were edited to remove obsolete diagnostics, to clarify earlier CMIP5 diagnostics where needed, and to add new diagnostics.

1. **Ocean physics** (*CMIP6_OMIP_physics_standard_output.xls*): The first spreadsheet is based on the diagnostics detailed in Griffies et al (2015), which focuses on physical ocean diagnostics. These diagnostics should be saved for all ocean models run as part of *any* CMIP6 experiment where an ocean component is included.
2. **Ocean inert chemistry** (*CMIP6_OMIP_chemistry_standard_output.xls*): Three inert chemical species, CFC11, CFC12, and SF_6, are requested for CMIP6/historical, CMIP6/OMIP-A, and CMIP6/OMIP-B simulations. These tracers build on past phases of OCMIP, most recently efforts undertaken during CMIP5. They help in evaluating elements of the ocean simulation through comparisons with the observational databases.
3. **Ocean biogeochemistry** (*CMIP6_OMIP_biogeochemistry_standard_output.xlsx*): Ocean biogeochemical diagnostics should be saved from *all* CMIP6 experiments involving comprehensive earth system models, as well as the CMIP6/OMIP-A and CMIP6/OMIP-B experiments described in this document. The ocean biogeochemical components of CMIP6 will be evaluated and compared, with some emphasis placed on assessing what is gained by moving to finer grid spacing. In addition, coupled and forced simulations with the same CMIP6 ocean biogeochemical models will be compared to assess how mean state and internal variability differ.

The union of these diagnostic spreadsheets represents the full suite of CMIP6 ocean related diagnostics as coordinated through OMIP.

II. OMIP: Providing ocean/sea-ice simulations for inter-comparison

The second portion of the OMIP project consists of one Tier 1 (OMIP-A) and one Tier 2 (OMIP-B) forced global ocean/sea ice simulation. Aspects of these two simulations are detailed in this section.

The physical component of OMIP is based on CORE-II

The physical component of the OMIP ocean/sea ice simulation follows the forcing protocol of Griffies et al. (2012), which has been used for the Coordinated Ocean-ice Reference Experiments (CORE) inter-annually varying simulation (CORE-II).

- CORE-II simulations use atmospheric data sets compiled by Large and Yeager (2009) to force physical ocean fields.
- CORE-II forcing covers the 62-year period from 1948-2009.
- CORE-II atmospheric forcing data sets are collaboratively supported by the U.S. National Center for Atmospheric Research (NCAR) and the NOAA Geophysical Fluid Dynamics Laboratory (NOAA/GFDL).
- All data sets, codes for the bulk flux formulae, technical report, and other support codes along with the release notes are available at http://www.clivar.org/clivar-panels/omdp/core-2

- CORE-II studies are coordinated by the WCRP Climate Variability and Predictability (CLIVAR) Ocean Model Development Panel (OMDP) (OMDP was formerly known as the Working Group on Ocean Model Development, WGOMD).

More than 20 international ocean and climate modelling groups have performed CORE-II hindcast simulations. The simulations are being analyzed in roughly ten separate studies, each focusing on a specific aspect of the ocean climate system. Examples include analysis of mean states in the North Atlantic with a focus on the Atlantic Meridional Overturning Circulation (Danabasoglu et al., 2014) and an assessment of global and regional sea level changes (Griffies et al., 2014). Manuscripts documenting further CORE-II studies will be published during 2015 in a special issue of the journal **Ocean Modelling**.

The inert chemical tracer component of OMIP is based on OCMIP2

The chemical component of OMIP includes the simulation of CFC11, CFC12, and SF_6 online. CFC11 is high priority (level=1), whereas CFC12 and SF_6 are level=2 priority. Simulation protocols are based on the OCMIP2 protocol available at

http://ocmip5.ipsl.jussieu.fr/OCMIP/phase2/simulations/CFC/HOWTO-CFC.html

The SF_6 simulation is identical to the CFC simulations, except that its atmospheric history, solubility, and Schmidt number differ. These inert passive-tracers are computed online along with the physical tracers temperature and salinity, and they are independent of the biogeochemical model. These species should be included for the CMIP6/historical and CMIP6/OMIP simulations.

The biogeochemical component of OMIP is based on OCMIP3

Details of OCMIP3 are available at

http://ocmip5.ipsl.jussieu.fr/OCMIP/phase3/simulations/NOCES/HOWTO-NOCES.html

Each group will use its own prognostic ocean biogeochemical model, coupled online to its physical ocean component. Each physical-biogeochemical simulation will be forced as described above with the CORE-II atmospheric state, if possible over the full 310 years of the five forcing cycles.

- Boundary conditions for the atmosphere include an imposed constant atmospheric concentration of O_2 (mole fraction of 0.20946), but a variable atmospheric CO_2 that follows observations.
- Models will add two tracers for a parallel preindustrial world where atmospheric CO_2 is fixed at 278 ppm. Thus biogeochemical models will include a second dissolved inorganic carbon (DIC) tracer and a second total alkalinity tracer (ALK) in order to isolate natural CO_2 and keep track of model drift.
- All models will be initialized with available climatologies for biogeochemical fields for oxygen as well as total dissolved inorganic nitrogen, phosphorus, and silica will rely on the World Ocean Atlas (2013).

- The DIC and ALK fields will be initialized from the GLODAPv2 gridded data (based on discrete measurements during WOCE and CLIVAR).
- Initial fields for iron will be based on the GEOTRACES database, with a gridded version of this database provided as part of OMIP.
- Initialization of other tracers is less critical, e.g., phytoplankton biomass is restricted to the top 200 m and equilibrates rapidly as do other biological tracers.
- Direct coupling between simulated chlorophyll and ocean dynamics is not mandatory. That is, the optical model used to compute the penetration of solar radiation in the water column is not required to be the same in the dynamical and biogeochemical components.
- Groups are free to use their preferred boundary conditions for the different sources of nutrients to the ocean (via atmospheric deposition, river discharge, and sediment mobilization) as well as lateral input of carbon (e.g., from river and groundwater discharge).
- All forcing files will be updated through 2009 on the OCMIP web sites.

Summary of the OMIP-A and OMIP-B experiments

The Tier 1 OMIP simulation (OMIP-A) consists of one global ocean/sea-ice/inert-chemical/biogeochemical experiment run for a minimum of five repeating cycles of the CORE-II forcing. OMIP-A involves two paths:

- The groups that are unable to run with biogeochemistry can participate in the physical/chemical portion. Here, potential (or Conservative) temperature, salinity and ideal age are required. In addition, the inert chemical tracers are included with CFC11 as a required field and CFC12 and SF_6 as optional variables.
- The second path adds an online biogeochemistry model to the first path.

With 62-years of CORE-II forcing data repeated five times, the total integration length is 310 years. Primary analyses will focus on the fifth cycle. However, archiving results for all five cycles is encouraged, e.g., in order to assess drift. This simulation is a Tier 1 experiment.

The Tier 2 OMIP simulation (OMIP-B) is requested from the same groups having biogeochemistry and able to afford a millennial-scale spin-up. Rather than using observed climatologies to initialize the biogeochemistry as in OMIP-A, this simulation will be initialized with model tracer fields that have been spun up for at least 1000 years, ideally for 5000 years. In addition, to evaluate deep-ocean circulation the OMIP-B simulation (and its spin-up) will include radiocarbon (abiotic DIC and $DI^{14}C$ following the OCMIP2 protocol), unlike the Tier 1 OMIP-A simulation. Otherwise, OMIP-B will be identical to OMIP-A.

Details of the atmospheric forcing datasets and the experimental protocol for the physical fields are available in Griffies et al. (2012) and Danabasoglu et al. (2014), and they follow the CORE-II protocol. Inert chemical tracers follow OCMIP2, and biogeochemical tracers follow OCMIP3.

1. **Initialization**:
   a. Potential temperature (or Conservative temperature) and salinity are initialized using January-mean observational-based climatology from a state of rest.

   b. Inert anthropogenic chemical tracers (CFC11, CFC12, SF_6) are initialized to zero on 01January 1936 (which occurs during the 4th CORE-II forcing cycle at model date 01 January 0237).
   c. Biogeochemical tracers are initialized according to details given above.
   d. Sea ice fields are generally initialized from an existing state taken from another simulation.
2. **Forcing**:
   a. Surface heat fluxes are determined by radiative fluxes from CORE-II (Large and Yeager 2009). Turbulent heat fluxes are computed based on the ocean state and CORE-II atmospheric state.
   b. Bulk formulae for turbulent fluxes must follow the Large and Yeager (2009) method.
   c. There is no restoring term applied to the surface temperature.
   d. Surface water fluxes are provided by CORE-II, including precipitation and river runoff.
   e. Surface salinity is generally damped to a monthly observational-based climatology. Each model group determines details for the salinity restoring used for their model. A unified salinity restoring method for all models is not feasible due to physical sensitivities in ocean-ice models related to high latitude processes (Griffies et al., 2009).
   f. Atmospheric inert-chemical tracers (CFC11, CFC12, and SF_6) are held to their preindustrial values (zero) for the first three cycles of the CORE-II forcing, then changed according to the OCMIP2 protocol starting at model date 01 January 0237, corresponding to calendar date 01 January 1936. Alternatively, to save computational resources, the inert chemical traces could be activated only on that date, starting from zero concentrations in the atmosphere and ocean.
   g. Biogeochemical tracers are forced according to OCMIP3.

Planned evolution of the OMIP simulation protocol

As proposed here, the two CMIP6/OMIP experiments (OMIP-A and OMIP-B) use the Large and Yeager (2009) forcing data sets, following the Griffies et al (2012) protocol. Additionally, new atmospheric forcing data products, anticipated in 2016, will be incorporated into future phases of CMIP6/OMIP. Having multiple forcing products will enable a more robust understanding of sensitivities to forcing, and will provide feedback to those developing the observational-based atmospheric products. The OMIP Scientific Steering Committee will update the OMIP protocol as new developments and extensions become available.

III. Closing remarks

The oceanography and climate science communities have expressed great interest in analyzing ocean output from *all* CMIP6 simulations that include an ocean model component. Such analyses include physical, chemical, and biogeochemical diagnostic elements. OMIP provides the framework for CMIP6 ocean diagnostics by detailing diagnostics to be output from CMIP6 simulations using an ocean component. Furthermore, the two (one Tier 1 and one Tier 2) OMIP global ocean/sea ice experiments allow for systematic inter-comparison of the ocean/sea-ice/chemical/biogeochemical components of CMIP6 climate and earth system models. Ocean and climate modeling groups participating in CMIP6/OMIP experiments will be keenly interested in evaluating their ocean simulations using the OMIP ocean/sea ice protocol. This evaluation will help to improve our understanding of ocean/sea-ice/chemical/biogeochemical components used in coupled climate and earth system models.

References

Danabasoglu, G., S. G. Yeager, D. Bailey, E. Behrens, M. Bentsen, D. Bi, A. Biastoch, C. Boning, A. Bozec, V. Canuto, C. Cassou, E. Chassignet, A. C. Coward, S. Danilov, N. Diansky, H. Drange, R. Farneti, E. Fernandez, P. G. Fogli, G. Forget, Y. Fujii, S. M. Griffies, A. Gusev, P. Heimbach, A. Howard, T. Jung, M. Kelley, W. G. Large, A. Leboissetier, J. Lu, G. Madec, S. J. Marsland, S. Masina, A. Navarra, A. J. G. Nurser, A. Pirani, D. Salas y Melia, B. L. Samuels, M. Scheinert, D. Sidorenko, A.-M. Treguier, H. Tsujino, P. Uotila, S. Valcke, A. Voldoire, and Q. Wang, 2014: North Atlantic simulations in Coordinated Ocean-ice Reference Experiments phase II (CORE-II). Part I: Mean states. *Ocean Modelling*, **73**, 76-107, doi:10.1016/j.ocemod.2013.10.005.

Griffies, S. M., A. Biastoch, C. Boning, F. Bryan, G. Danabasoglu, E. P. Chassignet, M. H. England, R. Gerdes, H. Haak, R. W. Hallberg, W. Hazeleger, J. Jungclaus, W. G. Large, G. Madec, A. Pirani, B. L. Samuels, M. Scheinert, A. S. Gupta, C. A. Severijns, H. L. Simmons, A. M. Treguier, M. Winton, S. Yeager, and J. Yin, 2009: Coordinated Ocean-ice Reference Experiments (COREs). *Ocean Modelling*, **26**, 1-46, doi:10.1016/j.ocemod.2008.08.007.

Griffies, S. M., A. Adcroft, H. Aiki, V. Balaji, M. Bentson, F. Bryan, G. Danabasoglu, S. Denvil, H. Drange, M. England, J. Gregory, R. Hallberg, S. Legg, T. Martin, T. J. McDougall, A. Pirani, G. Schmidt, D. Stevens, K. Taylor, and H. Tsujino, 2009: Sampling physical ocean fields in WCRP CMIP5 simulations. ICPO Publication Series 137, WCRP Informal Report No. 3/2009.

Griffies, S. M., M. Winton, B. Samuels, G. Danabasoglu, S. Yeager, S. Marsland, H. Drange, and M. Bentsen, 2012: Datasets and protocol for the CLIVAR WGOMD Coordinated Ocean-ice Reference Experiments (COREs). *WCRP Report* No. 21/2012, pp.21.

Griffies, S. M., J. Yin, P. J. Durack, P. Goddard, S. C. Bates, E. Behrens, M. Bentsen, D. Bi, A. Biastoch, C. W. Boning, A. Bozec, E. Chassignet, G. Danabasoglu, S. Danilov, C. M. Domingues, H. Drange, R. Farneti, E. Fernandez, R. J. Greatbatch, D. M. Holland, M. Ilicak, W. G. Large, K. Lorbacher, J. Lu, S. J. Marsland, A. Mishra, A. J. G. Nurser, D. Salas y Melia, J. B. Palter, B. L. Samuels, J. Schroter, F. U. Schwarzkopf, D. Sidorenko, A. M. Treguier, Y.-H. Tseng, H. Tsujino, P. Uotila, S. Valcke, A. Voldoire, Q. Wang, M. Winton, and X. Zhang, 2014: An assessment of global and regional sea level for years 1993-2007 in a suite of interannual CORE-II simulations. *Ocean Modelling*, **78**, 35-89, doi: 10.1016/j.ocemod.2014.03.004.

Griffies, S. M., A. J. Adcroft, V. Balaji, G. Danabasoglu, P. J. Durack, P. J. Gleckler, J. M. Gregory, J. P. Krasting, T.J. McDougall, R. J. Stouffer, and K. E. Taylor, 2015: Sampling the physical ocean in CMIP6 simulations. (draft)

Large, W. G., and S. G. Yeager, 2009: The global climatology of an interannually varying air-sea flux data set. *Climate Dynamics*, **33**, 341-364, doi: 10.1007/s00382-008-0441-3.

Application for CMIP6-Endorsed MIPs

Please return to CMIP Panel Chair Veronika Eyring (email: Veronika.Eyring@dlr.de)

Proposals from MIPs should include the following information:

**	Preliminary information used to determine whether a MIP should be endorsed for CMIP6 or not.*

***	Information that must be provided later (and before the panel can determine which experiments, if any, will be incorporated in the official CMIP6 suite).*

- Name of MIP*

  Paleoclimate Modeling Intercomparison Project

  (This will be the fourth phase of PMIP)

- Co-chairs of MIP (including email-addresses)*

  - Pascale Braconnot (Pascale.Braconnot@lsce.ipsl.fr)
  - Sandy Harrison (s.p.harrison@reading.ac.uk)

- Members of the Scientific Steering Committee*

  - Pascale Braconnot / LSCE, France (model and model-data)
  - Sandy P. Harrison / University of Reading, UK and Macquarie University, Australia (data and model-data)
  - Ayako Abe-Ouchi / AORI, University of Tokyo (ice-sheet and PCMIP)
  - Pat Bartlein / University of Oregon, USA (Continental data)
  - Alan Haywood / University of Leeds, UK (Mid-pliocene)
  - Sylvie Joussaume / LSCE, France
  - Johann Jungclaus / MPI-M, Germany (Last millennium)
  - Michal Kucera / MARUM, Germany (Ocean data)
  - Bette Otto-Bliesner / NCAR, USA (warm climates)
  - Gilles Ramstein / LSCE, France (glacial and ice sheet)
  - Karl Taylor / PCMDI, USA (Link with CMIP5)
  - Paul Valdes / BRIDGE, UK (abrupt changes)

- Link to website (if available)*

http://pmip3.lsce.ipsl.fr + http://pmip.lsce.ipsl.fr (PMIP1) and http://pmip2.lsce.ipsl.fr (PMIP2)

Goal of the MIP and a brief overview*

Since the 1990s, PMIP has developed with the following objectives:

- to evaluate the ability of climate models used for climate prediction in simulating well-documented past climates outside the range of present and recent climate variability
- to understand the mechanisms of these climate changes, in particular the role of the different climate feedbacks

To achieve these goals, PMIP has actively fostered paleo-data syntheses, multi-model analyses, including analyses of relationships between model results from past and future simulations, and model-data comparisons. These have first been focusing on the results from Atmospheric General Circulation Models (PMIP1) and then been extended to coupled Ocean-Atmosphere General Circulation Models and AOGCM including carbon cycle feedbacks, thereby closely following model developments for CMIP (PMIP2 and PMIP3). Three PMIP3 simulations were part of the CMIP5 ensemble of simulations: the last millennium, the mid-Holocene (~6,000 years ago) and the Last Glacial Maximum (~21,000 years ago), hence allowing, for the first time, the rigorous comparison of model results for past and future climates. The rationale for considering these periods was:

- for the Last Glacial Maximum, to evaluate the models on a well-documented climatic extreme, especially in terms of temperatures, and study the role of forcings and feedbacks in establishing this climate;
- for the mid-Holocene, to evaluate and analyse the models on a climate "optimum" for the northern hemisphere, characterized by enhanced monsoons, extra-tropical continental aridity and much warmer summers;
- for the last millennium, to study the mechanisms (natural variability vs impact of solar, volcanic and anthropogenic forcings) of decadal to centennial climate variability and evaluate the models' performance w.r.t numerous detailed records.

For CMIP6, we propose to include two new warm periods in the PMIP/CMIP set of experiments: the Last Interglacial and the Mid-Pliocene, for which simulations have been performed and significantly contributed to AR5.

PMIP3/CMIP5 and PlioMIP have been very successful in terms of participation and publications. 19 groups have contributed to PMIP3/CMIP5 simulations, 12 groups have taken part in PlioMIP. PMIP3/CMIP5 simulations have been used in more than 40 publications (as of Sept 11th, 2014) and PlioMIP simulations have been the topic of more than 20 publications. PMIP simulations have brought strong contribution to 2 IPCC AR5 chapters: chapter 5 "information from paleoclimate archives" and chapter 9 "evaluation of climate models".

PMIP simulations specifically address CMIP6 key question on "How does the Earth System respond to forcing" for a variety of forcings and with possible comparisons to data for climates states very different from the current or historical climate. PMIP also addresses question 2 ("What are the origins and consequences of systematic model biases?") about systematic model biases, with the perspective given by documented climates different from today: PMIP simulations, with comparisons to data, can help assessing whether the biases for present-day are also found for other climate states and whether present-day biases have an impact on the magnitude of simulated climate changes. Finally, PMIP is relevant for question 3 ("How can we assess future climate changes given climate variability, predictability and uncertainties in scenarios?"), by examining these very questions for documented past climate cases and via the use of the last millennium simulations as reference state for natural variability.

PMIP simulations are being analyzed within the Grand Challenge "Clouds, Circulation and Climate Sensitivity". They can also provide valuable input for other grand challenges, such as those on the Cryosphere and on Regional Climate Information, with the challenge of paleoclimate modelling at fine scale. Indeed, PMIP model output is increasingly used in "paleo-impact studies", on biodiversity or on understanding the potential impact of climate and environmental changes on early Humans. Several initiatives have already been proposed along these themes and will be reinforced in the future (e.g. Future Earth "Fast Track Initiatives and Cluster Activities" project "Making better use of the Paleoclimate Modeling Intercomparison Project simulations (MAPS)" led by P. Braconnot, a project concerning WGCM, PAGES, CLIVAR, CLiC and bioDISCOVERY).

The five proposed experiments constitute a reference ensemble for further studies within PMIP: single forcing experiments, transient experiments (testing the models on abrupt climate change and on glacial-interglacial transitions).

- References*

PMIP1:

Joussaume, S. and Taylor, K. E., 1995. Status of the Paleoclimate Modeling Intercomparison Project (PMIP), Proceedings of the first international AMIP scientific conference (Monterrey, California, USA, 15-19 May 1995), WCRP report 92, 425-430. Text available at: https://pmip.lsce.ipsl.fr/publications/overview.html

PMIP2:

Braconnot, P., B. Otto-Bliesner, S. Harrison, S. Joussaume, J.-Y. Peterchmitt, A. Abe-Ouchi, M. Crucifix, E. Driesschaert, T. Fichefet, C. D. Hewitt, M. Kageyama, A. Kitoh, A. Laîné, M.-F. Loutre, O. Marti, U. Merkel, G. Ramstein, P. Valdes, S. L. Weber, Y. Yu, and Y. Zhao, 2007. Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum – Part 1: experiments and large-scale features. Climate of the Past, 3, 261–277, www.clim-past.net/3/261/2007/.

Braconnot, P., B. Otto-Bliesner, S. Harrison, S. Joussaume, J.-Y. Peterchmitt, A. Abe-Ouchi, M. Crucifix, E. Driesschaert, T. Fichefet, C. D. Hewitt, M. Kageyama, A. Kitoh, M.-F. Loutre, O. Marti, U. Merkel, G. Ramstein, P. Valdes, L.Weber, Y. Yu, and Y. Zhao, 2007. Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum – Part 2: feedbacks with emphasis on the location of the ITCZ and mid- and high latitudes heat budget. Climate of the Past, 3, 279–296, www.clim-past.net/3/279/2007/.

PMIP2/PMIP3:

Braconnot, P., S. P. Harrison, M. Kageyama, P. J. Bartlein, V. Masson-Delmotte, A. Abe-Ouchi, B. Otto-Bliesner and Y. Zhao, 2012. Evaluation of climate models using palaeoclimatic data, Nature Climate Change, DOI: 10.1038/NCLIMATE1456

Schmidt, G.A., J. D. Annan, P. J. Bartlein, B. I. Cook, E. Guilyardi, J. C. Hargreaves, S. P. Harrison, M. Kageyama, A. N. LeGrande, B. Konecky, S. Lovejoy, M. E. Mann, V. Masson-Delmotte, C. Risi, D. Thompson, A. Timmermann, L.-B. Tremblay, and P. Yiou, 2014. Using paleo-climate comparisons to constrain future projections in CMIP5, Climate of the Past, 10, 221-250

An overview of the proposed experiments*

The following table summarizes the experiments proposed by PMIP for CMIP6. These experiments all build from the DECK experiments and are part of the core of PMIP simulations (~10), which will themselves constitute a basis for other PMIP experiments (sensitivity analyses, transient simulations starting from the core ones). Within PMIP, each PMIP working group will organize their set of simulations, as PMIP also federates focused MIPs such as PlioMIP on the Pliocene climate, LIGMIP on the Last Interglacial, PAST2K on the last two millennia.

Table 1: summary of proposed experiments. In yellow: PMIP3/CMIP5 experiments. In green: new experiments for CMIP6. The PMIP3/CMIP5 experiment names in the ESFG nomenclature are indicated in italic below each period name.

Period	**Purpose**	**Imposed boundary conditions**	**# of years**
Last millennium (*past1000*) 850-1850 CE	a) Evaluate the ability of models to capture observed variability on multi-decadal and longer time-scales. b) Determine what fraction of the variability is attributable to "external" forcing and what fraction reflects purely internal variability. c) Provides a longer-term perspective for detection and attribution studies	• Solar variations • Volcanic aerosols • Atmospheric concentration of well mixed greenhouse gases • Land use • Orbital parameters	1000 (after spin-up period)
Mid-Holocene (*midHolocene*) 6 kyr ago	a) Compare with paleodata the model response to known orbital forcing changes and changes in greenhouse gas concentrations. b) Relationships between changes in mean state and variability	• Orbital parameters • Atmospheric concentration of well-mixed greenhouse gases	≥100 (after spin-up period)
Last Glacial Maximum (*lgm*) 21 kyr ago	a) Compare with paleodata the model response to ice-age boundary conditions. b) Attempt to provide empirical constraints on global climate sensitivity.	• Ice-sheet and land-sea mask • Atmospheric concentration of well-mixed greenhouse gases • Orbital parameters	≥100 (after spin-up period)
Last Interglacial *lastInterglacial* 127 kyr ago	a) Evaluate climate model for warm period, high sea-level stand b) Impacts of climate on sea ice and ice sheets	• Orbital parameters • Atmospheric concentration of well-mixed greenhouse gases	≥100 (after spin-up period)
Mid-Pliocene Warm Period 3.2 Ma ago	a) How does the Earth System respond in the long term to CO_2 forcing analogous to that of the modern? b) What is the significance of CO_2 induced polar amplification for the stability of the ice sheets, sea-ice and sea-level?	• Ice-sheet and land-sea mask, topography (smaller ice-sheets) • Atmospheric concentration of well-mixed greenhouse gases • Orbital parameters	≥100 (after spin-up period)

For all these periods the model to be used is the same as the one used for future climate projections. Therefore depending on the groups the model will be only atmosphere-ocean coupled models or Earth System models. The reference for the analyses will be the CMIP6 pre-industrial simulation. Hereafter, we shortly describe the Mid-Holocene and Last Glacial Maximum, periods which have already been a focus of PMIP since its start and which have been part of the PMIP3-CMIP5 simulations. More details are given below on the Last Millennium (part of PMIP3-CMIP5 as well) and on the two new periods proposed for CMIP6: the Last Interglacial and the Mid-Pliocene Warm Period.

➢ Mid-Holocene (*midHolocene*) and Last Glacial Maximum (*lgm*):

The mid-Holocene (~6000 years ago) and the Last Glacial Maximum (LGM, ~21000 years ago) constitute the most recent quasi-stable climatic extremes: the mid-Holocene is often described as a warm state, or "climate optimum", in which dominant features of the global hydrological cycle, such as the North African and Asian monsoon, were amplified; the LGM is a cold extreme in which greenhouse gas concentrations were at their minimum and continental ice-sheet at their maximum size, covering large areas of northern North America and northwestern Eurasia.

These periods have been the focus for paleo-data syntheses since the beginning of the PMIP project and therefore are well documented in terms of temperature, hydrological cycle and land surface type. Some long standing model-data disagreement are echoing preoccupations for future climate change, such as the systematic underestimation of the northward penetration of the African monsoon rainfall onto the continent compared to available records for the Mid-Holocene. The LGM is relevant for studying feedback mechanisms at work in establishing a temperature response as large as (although with an opposite sign) as that predicted for the end of the 21st century. Both periods constitute test cases for our understanding of mechanisms of climate change, such as the interplay between circulation changes and radiation/cloud changes, the respective strengths of feedbacks from different components of the climate system, and for our understanding of the connections between global and regional climate changes.

Compared to the previous phases of PMIP a particular amphasis will be put on the impact of dust on the mean climate and climate feedbacks, as well as on uncertainties in boundary conditions or surface feedbacks related to the vegetation or interactive carbon cycle.

The reference experiments for both the *midHolocene* and *lgm* simulations are the pre-industrial control and it is very interesting to compare those experiments with an idealized experiment designed to study mechanisms of future climate change, such as abrupt4xCO2. PMIP-CMIP6 will benefit from idealized experiments proposed by CFMIP, such as AMIPminus4K or abrupt0.5CO2 which will help comparing feedbacks at work in setting up a cold climate vs. those at work for a warm climate. Similarly, an AMIP experiment with insolation prescribed at a 6ky BP value will be very useful to analyze the strengths of forcings and feedbacks within the climate system and the mechanisms for common/different responses for past and future climates. These sensitivity experiments will be discussed as part of PMIP in the Past to Future working group. They would echo PMIP1 simulations (http://pmip.lsce.ipsl.fr), while ESM simulations would echo PMIP2 and PMIP3 (http://pmip2.lsce.ipsl.fr and http://pmip3.lsce.ipsl.fr) simulations, hence allowing a characterization of the models' evolution in their ability to represent documented large climate changes.

➢ Last Millennium (*past1000*):

The last millennium is the best-documented period of climate change in a multi-century time frame. Climate has varied considerably during the late Holocene and these changes left their traces in history (Medieval Climate Optimum, Little Ice Age). However, the relative magnitude of natural fluctuations due to internal variability of the Earth's climate system and to variations in the external forcings (Sun, orbital, volcanic) and the present global warming, attributed to anthropogenic greenhouse gases, is still under debate. Simulations of the last millennium (LM) therefore directly address the first CMIP6 key scientific question "How does the Earth System respond to forcing". Investigating the response to (mainly) natural forcing under climatic background conditions not too different from today is crucial for an improved understanding of climate variability, circulation, and regional connectivity. LM simulations also allow assessing climate variability on decadal and longer scales and provide information on predictability under forced and unforced conditions. These are crucial for near-term predictions and thus address the third CMIP5 scientific question "How can we assess future climate changes given climate variability, predictability and uncertainties in scenarios". In providing in-depth model evaluation with respect to observations and paleoclimatic reconstructions in particular addressing details of response to forcing, LM simulations serve to "understand origins and consequences of systematic model biases", thus addressing also the second CMIP6 scientific question.

LM will build on DECK experiments, in particular the pre-industrial control simulation as unforced reference and the historical simulations. Moreover, LM provide initial conditions for historical simulations starting in the 19th century that are considered superior to the *piControl* state as it includes integrated information from the forcing history (e.g. large volcanic eruptions in the early 19th century).

Within PMIP, a considerable number of individual researchers and modelling groups is committed to perform LM simulations. The simulations will base on experience gained in PMIP3/CMIP5 where more than a dozen modelling groups participated and a total of 15 LM experiments where stored in the ESGF database. Several studies, partly reflected by entries in the AR5 chapter 5, have highlighted the value of the LM multi-model ensemble. The PMIP3 LM working group (WG Past2K) is closely cooperating with the PAGES initiative PAGES2k promoting regional reconstructions of climate variables and variability modes. Collaborative work has focused on reconstruction-model intercomparison (e.g. Bothe et al., 2013) and assessment of variability modes (e.g. Raible et al., 2014). Integrated assessment of reconstruction and simulations has led to progress in model evaluation and process understanding (e.g. Lehner et al., 2013; Sicre et al., 2013; Jungclaus et al., 2014). WG Past2K will promote future common analyses and workshops bringing together observational and modelling expertise.

For CMIP6 progress is expected owing to new, more comprehensive reconstructions of volcanic forcing (Sigl et al., in preparation), improved models, and an experimental protocol that ensures seamless simulations from the pre-industrial past to the future. Higher-resolution simulations will allow assessing more regional details and processes, e.g. storm-tracks, precipitation.

➢ Last Interglacial:

The Summary for Policymakers for both the IPCC WG1 AR4 and AR5 included statements on the Last Interglacial (LIG):

> *"There is very high confidence that maximum global mean sea level during the last interglacial period (129,000 to 116,000 years ago) was, for several thousand years, at least 5 m higher than present, and high confidence that it did not exceed 10 m above present. During the last interglacial period, the Greenland ice sheet very likely contributed between 1.4 and 4.3 m to the higher global mean sea level, implying with medium confidence an additional contribution from the Antarctic ice sheet. This change in sea level occurred in the context of different orbital forcing and with high-latitude surface temperature, averaged over several thousand years, at least 2°C warmer than present (high confidence)."*

Yet the AR4 and AR5 had no coordinated simulations for the LIG to assess the interplay of polar amplification of temperature, seasonal memory of sea ice, and precipitation/storm track changes on the stability of the Greenland ice sheet and its contribution to the sea level high stand nor the interplay of oceanic and atmospheric temperatures and circulation on the stability of the Antarctic ice sheet. Climate model simulations for the LIG assessed in the AR5, although completed by many modeling groups, varied in their forcings and often were not made with the same model/same resolution as the CMIP5 future projections, thus providing a useful but incomplete means for assessment (Chapter 5; Lunt et al., 2013). Similarly, Greenland ice sheet simulations assessed in the AR5 used offline models with a variety of climate forcing setups (Chapter 5). No simulations were available to assess the Antarctic ice sheet (particularly, the West Antarctic Ice Sheet) contribution to the LIG sea level high stand.

We propose a PMIP-CMIP6 time-slice experiment for the LIG (*lastInterglacial*) to determine the interplay of warmer atmospheric and oceanic temperatures, changed precipitation, and changed surface energy balance on ice sheet thermodynamics and dynamics during this period. The output from these time-slice experiments will be used by ISMIP6 to force standalone ice sheet experiments (*lastIntergacialforcedism*). Still uncertain are how well ice sheet and climate models can predict the stability of the ice sheets and if thresholds may be passed this century. A LIG simulation will be of high societal relevance because of implications for sea level changes as well as sea ice and monsoons. The LIG is the most suitable of the warm past interglacials for a CMIP6 assessment because of the relative wealth of data including: ice cores providing measurements of well-mixed greenhouse gases, aerosols including dust and sea salt, and stable water isotopes as a proxy for temperature, as well as for Greenland, ice sheet elevation and extent; marine records for ocean temperatures and geotracers that can be interpreted in terms of water masses and overturning strength; speleothems that provide indication of monsoon strength; and terrestrial records that indicate temperature and vegetation. As well, new records are refining our knowledge of sea ice extent, fire, and biodiversity.

The proposed PMIP-CMIP6 simulation for the LIG is particularly relevant to the WCRP Grand Challenges: Changes in Cryosphere and Regional Sea-level Rise, but also to Regional Climate Information and Clouds, Circulation and Climate Sensitivity because of the large forcings and thus large regional responses as recorded in the data. It addresses well the broad scientific questions: 1. How does the Earth System respond to forcing? and 2. What are the origins and consequences of systematic model biases (especially at high latitudes and relevant to the stability of the ice sheets)? As part of PMIP and ISMIP6, some groups will additionally perform transient coupled ice sheet-climate simulations as well as other standalone ice sheet simulations that will provide rates of change for sea level, including regional sea level if offline GIA models applied, as well as a measure of the capability of these models to initiate the next glacial inception.

➢ Pliocene warm period

The Pliocene epoch was the last time in Earth history when atmospheric CO_2 concentrations approached modern values (~400 ppmv) whilst at the same time retaining a near modern continental configuration. The IPCC 5th Assessment report chapter 5 (Masson-Delmotte et al., 2013) states that that model–data comparisons for the Pliocene provide high confidence that mean surface temperature was warmer than pre-industrial (Dowsett et al., 2012; Haywood et al., 2013). Global mean sea surface temperatures have been estimated to be +1.7°C above the 1901–1920 mean based on large data syntheses (Lunt et al., 2010; Dowsett et al., 2012). Exiting climate model simulations have produced a range of global mean surface air temperature of +1.9°C and +3.6°C relative to the 1901–1920 mean (Haywood et al., 2013). Model simulations have indicated that meridional temperature gradients were reduced (due to high latitude warming), which has significant implications for the stability of polar ice sheets and sea level in the future (e.g. Miller et al. 2012). Compilations of vegetation (Salzmann et al., 2008) have indicated that the global extent of arid deserts decreased and boreal forests replaced tundra, and climate models predict an enhanced hydrological cycle, but with a large inter-model spread (Haywood et al., 2013). The East Asian Summer Monsoon, as well as other monsoon systems, may also have been enhanced (Zhang et al. 2013). Although climate model simulations for the Pliocene were assessed in the AR5, these simulations were not derived from the same model/same resolution as the CMIP5 future projections, thus reducing the communities' ability to assess and compare changes in global and regional Pliocene climates, vis-à-vis similar predictions of future climate change (Haywood et al., 2013).

We propose a CMIP6 time-slice experiment for the Pliocene to understand the long term response of the Earth's climate system to a near modern concentration of atmospheric CO_2 (longer term climate sensitivity or Earth System Sensitivity), and to understand the response of ocean circulation, Arctic sea-ice, modes of climate variability (e.g. ENSO), as well as the global response in the hydrological cycle and regional changes in monsoon systems. A Pliocene simulation will be of high societal relevance because of its potential to inform policy makers on required emission reduction scenarios designed to prevent global annual mean temperatures increase by more than 2 to 3 °C in the long term (beyond 2100 AD).

The proposed CMIP6 simulation for the Pliocene is relevant to two of the WCRP Grand Challenges. This includes Clouds, Circulation and Climate Sensitivity because of the enhanced CO_2 forcing (contemporaneous with modern CO_2 forcing), providing a unique opportunity to examine an equilibrium climate state to a near modern concentration of atmospheric CO_2. The pattern of polar amplification preserved Pliocene climate archives can be compared directly with the latest generation of CMIP models making a valuable contribution towards addressing the potential polar amplification problem. Through the analysis of Pliocene polar amplification in CMIP models, and examining the geological interpretation of a seasonally sea-ice free Arctic Ocean during the Pliocene, our CMIP6 simulation will also address the WCRP Grand Challenge of Changes in the Cryosphere. Whilst uncertainty exists in Pliocene sea level reconstruction, IPCC AR5 states with high confidence that Pliocene sea-levels were higher than the pre-industrial era, with a number of independent methods indicating a sea-level rise of between 10 and 20 m. This indicates potential long term instability of both the Greenland and Antarctic Ice Sheets (Miller et al. 2012) with CO_2 concentrations at approximately 400 ppmv.

CMIP6 Pliocene experiments will be used within the Pliocene Ice Sheet Model Intercomparison Project in order to better constrain the climatological forcing in ice sheet model simulations for the Pliocene in the future. There is a well-organized and highly active of community of Pliocene climate modellers within PMIP, with the Pliocene working group being one of the most successful working groups within PMIP3. The working group is closely associated with the United States Geological Survey (USGS) who has had a highly productive core program focused on Pliocene environmental reconstruction for the last 25 years, and their data has been used to underpin almost all model-data comparisons performed for the Pliocene. Thus, CMIP6 can expect a high degree of continued support and new Pliocene data sets from the USGS for comparison with model outputs.

The experiment will address the broad scientific questions: 1 How does the Earth System respond in the long term to CO_2 forcing analogous to that of the modern? and 2 What is the significance of CO_2 induced polar amplification for the stability of the ice sheets, sea-ice and sea-level?

An overview of the proposed evaluation/analysis of the CMIP DECK and CMIP6 experiments*

PMIP benefits from strong links with PAGES and INQUA (International Union for Quaternary Science) for model-data comparisons.

midHolocene* and *lgm: evaluation w.r.t available data (systematic benchmarking, cf. Harrison et al, Climate Dynamics, 2013), both in terms of temperature and hydrological cycle. These evalutions make use of independent climate reconstructions over land and ocean. A specific focus will be put on the link with model biases and model results for future climate. Specific working groups in PMIP have been set up to improve the comparisons with marine data (COMPARE group) and isotopic data (cf. http://pmip3.lsce.ipsl.fr/, "working groups" tab). This provides new methodologies and new possibilities for quantitative model assessments.

***past1000*:** In-depth analyses using novel statistical approaches (Sundberg et al., 2012; Moberg et al., 2014; Bothe et al., 2013) and detection/attribution techniques (Schurer et al., 2014).

Process-oriented analyses on variability and changes in circulation modes. Partly supported by dedicated sensitivity studies, e.g. in VolMIP.

***lastInterglacial*:** The CMIP6 experiment will analyse the strength of feedbacks at work in the Arctic, and their potential implications for the stability of the Greenland ice sheet. A particular emphasis will be put on the annual redistribution heat by the ocean circulation and the potential role of the transmission of the subsurface warming from North Atlantic to Southern ocean, with implication for basal melting of West Antarctic ice Sheet. High latitude feedbacks from sea-ice, water vapor and cloud will be a focus, as well as the relative changes between the tropical and high latitude water cycle.

Pliocene Warm Period: The CMIP6 experiment will evaluate the ability of models to simulate a recent interval of CO_2-induced global warmth, and assess the response of critical components of the climate system to near modern CO_2 forcing in the long term (sea-ice, modes of variability, monsoons, storm tracks, vegetation). Unlike other warm periods or interglacials the Pliocene retains critical modern boundary conditions such as the continental configuration and astronomical forcing. The signal of change in Pliocene is large and therefore the signal to uncertainty ratio enables model-predicted changes to be attributed with confidence.

Some of these diagnostics and model evaluations will be performed as part of PMIP transverse analyses groups. In particular, the PMIP "Past2Future" working group aims at identifying and understanding relationships between model simulations for past and future climates and at using available paleodata to evaluate the consistency of these relationships. Its work is therefore potentially based on all PMIP simulations together with selected simulations relevant for future climate change.

Proposed timing*

Ideally past1000 should be run before the historical simulations and forcings should be continuous over the transition between the past1000 and historical runs.

All other experiments can be run as soon as the reference simulation in DECK is run.

Experimental design of proposed CMIP6 experiments

The baseline protocol is ready and described below. We have opened discussions on this protocol on the PMIP web page:

https://wiki.lsce.ipsl.fr/pmip3/doku.php/pmip3:cmip6:design:index

The final experimental design, decided after interactions with the modelling groups, will be published in GMD. We plan to open a PMIP special issue to describe the CMIP6 experiments and how they articulate with other PMIP simulations. Participating modelling groups will be involved in these description papers so as to give them first credit of their contribution.

➢ midHolocene and lgm

(taken from Braconnot et al, Nature Climate Change, 2012) + http://pmip3.lsce.ipsl.fr

Mid-Holocene (MH) and Last Glacial Maximum (LGM) simulations are equilibrium experiments, presenting a “snapshot” of climate at a specific time. Table 2 summarises the boundary conditions used for MH and LGM experiments during the various phases of PMIP. The ultimate external forcing (or driver) of climate is change in incoming solar radiation (insolation) as determined by changes in the Earth’s orbit. These changes can be specified precisely. Due to the slow variations of Earth’s orbital parameters, the seasonal and latitudinal distribution of MH insolation was different from present (1950 C.E), enhancing the magnitude of the seasonal contrast in the Northern Hemisphere by about 60 Wm^{-2}. Insolation forcing at the LGM was very similar to present. When models do not explicitly simulate slow processes such as the build up of ice sheets, concomitant changes in land-sea distribution, or the evolution of atmospheric composition, all of which lead to changes that have to be considered as climate forcings on shorter timescales, then these boundary conditions (hereafter forcings) have to be prescribed in the MH and LGM experiments. As models have evolved in complexity, so the set of forcings that has to be prescribed has also evolved. In the first phase of the PMIP (PMIP1), the experiments were performed with atmospheric general circulation models and the state of the ocean was prescribed as a forcing. In the second phase of PMIP (PMIP2), some models incorporated vegetation dynamics but vegetation cover and albedo still had to be specified for the coupled ocean-atmosphere general circulation models (OAGCMs). Some processes, such as those associated with the terrestrial and marine carbon cycle, have been ignored in the earlier PMIP experiments, but will be included as interactive components of some of the models used in PMIP3. In all experiments the atmospheric composition is prescribed using results from ice-cores.

The next phase of PMIP will make use of the PMIP3 boundary conditions whenever possible. A major foreseen evolution is related to the interactive computation of the dust cycle in atmospheric models, for which changes in vegetation also have to be taken into account. PMIP2 and PMIP3 recommended the use of either interactive vegetation or prescribed pre-industrial vegetation. For PMIP-CMIP6, those models which include an interactive representation of the dust cycle will have to account for changes in vegetation. This particular topic will be discussed with the modelling groups.

The *lgm* experiment will be the reference from which sensitivity experiments to uncertainties in boundary conditions will be developed. In particular, the sensitivity to ice sheet reconstructions will be tested within the PMIP working group on LGM ice sheet uncertainties. These PMIP-CMIP6 LGM experiments will also be starting points for transient deglaciation experiments (coordinated by the working group on the deglaciation).

Table 2 : Evolution of the boundary conditions prescribed in the different phases of the PMIP project. Boundary conditions that remain the same between different sets of simulations are highlighted in yellow; blue highlighting shows boundary conditions that are not included in a given set of experiments. More details on the protocols used in PMIP3 can be found on the PMIP3 web site (see http://pmip3.lsce.ipsl.fr/), which also provides links to the webpages detailing the protocols used in PMIP1 and PMIP2. Note that in the MH experiment the CO2 concentration is the pre-industrial one. CO2ctrl refers to the CO2 concentration of the present-day control simulation.

	PMIP1	PMIP2	PMIP3
Mid Holocene (6000 years BP)* *In this experiment ice-sheet, coastline, solar constant and aerosols are prescribed as in the PI simulation.			
Insolation	eccentricity = 0.018682 **obliquity** = 24.105° perihelion-180° = 0.87°	eccentricity = 0.018682 **obliquity** = 24.105° perihelion-180° = 0.87°	eccentricity = 0.018682 **obliquity** = 24.105° perihelion-180° = 0.87°
Trace gases	**CO_2** = 280 ppm or 280/345* CO_2ctrl **CH_4** = 650 ppb **N_2O** = 270 ppb CFC = 0 **O_3** = not considered	**CO_2** = 280 ppm **CH_4** = 650 ppb **N_2O** = 270 ppb CFC = 0 **O_3** = not considered	**CO_2** = 280 ppm **CH_4** = 650 ppb **N_2O** = 270 ppb CFC = 0 **O_3** = same as in CMIP5 PI
Vegetation and land surface	Prescribed to be the same as modern vegetation	Either prescribed to be the same as modern vegetation or computed using a dynamical vegetation module	Computed using a dynamical vegetation module, Or prescribed as in PI, with phenology computed for models with active carbon cycle or prescribed from data
Carbon cycle	Not considered	Not considered	Interactive, with atmospheric concentration prescribed and ocean and land carbon fluxes diagnosed as recommended in CMIP5
Last Glacial Maximum (21000 years BP) * * In this experiment solar constant and aerosols are prescribed as in the PI simulations.			
Insolation	eccentricity = 0.018994 **obliquity** = 22.949° perihelion-180° = 114.42°	eccentricity = 0.018994 **obliquity** = 22.949° perihelion-180° = 114.42°	eccentricity = 0.018994 **obliquity** = 22.949° perihelion-180° = 114.42°

Trace gases	**CO_2** = 200 ppm or (200/280) * CO_2ctrl **CH_4** = 350 ppb **N2O** = 190 ppb CFC =0 **O_3** = same as in PI	**CO_2** = 185 ppm **CH_4** = 350 ppb **N_2O** = 200 ppb CFC =0 **O_3** = same as in PI	**CO_2** = 185 ppm **CH_4** = 350 ppb **N_2O** = 200 ppb CFC =0 **O_3** = same as in PI
Ocean	SST prescribed from CLIMAP (1981) Or SST computed using a slab ocean model	3D Ocean model and sea-ice	3D ocean model and sea-ice
Ice sheet	ICE-4G (Peltier et al, 1994)	ICE-5G (Peltier et al, 2004)	PMIP3 Blended ice sheet
Land-sea mask	–105 m sea level	Prescribed following Peltier (2004) land-sea mask –120 m	Prescribed from the blended ice-sheet land-sea mask. Sea-level change consistent with the change in land-sea mask.
Freshwater		Excess LGM freshwater added to the ocean in 3 different regions	Excess LGM freshwater added to the ocean in 3 different regions
Ice sheet ice streams	Not considered	Not considered	Not considered
River runoff	Not considered	As in CTRL or river pathway modified	As in PI or river pathway modifier according to PMIP protocol
Mean ocean salinity	Not considered	Not considered	+1 PSU everywhere
Carbon cycle	Not considered	Not considered	Interactive, with atmospheric concentration prescribed and ocean and land carbon fluxes diagnosed as recommended in CMIP5 For PCMIP: fully interactive with atmospheric concentration computed by the model

➢ Last Millennium

Updated PMIP3 protocol (http://pmip3.lsce.ipsl.fr) based on Schmidt et al (2011, 2012):

Schmidt, G. A. et al. (2011). Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1.0), Geosci. Model Dev., 4, 33–45.

Schmidt, G. A. et al. (2012). Climate forcing reconstructions for use in PMIP simulations of the Last Millennium (v1.1), Geosci. Model Dev., 5, 185–191

Transient simulations 850-1849 followed by historical experiments, set of boundary conditions for solar, volcanic, land-cover-change, greenhouse gases to be blended with those for historical

(1850-2010) simulations. The continuity between the past1000 and historical scenarios has to be improved and fully discussed within CMIP6.

Last Interglacial

Based on the protocol discussed within PMIP3.

For CMIP6, we propose a simulation for the 127ky BP time slice - large orbital forcing (Figure 1), large responses.

- Orbital parameters set to 127ka.
- Greenhouse gas concentrations well-known from ice cores [CO_2 275ppm; CH_4 709ppb; N_2O 266 ppb].
- modern geography, ice sheets, and vegetation;
- Initialize from CMIP6 pre-industrial DECK simulation;
- length: ≥ 100 years after spinup.

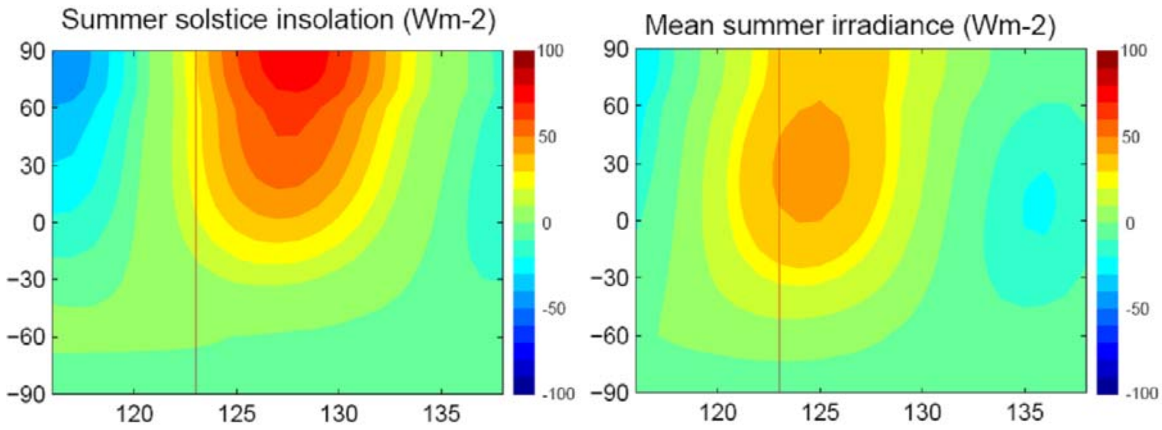


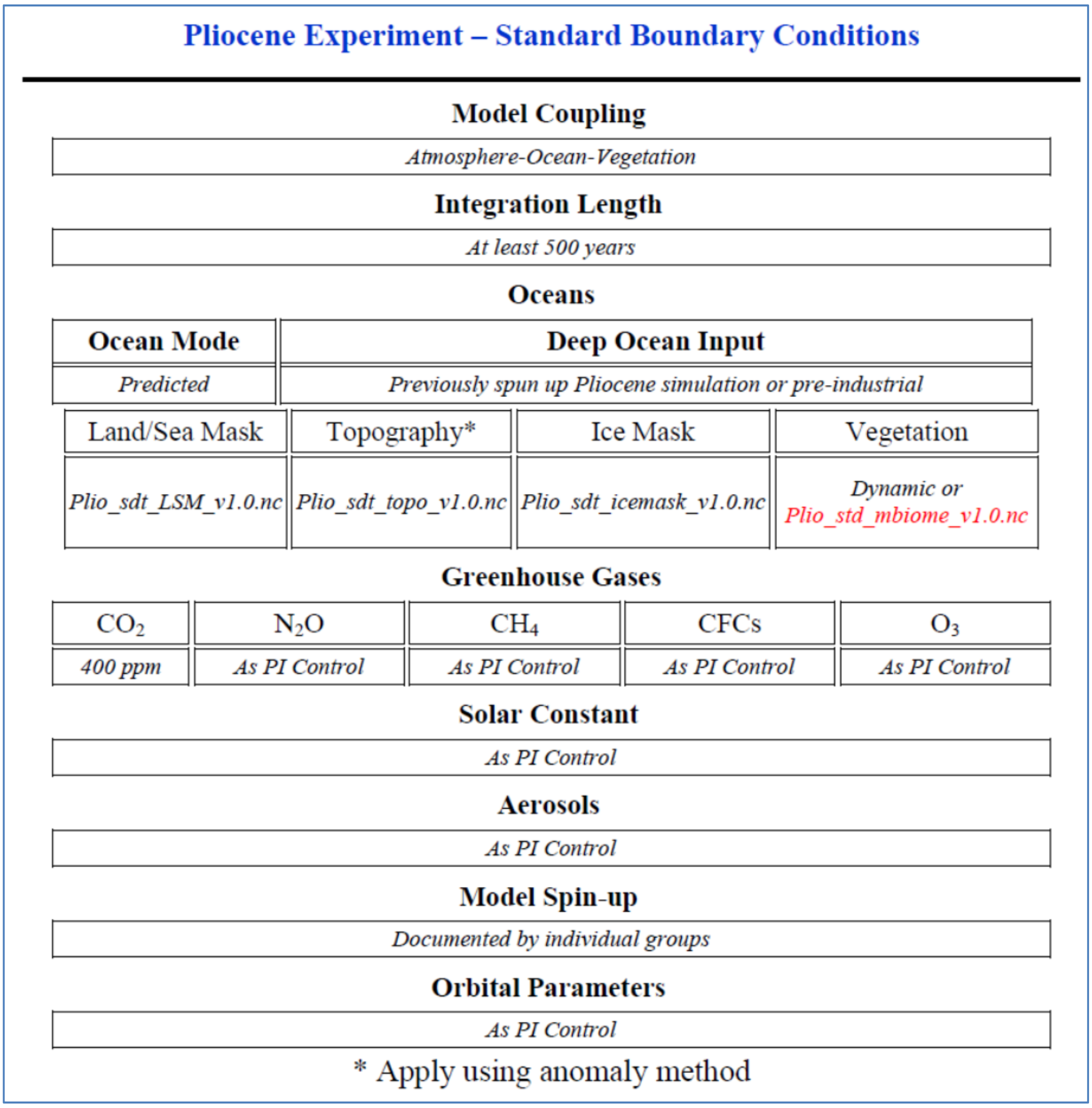
Figure 1. Anomalies of summer solstice insolation (left) and mean summer irradiance (right) as compared to present. Left axes are latitude and bottom axes are in thousands of years before present.

This simulation will constitute a reference for additional PMIP LIG simulations: other time slices within the last interglacial (125, 122 ky BP), transient simulations for the whole interglacial. This will also be a target period for testing AOGCMs coupled with polar ice-sheet models, in coordination with ISMIP6.

Pliocene warmth

Time slice equilibrium climate experiment modifying CO_2 (to 400 ppmv), topography, ice sheet extent and running with dynamic vegetation.

Updated from PlioMIP experiment 2 (Haywood et al, GMD, 4, 571-577, 2011). The proposed PMIP-CMIP6 experiment is the “standard experiment” on the PlioMIP 2 site, with boundary conditions summarized in the following table.



Pliocene Experiment – Standard Boundary Conditions

Model Coupling

Atmosphere-Ocean-Vegetation

Integration Length

At least 500 years

Oceans

Ocean Mode	**Deep Ocean Input**
Predicted	*Previously spun up Pliocene simulation or pre-industrial*

Land/Sea Mask	Topography*	Ice Mask	Vegetation
Plio_sdt_LSM_v1.0.nc	Plio_sdt_topo_v1.0.nc	Plio_sdt_icemask_v1.0.nc	Dynamic or Plio_std_mbiome_v1.0.nc

Greenhouse Gases

CO_2	N_2O	CH_4	CFCs	O_3
400 ppm	*As PI Control*	*As PI Control*	*As PI Control*	*As PI Control*

Solar Constant

As PI Control

Aerosols

As PI Control

Model Spin-up

Documented by individual groups

Orbital Parameters

As PI Control

* Apply using anomaly method

As for the other proposed CMIP6 experiments, this Pliocene experiment is the basis for a full range of experiments coordinated within PMIP by the PlioMIP working group. In particular, the sensitivity of the results to insolation, ice sheet configuration and other boundary conditions will be investigated. Modelling groups wishing to take part in these plans are advised to visit the PlioMIP2 web site: http://geology.er.usgs.gov/egpsc/prism/7.2_pliomip2_data.html.

Science question and/or gap addressed with the PMIP experiments

Cf. introduction and summary excel table.

New foci for analyses will be:

- Forced vs. internal variability, putting in context climate changes in the industrial historical period
- Clouds/Circulation: WCRP Grand challenge Initiative on Leveraging the past record (http://www.wcrp-climate.org/index.php/gc-clouds-circulation-activities/gc4-clouds-initiatives/116-gc-clouds-inititative4)
- Analyses of cryospheric feedbacks under natural forcings (transient simulations over the last millennium put in perspective the recent changes e.g. in Arctic Sea ice, coupling between ice-sheets and climate (*lgm*, *lastInterglacial*, Pliocene)
- Regional climate and decadal predictions:
    - Improved assessment of decadal to centennial variability as carrier of near-term prediction potential (*past1000*, *midHolocene*, *lgm*).
    - Regional assessment of response to natural forcing and interaction with variability modes and teleconnections (all experiments)
- Assessment of extremes under natural forcing, e.g. volcanoes. Natural variations in droughts in connection with paleo-reconstructions (*past1000*). Analyses of mechanisms of mega-droughts (*midHolocene*). → link with Grand challenges on extremes and on water availability.

Possible synergies with other MIPs

PMIP simulations can serve to interact with other MIPs on the following themes:

- CF-MIP (cloud feedbacks): dedicated common idealized sensitivity experiments to be run in aquaplanet set up: AMIP simulations with SSTs minus 4K, abrupt0.5xCO2, abrupt solar perturbation experiments, to be co-analysed in CF-MIP and PMIP.
- ICE SHEETS (ISMIP6): Assessment of the climate and cryosphere interactions and the sea level changes associated with large ice sheets. In particular, the Last Interglacial simulations will be used to force ice sheet models in ISMIP6. Additional experiments co-designed by the PMIP and ISMIP groups are foreseen outside the CMIP6 exercise: transient interglacial experiments, with climate model output forcing an ice sheet model, and coupled climate-ice sheet experiments.
- OCEAN/SEA_ICE: Mutual assessment of the role of the ocean in low-frequency variability, e.g. multi-decadal changes in ocean heat content or heat transport. Provide initial conditions for the ocean including long-term forcing history.
- CARBON CYCLE (C4MIP): Assessment of carbon-cycle evolution and feedbacks between sub-components of the Earth System. Evaluation of paleo reconstructions of carbon storage.
- LAND USE: Links should be reinforced for better connecting past1000 to historical simulations. Useful for analysis of past1000 simulations, for biophysical as well as carbon cycle aspects.
- VolMIP (volcanic forcing): analysis of specific volcanic events very useful for critical analysis of past1000 simulations. VolMIP would systematically assess uncertainties in the climate response to volcanic forcing, whereas LM simulations describe the climate response to volcanic forcing in long transient simulations where related uncertainties are due to chosen input data for volcanic forcing: mutual assessment of forced response.
- SOLARMIP: similar interests as for VolMIP, interactions are foreseen for the analysis of the past1000 experiments.
- DETECTION/ATTRIBUTION: long millennium simulations can be very useful for this topic.

Potential benefits of the experiments

Potential benefits of the experiments to

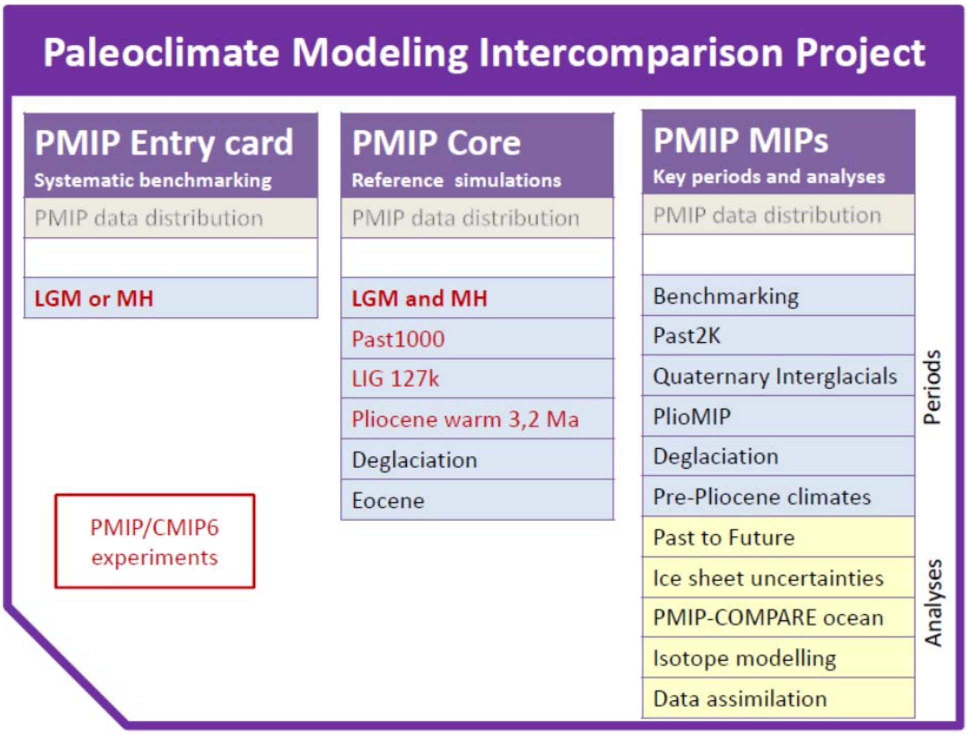
- (A) climate modeling community
    - Improved assessment of forced response and forced vs. internal variability
    - Improved knowledge on which processes are important in the forced response to natural forcing (e.g. ozone changes owing to solar radiation changes for the past1000 experiment)
- (B) Integrated Assessment Modelling (IAM) community,
- (C) Impacts Adaptation and Vulnerability (IAV) community
    - All experiments:
    - Identification of thresholds for ecosystems and water availability under different climate conditions
    - Improved assessment of natural variability including extreme events under pre-industrial boundary conditions. Identification of regions where, under natural forcing, changes, changes lead to specific vulnerability (e.g. regional sea-level)
- , and (D) policy makers.
    - Quantification of magnitude and speed of a range of past climatic changes compared to the natural variability and recent and future climatic trends.Impact of these changes on water availability and ecosystems.

Prioritization of the proposed experiments

- If possible, a prioritization of the suggested experiments, including any rationale**

Each proposed PMIP experiment for CMIP6 can be run independently, because they focus on different time periods. The *midHolocene* and *lgm* experiments have been the focus of PMIP since its start and allow for an evaluation of new model versions since the first atmosphere-only GCMs in PMIP1. We therefore require one of these two simulations to be performed as an entry card to PMIP-CMIP6 experiments. All five PMIP experiments proposed for CMIP6 have equal priority, each experiment being the core of a set of sensitivity experiments to be run within PMIP.

The organization of the PMIP experiments w.r.t CMIP6 is given in the figure below. The complementary experiments will be designed by each PMIP working group (http://pmip3.lsce.ipsl.fr/groups.shtml) and modelling groups are invited to join the discussions on these experiments in this framework.



Model output

- All model output archived by CMIP6-Endorsed MIPs is expected to be made available under the same terms as CMIP output. Most modeling groups currently release their CMIP data for unrestricted use. If you object to open access to the output from your experiments, please explain the rationale.**

PMIP (all experiments): no objections

- List of output and process diagnostics for the CMIP DECK/CMIP6 data request**
    - whether the variable should be collected for all CMIP6 experiments, or only some specified subset and whether the output is needed from the entire length of each experiment or some shorter period or periods;

PMIP (all experiments): same set of CMOR variables as historical/scenario (but reduced set of high-frequency output given the length of experiment), some simulations with COSP simulator (subset of years)

- whether the output might only be relevant if certain components or diagnostic tools are used interactively (e.g. interactive carbon cycle or atmospheric chemistry, or only if the COSP simulator has been installed);

PMIP all-experiments: diagnostics needed for ESM (i.e all components of the ESM + forcings + feedback analyses) + tracers and isotopes (lists to be discussed, first draft sent along with variable lists)

- whether this variable is of interest to downstream users (such as impacts researchers, WG2 users) or whether its principal purpose is for understanding and analysis of the climate system itself. Be as specific as possible in identifying why the variable is needed.

PMIP all periods: subset of variables for driving regional climate models, ice-sheet models (ISMIP) or ecological models (land surface variables) or dust models.

PMIP *past1000* and *midHolocene*: subset of variables for investigating extreme events or variability

- whether the variables can be regridded to a common grid, or whether there is essential information that would be compromised by doing this;

PMIP: same as for CMOR variables from historical/scenario;

- the relative importance of the various variables requested (indicated by a tiered listing) is required if the data request is large.

The list of required variables and their priority has been sent separately.

See also previous PMIP requests (CMIP5 or PMIP3 ESGF): same set of CMOR variables as historical/scenario (possibly reduced set of high-frequency output owing to length of experiment)

- Any proposed changes from CMIP5 in NetCDF metadata (controlled vocabularies), file names, and data archive (ESGF) search terms.**

Needs to discussed with all MIPs

- Explanation of any proposed changes (relative to CMIP5) that will be required in CF, CMOR, and/or ESGF.**

PMIP benefits from two entries on ESGF: via the CMIP5 Project for PMIP3-CMIP5 experiments or the PMIP3 project for other PMIP3 experiments or for groups which do not take part in CMIP5. It would be very convenient to still be able to search through both (or indeed multi-MIP) data bases on the same system, as can be done now.

Proposed contributions for model diagnostics and evaluation

- Any proposed contributions and recommendations for**
    - model diagnostics and performance metrics for model evaluation;

past1000: use diagnostics that have been defined for DECK historical/scenario simulations. In addition to integrated quantities such as hemispheric temperature averages, *past1000* experiments will increasingly be analysed w.r.t. circulation regimes, extreme events etc.

midHolocene, lgm: PMIP specific diagnostics have been developed for benchmarking. A working group is dedicated to this topic. cf. Harrison et al (2014), Climate Dynamics, 43, 671–688

- observations/reanalysis data products that could be used to evaluate the proposed experiments. Indicate whether these are available in the obs4MIPs/ana4MIPs database or if there are plans to include them;

past1000 simulations will benefit from observations to be extended to the early 19th century. *past1000* simulations will be compared, mutually analysed with paleo reconstructions, most importantly the growing set of PAGES2K reconstructions that are available through Paleodata data bases.

PMIP data syntheses for *midHolocene* and *lgm* (http://pmip3.lsce.ipsl.fr/synth/)

new syntheses will be available for characterizing high resolution variability during the Holocene (paleoVar PMIP working group)

- tools, code or scripts for model benchmarking and evaluation in open source languages (e.g., python, NCL, R).

Common analyses scripts are being discussed within PMIP.

For *past1000*: in the framework of the PMIP working group Past2K advanced statistical analyses and evaluation tools have been developed (e.g. Bothe et al., 2013; Moberg et al., 2014).

Expression of interest from modelling groups

On Nov 28th 2014, PMIP has received the expression of interest from 9 modelling groups for the *lastInterglacial* experiment, 10 modelling groups for the *midHolocene*, *lgm* and *past1000* experiments and 11 modelling groups for the Pliocene Warm Period experiment.

Planned publications

The PMIP-CMIP6 protocols will be published in the CMIP6 and PMIP special issues to be opened in GMD. We also plan “first results papers” co-authoring contacts of the participating modelling groups, either in GMD or in Climate of the Past.

Appendix: message sent to all PMIP participants in June 2015

Dear all,

The PMIP proposal for CMIP6 has been endorsed by the CMIP panel. You can find the PMIP/CMIP6 document on the PMIP website (https://wiki.lsce.ipsl.fr/pmip3/doku.php/pmip3:cmip6:design:index for the general presentation and https://wiki.lsce.ipsl.fr/pmip3/lib/exe/fetch.php/pmip3:cmip6:design:cmip6-endorsedmips_pmip.pdf for the proposal itself).

By doing this we committed:
-to a paper on the experimental design of PMIP in the CMIP6 special issue of GMD,
-to report our results by co-authoring a paper jointly with the modelling groups who contribute to PMIP.

This email is to provide some information on PMIP plans for the year. We have both to fulfill the CMIP requirements and to make sure PMIP activity as a whole is well advertised and understood by the community.

First of all, it is important that PMIP participants contribute to the definition of the boundary conditions for the different experiments. Please post your comments on the PMIP wiki (https://wiki.lsce.ipsl.fr/pmip3/doku.php/pmip3:cmip6:design:index). You will find dedicated pages for the different experiments. We still have some cross checking to do between experiments and to fix some of the details.

- PMIP contribution to CMIP6 GMD special issue

A GMD paper is planned (led by Masa Kageyama, who coordinated the PMIP response to CMIP6). This manuscript will describe the PMIP/CMIP6 experiments in terms of objectives and protocol. The deadline for submission of this paper is December 2015.

A second paper (coordinating author not identified yet) will document how each modeling groups has implemented the protocol and performed the experiments, and present preliminary results of the simulations. This paper will involve all the contributing modeling groups, in order to give credit to the modeling groups who have elected to run the palaeo-simulations. We plan to submit this paper by mid-2016, and will be contacting people for information and outputs.

- PMIP GMD special issue

In parallel to the CMIP6 focus, we will open a PMIP GMD special issue as a vehicle for publishing manuscripts presenting each PMIP period and corresponding experiments (PMIP MIPs). We anticipate that these PMIP description papers will be finalized by the end of 2016. For each PMIP MIP (PMIP period) the manuscript will discuss the objectives of the experiments and the boundary conditions, describe the reference simulations, discuss major uncertainties requiring sensitivity experiments, and outline any other coordinated or sensitivity experiments. The plan is for each manuscript to be led by the coordinators of the MIP; all participants in the MIPs will be invited to contribute to the manuscript. The contributors to the definition of the boundary conditions will also be included as co-authors.

PMIP has working groups of different types, some organized as MIP and others more oriented towards analyses. Analytical working groups are also invited to commit to the GMD special issue highlighting the key questions they would like to address and the methodological questions. A wiki page has now been opened for each of the working groups and PMIP participants are invited to contribute and/or lead some of the analyses.

Data syntheses are required for data-model comparisons and model evaluation, both for the CMIP6/PMIP simulations and for the other PMIP simulations. The “benchmarking” working group will identify what major data syntheses are being undertaken and what additional efforts are required. We plan to produce a paper for the GMD CMIP6 special issue (coordinated by Sandy Harrison and Patrick Bartlein) that will document the availability of data for the five CMIP6/PMIP time periods and discuss communalities and differences between the data sets in the treatment of age models, quality control and uncertainties. Everybody who is working on data syntheses for these time periods is invited to contribute to this paper, with the first step being to identify data contributions visa the “benchmarking” wiki page. As a second step, there will be a similar paper (coordinator to be identified) on data syntheses for the wider set of PMIP experiments designed to go into the PMIP special issue.

So, there is quite a lot of work in front of us. I hope it will foster lots of interactions within PMIP.

Pascale, for PMIP coordination

Pascale Braconnot <pascale.braconnot@lsce.ipsl.fr>

IPSL/LSCE, unité mixte CEA-CNRS-UVSQ

- Name of MIP*

Radiative Forcing Model Intercomparison Project (RFMIP)

- Co-chairs of MIP (including email-addresses)*

Robert Pincus, University of Colorado, US; Robert.Pincus@colorado.edu
Piers Forster, University of Leeds, UK; P.M.Forster@leeds.ac.uk
Bjorn Stevens, Max Planck Institue for Meteorology, Germany; Bjorn.Stevens@mpimet.mpg.de

- Members of the Scientific Steering Committee*

Viktor Brovkin, Max Planck Institute for Meteorology, Germany (representing LUMIP)
Gunnar Myhre, CICERO, Norway (representing AerChemMIP)
Hideo Shiogama, National Institute for Environmental Studies, Japan (representing DAMIP)
Karl Taylor, Program for Climate Model Diagnosis and Intercomparison, US (general expertise)

Jean-Louis Dufresne, LMD/IPSL, France (representing IPSL)
James Manners, UK Met Office, UK (representing UKMO)
Miho Sekiguchi, Tokyo University of Marine Science and Technology, Japan (representing MIROC)

- Link to website (if available)*

http://www.wcrp-climate.org/modelling-wgcm-mip-catalogue/modelling-wgcm-mips/418-wgcm-rfmip, though this is out of date

- Goal of the MIP and a brief overview*

RFMIP aims to understand the radiative forcing to which models are subject. The project will assess the accuracy of instantaneous radiative forcing calculations for greenhouse gases and aerosols in each model by comparing these to reference calculations across a range of states representative of present-day, past, and future climates. We will increase the accuracy and spatial detail with which effective radiative forcing is known for each model and for each DECK or other experiment by requesting and analyzing matching simulations designed for this purpose, carefully diagnosing the degree to which the diversity in effective radiative forcing is due to variations in rapid adjustments, instantaneous radiative forcing and climatological base state. We will close the circle by requesting historical-to-near-future simulations in which anthropogenic aerosol optical and cloud-active properties are tightly controlled, allowing us to determine which aspects of the observed historical record consistently emerge and so can be attributed to aerosol forcing.

The project is aligned with CMIP6 criteria as follows:

The MIP addresses at least one of the key science questions of CMIP6

One of the guiding questions for CMIP6 centers on models' response to forcing; it is not possible to answer this question without the ability to quantify the forcing precisely. RFMIP is central to the question of response to forcing and relevant for the other two CMIP questions. Accurate diagnosis of forcing and its errors in CMIP models is key to understanding the spread of models response across simulations and, for example, understanding any possible model bias. Interpreting projections also requires an understanding of the forcings to which models are subject and understanding the degree to which the treatment of radiative transfer introduces errors is central to evaluating a key source of model biases.

The MIP follows CMIP standards in terms of experimental design, data format and documentation

Model integration requests for calculations of effective radiative forcing follow traditional CMIP protocols.

Integrations with prescribed aerosol properties may require small code changes to existing models but will largely follow a protocol that has been already implemented and tested by a number of modeling centers through the Easy Aerosol project (http://www.wcrp-climate.org/index.php/gc-clouds-circulation-activities/gc4-clouds-initiatives/368-gc-clouds-inititative3-easy-aerosol).

The data request for the assessment of aerosol optical properties is unusual in requesting spectrally-detailed information at a few snapshots, but otherwise is a simple request for diagnostic information.

Requests for offline radiative transfer calculations will build on the "site" infrastructure. Reference line-by-line calculations will be performed by RFMIP, formatted to comply with CMIP conventions, and made available to participating modeling centres on the ESG.

A sufficient number of modeling groups have agreed to participate in the MIP

These requests were developed through preliminary email discussions and a dedicated workshop (Hamburg, Sept, 2014). Input has been solicited from climate modeling groups (HadGEM, GFDL, CCCM, CanESM, ECHAM, IPSL, NorESM, MIROC), AeroCom/AerChemMIP, and the radiative forcing community. The second two phases of RFMIP have been developed in consultation with the EC-Earth and CNRM communities. There have been several rounds of interactions with the CMIP panel.

Eleven modeling groups have agreed perform all Tier 1 RFMIP calculations.

The MIP builds on the shared CMIP DECK experiments

The project will precisely quantify forcing in 4xCO_2 and historical integrations, It may lead to decreased spread in forcing by uncovering errors in radiative transfer and will, by design, narrow the spread in forcing for one set of experiments.

A commitment to contribute to the creation of the CMIP6 data request and to analyze the data

Data requests have been provided and refined. The US DoE has funded two of the coordinators (Pincus, Forster) to perform part of the analyses. The same proposal will support reference radiative transfer calculations (Mlawer, Collins, Ramaswamy). Several groups have committed to the analyses (CICERO/UiO, Leeds, UKMO, GFDL, Colorado, Berkley, PCMDI, MPI, LMD/CNRS, GISS).

A commitment to identify observations needed for model evaluation and improved process understanding, and to contribute directly or indirectly to making such datasets available as part of obs4MIPs

Reference calculations by line-by-line models are the radiative forcing equivalent of observations because they have been extensively tested against well-calibrated, spectrally detailed observations. These reference calculations will be made available on the ESG so models can run ongoing tests on their radiative transfer.

Observations of clear-sky flux will constrain radiative forcing for the prescribed aerosol runs of RFMIP and are already distributed through the ESG.

The proposed experiment is of central importance to CMIP6

Without quantifying forcings the community will not be able to understand the spread of model responses.

The proposed experiment has been run at least by two modeling groups already

Fixed SST simulations were performed by nine modeling groups in CMIP5.

The Easy Aerosol methodology has been tested by eight modeling centers (MPI, UKMO, CAM, NCAR, GFDL, IPSL, CICERO).

A pilot assessment of radiative parameterization accuracy using offline calculations under 4xCO_2 conditions attracted 6 GCM codes (representing roughly 16 modeling centers) and 2 reference codes. RFMIP differs from this prototype only in scale.

The proposed experiment is useful in a multi-model context and to a number of climate researchers.

Our focus is on understanding the multi-model spread in response, especially the degree to which this is due to spread in forcing as opposed to spread in sensitivity. This is by definition essential for establishing origins of systematic biases. We expect a very wide community to take full advantage of the data we request and the reference calculations we make. We have support from other proposed MIPs (e.g. AerChemMIP, LUMIP, PDRMIP, VolMIP, ScenarioMIP).

A commitment to scientifically analyze, evaluate and exploit the proposed experiment.

The PIs and their groups are very active publishers of CMIP and related model results. We expect the results of the project to make important contributions to the refereed literature.

- References (if available)*

Shindell, D. T., and Coauthors, 2013: Radiative forcing in the ACCMIP historical and future climate simulations. *Atmos. Chem. Phys*, **13**, 2939–2974, doi:10.5194/acp-13-2939-2013.

Forster, P. M., T. Andrews, P. Good, J. M. Gregory, L. S. Jackson, and M. Zelinka, 2013: Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. *J. Geophys. Res*, **118**, 1139–1150, doi:10.1002/jgrd.50174.

Stevens, B, 2015: Rethinking the lower bound on aerosol radiative forcing. *J. Climate*, doi: 10.1175/JCLI-D-14-00656.1

Pincus, R., and coauthors, 2015: Radiative flux and forcing parameterization error in aerosol-free clear skies, *Geophys. Res. Lett.*. doi:10.1002/2015GL064291.

- An overview of the proposed experiments*

RFMIP-IRF-AER: We will request *detailed diagnostic information* on the vertically- and spectrally-resolved optical properties of aerosols and the surface, along with detailed information about atmospheric physical and chemical state, for snapshots at present-day and pre-industrial conditions. These will be paired with requests for model computations of clear-sky instantaneous radiative forcing. These will allow the diversity in relationships between burden and optical properties to be quantified and will provide data required for accuracy assessments.

RFMIP-IRF-GHG: We will request calculations of vertically-resolved broadband-integrated longwave and shortwave fluxes made with *off-line radiative transfer models* identical to the model's online version, using specified atmospheric states (distribution of temperature and humidity) and surface properties over many profiles. A series of such calculations will assess the accuracy of radiative transfer approximations for gases under various conditions.

RFMIP-ERF: We will request a series of 30-year *uncoupled (atmosphere+land only) simulations with prescribed sea-surface temperatures* and sea-ice concentrations from the preindustrial AOGCM control simulation of the model. Each simulation is matched to an existing simulation from the DECK or some other MIP and will enable the accurate diagnosis of effective radiative forcing for this experiment. Direct radiative forcing and rapid adjustments can be analyzed using newly developed kernel methodologies. Longer runs with fixed sea-surface temperature, paired with the proposed DAMIP AOGCM historical and RCP8.5 simulations, will allow us to calculate transient radiative forcings.

RFMIP-Historical: We will request *small ensembles of "historical+scenario" coupled runs* from 1850 to 2020 in which spectrally-dependent anthropogenic aerosol optical properties and cloud interactions are directly prescribed. Three types of simulations are being requested: historical simulations with all forcings, historical simulations with only natural forcings (here the DAMIP simulations will be used) and, at lower priority, historical simulations with only non-GHG forcing. The time-dependent spatially-patterned prescription will be based on a hybrid model-observation climatology. Such an approach has already been developed and implemented within the MPI-ESM. A non-stationarity factor will account for day-to-day variations. These simulations will be used to assess which aspects of the historical period emerge robustly from ensembles in which aerosol direct and indirect effects are constrained.

- An overview of the proposed evaluation/analysis of the CMIP DECK and CMIP6 experiments*

We will perform no direct analysis of DECK or other experiments in isolation.

The assessment of radiative parameterization accuracy (RFMIP-GHG, RFMIP-AER) is applicable to most DECK experiments, but especially the AMIP, coupled historical, and 4xCO_2 experiments.

Thirty year fixed SST/sea-ice simulations (RFMIP-ERF) will allow us to determine accurate regional variations of effective radiative forcing , rapid adjustments and direct radiative forcing for the DECK's coupled historical, and 4xCO_2 experiments, and for simulations requested by the Detection and Attribution MIP (anthropogenic, well-mixed greenhouse gases, aerosols+ozone, natural) and Land Use MIP (land use).

Simulations with prescribed anthropogenic aerosol optical properties (RFMIP-Historical) will provide a useful complement to historical coupled simulations present in the DECK and will complement activities under the Detection and Attribution MIP.   They will be used to evaluate the hypothesis that a present day aerosol forcing stronger than -1 W/m^2 is incompatible with the temperature record prior to 1950.

- Proposed timing*

Requests for simulations required to diagnose effective radiative forcing, including specification of the diagnostic fields are required, are available now.

Detailed requests (CMOR tables) for aerosol diagnostic outputs are being finalized.

Atmospheric specifications for off-line radiation calculations will be available by the end of 2015.

Prescriptions of aerosol optical properties to be used in specified-aerosol simulations will be available by the end of 2015.

- For each proposed experiment to be included in CMIP6**
  - the experimental design;
  - the science question and/or gap being addressed with this experiment;
  - possible synergies with other MIPs;
  - potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, and (D) policy makers.

- If possible, a prioritization of the suggested experiments, including any rationale**

Aerosol diagnostic requests for RFMIP-IRF-AER

These requests will allow for complete characterization of the diversity in aerosol properties that contribute to the diversity in aerosol direct radiative forcing. They will also provide the basis for the sampling of aerosol optical properties needed to assess radiative transfer errors in the project's next stage.

TIER 1

These requests are unusual in that they require three-dimensional fields at a few instants, and that the spectral detail is requested.

Snapshots of aerosol optical properties (optical thickness, single-scattering albedo, asymmetry parameter) on the model grid (including the spectral grid), along with surface conditions (spectral albedo, temperature), top-of-atmosphere insolation, and fields of temperature, pressure, and water vapor content or relative humidity.

1. Present-day (2015)
2. Pre-industrial (1850)

Offline radiative transfer calculations for RFMIP-IRF-GHG

Offline radiative transfer calculations are intended to identify the degree to which errors in radiative transfer contribute to diversity in estimates of instantaneous radiative forcing, and how this error depends on the background state of the atmosphere and on the forcing itself.

These calculations require an investment in infrastructure (the creation or adaptation of an off-line radiative transfer code traceable to the on-line version) but are a very small computational burden.

TIER 1

These calculations provide baseline error estimates at present-day (PD, 2015) and pre-industrial (PI, 1850) conditions, for the forcing used to assess equilibrium climate sensitivity (4xCO_2) and under conditions used to assess rapid adjustments (PD+4K). We request the following clear sky calculations for a set of prescribed atmospheric conditions.

1. PD greenhouse gas concentrations with PD atmospheres (aerosol-free)
2. PI greenhouse gas concentrations with PD atmospheres (aerosol-free)

3. 4 x PI CO_2, other gases at PD concentrations, in PD atmospheres (aerosol-free)
4. PD greenhouse gas concentrations in PD+4K conditions, assuming constant relative humidity and using a vertical shift transform to map surface warming to atmospheric structure (aerosol-free)
5. “Future” combining simultaneous increases in CO_2, temperature, and humidity (aerosol free)

TIER 2a

These calculations explore the accuracy of CO_2 forcing across a range of concentrations relevant to past values (relevant for Last Glacial Maximum calculations for PMIP) and future concentrations. All use aerosol-free clear-sky PD atmospheric conditions with CO_2 concentrations drawn from {0.25, 0.5, 2, 3, 8} times the PI value.

TIER 2b

These calculations explore the error in radiative forcing estimates from different well-mixed greenhouse gases. Calculations are for aerosol-free clear-sky conditions using PD atmospheres and greenhouse gas concentrations with one (or all) set to its PI value:

1. CH_4
2. N_2O
3. CO
4. HC
5. O3
6. All

Model integrations to diagnose effective radiative forcing for RFMIP-ERF

Model integration requests are designed to give consistent radiative forcings in both emission-based and prescribed-concentration frameworks and we recommend that CMIP6 standardizes to this methodology for computing radiative forcings across all MIPs and that other MIPs link to RFMIP as far as possible. RFMIP concentrates on understanding forcing changes from preindustrial (PI, 1850) to present day (PD, 2015).

The protocol for time-slice experiments is a single 30-year uncoupled simulation in which sea surface temperature (SST) and sea ice distributions are specified and vegetation may interact. Sea ice and anthropogenic forcing agents are specified at PI values unless noted. Tests of this method demonstrate that 30-year integrations constrain broad regional patterns of forcing to better than 0.1 Wm^{-2}.

TIER 1

These integrations will allow us to quantify the radiative forcing at present day. We request the following integrations:

a) 30 year PI control with monthly averaged fixed SST and sea-ice climatology from the PI AOGCM control integration. To be used as control for all other integrations
b) As a) with $4xCO_2$
c) As a) with PD ANTHROPOGENIC forcers
d) As a) with PD WMGHGs (+indirect effects in emission based models)
e) As a) with PD AEROSOLS and OZONE (linking to additional forcing estimates from AerChemMIP)

f) As a) with PD LAND USE (surface albedo/roughness, transpiration). Change vegetation but fix GHG concentrations and preindustrial levels. (linking to LUMIP)

TIER2a

Non-linearity of aerosol cloud interactions are particularly important for understanding historical forcing evolution (Carslaw et al., 2013). These two experiments are designed to quantify this across the models.

g) As a) with PD AEROSOLS and OZONE x 0.1
h) As a) with PD AEROSOLS and OZONE x 2.0

TIER2b

Forcings at periods other than present are important for understanding aspects of historical and future change. Different applications will likely require different time-slices. It is also difficult to evaluate transitory volcanic and solar forcing using a time-slice methodology of Tier 1. Therefore we propose to evaluate forcings over the full 1850-2100 period, concentrating on natural forcings not evaluated in Tier 1.

Transient forcings from CMIP5 models relied on a crude two-step residual method to diagnose globally averaged ERF estimates (Forster and Taylor, 2006; Forster et al., 2013). This method assumed that a CO_2-based climate sensitivity was applicable across scenarios and the method had significant errors due to noise. The two-step method is a possible fall-back for CMIP6.

The two-step method can be improved on by using transient climate model integrations with fixed SST and sea ice (PD ctrl). Preliminary work with HadGEM2 indicates that three ensemble members with this method would be needed to constrain interannual variation on global forcing to better than 0.1 Wm^{-2}. Further work is needed to verify the number of accuracy of different ensemble sizes.

We will request small ensembles of 1850-2100 fixed SST and sea ice climate integrations (including interactive vegetation). These requests are matches of DAMIP AOGCM requests to have pairs of consistent forcing and response integrations, maximizing the science benefit.

i) Historical + RCP8.5 (all forcing integration) with forcings added to PI control from a)
j) Natural (solar and volcanoes) forcings added to a), only to 2015
k) Aerosols (and their indirect effects) added to a),
l) WMGHG changes added to a)

TIER 3

Transient WMGHG forcing can likely be inferred from tier 1 results. However, a transient AOGCM WMGHG run is proposed by DAMIP and it is useful to check its forcing in a few models.

Model integrations with specified aerosol properties for RFMIP-Historical

Here we request integrations in which aerosol optical properties including cloud-radiation interactions are prescribed. The prescription is expected to be a single analytically-described spatial pattern with time-constant spectral variation (building on the Easy Aerosols experience) and time-varying strength.

Our goal is to determine which aspects of the historical record robustly emerge from ensembles in which the radiative forcing by anthropogenic aerosols is tightly constrained.

One of the central aims of RFMIP it to ensure that all models use a common prescription of aerosol optical properties. In the past there has been the idea that one could use a concentration based aerosol prescription of aerosol properties derived, for instance, from the ensemble average of emission based aerosol models. This strategy has several drawbacks that have led to confusing inconsistencies in previous iterations of CMIP:

(1) AerChemMIP has an ambitious experimental design so that climatological forcing from the multi-model mean will only become available very late and there will not be a chance to thoroughly evaluate it.

(2) The AerChemMIP models have a very heterogeneous description of the aerosol so a multi-model mean implies different averages over different sub-ensembles. This is unsound because of correlations in varying aerosol components in the different sub-ensembles (i.e., models with nitrate usually have smaller sulfate forcing because these species are linked).

(3) The data volumes for a concentration-based are large. Past experience has shown that interpolating them to the different model grids is error prone.

(4) The forcing fields derived from the emission-based models are inconsistent with present understanding of the aerosol distribution. To put this in context it would be like forcing the historical simulations with the multi-model mean of the CO2 concentrations from the emission based models rather than tying the CO2 concentrations to our understanding of the present and past record.

For these reasons we see AerChemMIP as focusing on the important problem of our understanding of how emissions effect concentrations and the radiative forcing associated with the concentrations predicted by different models but that, for simulations designed to explore the historical record and to satisfy DAMIP, the CMIP panel should recommend that modeling centers use the RFMIP prescription of aerosols. We are happy to work with the emissions community and AerChemMIP to extend the aerosol climatology proposed for use in CMIP to the future, which would be useful for models performing the Scenario runs or near term projections.

We are in close contact with HighResMIP; they are very interested in using the RFMIP prescription of the aerosol particularly in light of points (1) and (3).

TIER 1

A 4-member ensemble of coupled simulations from 1850-2020 (merging the historical period and RCP 8.5 or an appropriate scenario for GHGs for the 2015-2020 period) with all forcings including specified aerosols.

TIER 2

A set of 4-ensembles to be used in detection and attribution experiments

1. Hist-Nat, 1850-2020 (joint with DAMIP)
2. Hist-Aer 1850-2020 (with prescribed aerosol)
3. Corresponding prescribed SST ERF runs (see part III) to diagnose aerosol forcing

TIER 3

1. AMIP, 1980-2020 (with prescribed aerosol)
2. Single AMIP, 1980-2020, nudged to observed winds (with prescribed aerosol)

- All model output archived by CMIP6-Endorsed MIPs is expected to be made available under the same terms as CMIP output. Most modeling groups currently release their CMIP data for unrestricted use. If you object to open access to the output from your experiments, please explain the rationale.**
- List of output and process diagnostics for the CMIP DECK/CMIP6 data request**
  - whether the variable should be collected for all CMIP6 experiments, or only some specified subset and whether the output is needed from the entire length of each experiment or some shorter period or periods;
  - whether the output might only be relevant if certain components or diagnostic tools are used interactively (e.g. interactive carbon cycle or atmospheric chemistry, or only if the COSP simulator has been installed);
  - whether this variable is of interest to downstream users (such as impacts researchers, WG2 users) or whether its principal purpose is for understanding and analysis of the climate system itself. Be as specific as possible in identifying why the variable is needed.
  - whether the variables can be regridded to a common grid, or whether there is essential information that would be compromised by doing this;
  - the relative importance of the various variables requested (indicated by a tiered listing) is required if the data request is large.
- Any proposed contributions and recommendations for**
  - model diagnostics and performance metrics for model evaluation;
  - observations/reanalysis data products that could be used to evaluate the proposed experiments. Indicate whether these are available in the obs4MIPs/ana4MIPs database or if there are plans to include them;
  - tools, code or scripts for model benchmarking and evaluation in open source languages (e.g., python, NCL, R).
- Any proposed changes from CMIP5 in NetCDF metadata (controlled vocabularies), file names, and data archive (ESGF) search terms.**

  Aerosol diagnostics will require the addition of a model-dependent spectral dimension and the addition of roughly ten variables describing the spectrally-dependent characteristics of aerosols, surface properties, and top-of-atmosphere solar insolation. These are described in the data request.

- Explanation of any proposed changes (relative to CMIP5) that will be required in CF, CMOR, and/or ESGF.**

None as far as we are aware.

Scenario Model Intercomparison Project (ScenarioMIP)
Application for CMIP6-Endorsed MIPs

Date: 18 June 2015

Proposals from MIPs should include the following information:

* *preliminary information that will be used to determine whether a MIP should be endorsed for CMIP6 or not.*

** *information that must be provided later (and before the panel can determine which experiments, if any, will be incorporated in the official CMIP6 suite).*

- Name of MIP*

ScenarioMIP

- Co-chairs of MIP (including email-addresses)*

Brian O'Neill (boneill@ucar.edu), Claudia Tebaldi (tebaldi@ucar.edu), Detlef van Vuuren (detlef.vanvuuren@pbl.nl)

- Members of the Scientific Steering Committee*

Veronika Eyring (DLR, Germany), Pierre Friedlingstein (U of Exeter, UK); George Hurtt (U of Maryland, USA); Reto Knutti (ETH, Switzerland); Jean-Francois Lamarque (NCAR, USA); Jason Lowe (MetOffice, UK); Jerry Meehl (NCAR, USA); Richard Moss (Joint Global Change Research Institute, USA); Ben Sanderson (NCAR, USA)

- Link to website (if available)*

https://www2.cgd.ucar.edu/research/mips/scenario-mip

- Goal of the MIP and a brief overview*

Overall objectives

The goal of ScenarioMIP is to simulate future climate outcomes based on alternative plausible[1] future scenarios in order to:

(1) Facilitate integrated research leading to a better understanding not only of the physical climate system consequences of these scenarios, but also of the climate impact on societies, including considerations of mitigation and adaptation. The results of the ScenarioMIP experiments will provide new climate information for plausible future scenarios that will facilitate integrated research across multiple communities including the (1) climate science, (2) integrated assessment modeling (IAM) and mitigation, and (3) impacts, adaptation and vulnerability (IAV) communities.

(2) Provide a basis for addressing targeted science questions regarding the climate effects of particular aspects of forcing relevant to scenario-based research, e.g., the effect of different assumptions on land use and near-term climate forcers (NTCFs), such as tropospheric aerosols, ozone and methane, on climate change and its impacts.

---

[1] We define plausible in a narrow sense as outcomes feasible to achieve in a model and published in the peer reviewed literature as scenarios intended to be considered as plausible futures (as opposed to sensitivity analyses, thought experiments, or other characterizations of the future).

(3) Provide a basis for various international efforts that target improved methods to quantify projection uncertainties based on multi-model ensembles, taking into account model performance, model dependence and observational uncertainty (this extends the knowledge basis derived from the DECK experiments and the CMIP6 Historical Simulation and allows for the quantification of uncertainties on different timescales).

The first objective on providing “scenarios for integration” across the disciplines involved in climate research is considered to be the highest priority for the following reasons:

- Scenarios for integration serve a large scientific audience, underpinning hundreds of scenario-based studies addressing a wide variety of scientific questions regarding physical climate changes, mitigation, impacts, and adaptation. Having common climate and socioeconomic scenarios serves as a critical means to enhance direct comparability of a wide variety of studies, allowing synthetic conclusions to be drawn that would not be possible from a variety of uncoordinated studies.
- Climate simulations based on such broad-use scenarios are critical elements of the new scenario process established at Noordwijkerhout in 2007 (Moss et al., 2007; 2010); without climate simulations to support integrated studies that draw on both climate and societal futures, the scenario process cannot function. CMIP5 simulations will continue to underpin this process through 2020, and CMIP6 scenarios are seen as a critical continuation of that contribution.
- Scenarios for integration serve as a key means for connecting the assessments of different working groups of IPCC as well as in the Synthesis Report.
- A common set of scenarios for integration reduces the need for individual research projects to develop their own scenario information to support scenario-based studies. The availability of common scenarios reduces possible redundancy in efforts and makes scenario-based research feasible for many groups that otherwise would not be able to carry it out.

Because targeted questions regarding the climate effects of individual forcings and forcing pathways are also very important to scenario-based research, a set of variants of the scenarios proposed here are being proposed in other MIPs (see below) to address these targeted questions. Thus, the scenarios in ScenarioMIP not only serve the function of integration across research communities, but also serve as anchoring scenarios from which variants are designed to address targeted questions in AerChemMIP, C4MIP, DAMIP, GeoMIP, ISMIP6, LUMIP, and RFMIP. In addition, some of the scenarios proposed by ScenarioMIP will be used as a basis for short-term predictions by DCPP MIP, for regional downscaling in CORDEX, and for provision of climate information by ClimateServicesMIP. This document discusses how the experiments in the other MIPs relate to the design of ScenarioMIP.

Background

A ScenarioMIP Scientific Steering Committee was formed following the October 2013 WGCM17 meeting as an outcome of earlier discussions among the IAM, IAV and climate modeling communities at the annual meeting of the integrated assessment and impacts communities in Snowmass, CO, in July 2013, and the AGCI session on CMIP6 in Aspen, CO, in August 2013. The Scientific Steering Committee together with other communities (see below) systematically investigated a number of issues that could substantially influence the experimental design, including the possibility of statistically sampling climate model-scenario combinations, the potential for pattern scaling or other statistical emulators of climate model output to meet some of the demand for scenario-based climate information,[2] and the differences between scenarios (in terms of global average forcing or temperature change) that is required to produce climate outcomes that are significantly different at the grid-cell level.

---

[2] A three-day workshop on pattern scaling was organized by ScenarioMIP co-chairs and others in April 2014 to address this question, see https://www2.image.ucar.edu/event/PS2014.

Conclusions of these investigations were that a sparse statistical sampling approach to design was unworkable, that pattern scaling has not yet been demonstrated to be able to reliably replace the need for climate model simulations to generate information for impact studies (although it might play a limited role for some applications), and that scenario differences of at least 0.3 C in global average temperature are likely necessary to generate statistically significant differences in local temperature and precipitation outcomes over a substantial fraction of the surface.

Informed by these conclusions, a process was organized by the ScenarioMIP Scientific Steering Committee to develop this final protocol. This process included close interaction with the climate research, IAM and IAV communities through follow-up meetings at Snowmass and Aspen in summer 2014, discussions with representatives of the IAMC Working Group on Scenarios producing candidate scenarios for CMIP6, and discussions with key individuals in other relevant research communities, including through the International Committee On New Integrated Climate change assessment Scenarios (ICONICS) and the WCRP-IPCC WG1 meeting in Bern, Switzerland, in September 2014. That proposal was submitted to and discussed at the October 2014 WGCM18 meeting; revised to reflect feedback at that meeting, additional coordination with other MIP proposals, and feedback from a presentation of the proposal at the annual meeting of the Integrated Assessment Modeling Consortium (IAMC) in November; and resubmitted to the CMIP panel in early December 2014. That version of the proposal described a largely complete experimental design, but left open a number of details of the design for which final decisions required additional information from the IAV and IAM communities (particularly regarding the specific outcomes of IAM-based emissions and land use scenarios) as well as additional coordination with other MIPs. Since that time we have interacted closely with the IAM and IAV communities, jointly discussed experimental designs with several other MIPs, and also received feedback on the draft proposal from the CMIP review process as well as from our own circulation of the proposal to relevant groups including ICONICS, TGICA, and the WCRP Working Group on Regional Climate (WGRC). This final version of the proposal includes a number of changes to the previous submission based on this set of interactions and feedback, as well as on feedback from an IPCC expert meeting on Scenarios in May 2015 and from review comments provided by the CMIP panel.

Moving forward, and contingent on formal approval of ScenarioMIP by the CMIP Panel, the ScenarioMIP SSC plans to write and submit a paper before the end of 2015 on the ScenarioMIP experimental design as part of the CMIP6 special issue in Geoscientific Model Development. We also plan to co-author a paper with all participating modeling groups on the results of climate model simulations when they become available and to move forward research in the area of the scientific questions listed below.

- References (if available)*

- An overview of the proposed experiments*

Scientific questions

The scientific questions addressed by the experiments proposed in ScenarioMIP fall under two of the broad questions of interest to CMIP6, as follows:

1. **How does the Earth system respond to forcing?**
    - How does the Earth system respond to forcing pathways relevant to IAM and IAV research and to policy considerations?
    - What is the uncertainty in global and regional climate change due to *plausible* variations in future **land use** and **NTCFs** emissions, and how does it compare to multi-model uncertainty in the response to a given forcing pathway?
    - How much do plausible alternative shapes of forcing pathways (e.g. **overshoot**) matter to climate change outcomes, and therefore to questions about mitigation, impacts, and adaptation?

    - What is the uncertainty in global and regional climate as a result of model uncertainty (as opposed to scenario variations), and how can this be estimated from a model ensemble of opportunity without a specific design to sample uncertainty?
    - Can emergent constraints (i.e., statistical relationships between features of current and projected future climate that emerge from considering the multi-model ensemble as a whole) be used to recalibrate the ensemble and to reduce the uncertainty in the response to a given scenario of future forcing?
    - In which part of the Earth System, and when, are such constraints expected to emerge, how do they trace back to modelled processes, are those processes adequately represented, and how can this information be used to improve models, point to critical observations and monitoring programs, and link process understanding, detection and attribution, projections, and uncertainty quantification?

2. **How can we assess future climate changes given climate variability, climate predictability, and uncertainties in scenarios?**
    - How can we assess future climate changes for forcing pathways spanning a range of uncertainties in global and regional forcing relevant to IAM and IAV research, as well as to policy?

In addition, the ScenarioMIP experiments address the CMIP6 themes based on the WCRP grand challenges by addressing the following question:

> How will plausible future forcing pathways affect **climate extremes, global and regional climate information, regional sea level rise, water availability, and biospheric feedbacks**, and how will these effects influence mitigation and adaptation possibilities?

Overview of proposed design

The experimental design (see Table 1 for an overview) consists of seven 21st-century scenarios grouped into three tiers by priority. In addition, within the second tier, we propose an overshoot scenario in which a peak in radiative forcing occurs in the 21st century, additional ensemble members for one of the 21st century scenarios, and three long-term extensions that begin from the end points of two of the 21st century scenarios and extend to 2300.

Each 21st century scenario achieves a specific level of global average forcing by the end of the century. Choices of the global average forcing level for these scenarios were based on the ScenarioMIP objectives outlined above. These objectives imply that the global average forcing pathways should cover a wide range of plausible forcing levels, provide continuity with CMIP5 experiments, and fill in gaps in CMIP5 forcing pathways that would be of interest to the climate science, IAM, and IAV communities.

Based on these considerations, two types of 21st century scenarios were included in the design:

(1) “SSP-based RCPs”: new versions of the RCPs that are based on the Shared Socioeconomic Pathways (SSPs; O’Neill et al., 2014; van Vuuren et al., 2014) and new IAM model simulations derived from them. The SSPs are five new societal development pathways that have been developed as part of the parallel process and provide descriptions of future societal conditions that serve as the basis both for deriving forcing pathways and for characterizing vulnerability and mitigative capacity important for IAV/IAM studies.

(2) “Gap scenarios”: new forcing pathways not covered by the RCPs, including new unmitigated SSP baselines or new mitigation pathways.

The SSPs are named SSP1-SSP5. While they represent only one subset of wide range of possible socio-economic trajectories in the future, they were designed to span a range of low and high challenges for mitigation and adaptation. With one exception (RCP8.5 which is achievable only by assuming SSP5) all global average forcing pathways for the 21st century scenarios can be achieved by assuming different SSPs. The dominant hypothesis is that differences in climate outcomes produced by different SSPs for the same global forcing pathway are likely small relative to regional climate variability, uncertainty across climate models, and uncertainty in impact models used to investigate outcomes of interest to the IAV community. Therefore, climate simulations based on a forcing pathway produced with one SSP will be used in studies aimed at investigating the effects of that same global average forcing pathway but under future socioeconomic conditions given by a different SSP.

However, the degree to which this hypothesis is correct remains an open scientific question. We therefore choose an SSP for each global average forcing pathway by taking into consideration the possibility that the sensitivity of climate outcomes to SSP choice may be larger than anticipated. To account for that possibility, choices were based on one or more of the following goals:

(1) *facilitate climate research* to learn more about the climate effects of aspects of forcing that may vary by SSP for the same global average forcing pathway, particularly those from land-use changes and aerosol emissions.

(2) *minimize differences in climate* between the outcomes produced by the SSP chosen for a given global average forcing pathway and the climate that would have been produced by choosing other SSPs. These differences would be minimized by choosing an SSP with land use and aerosol pathways that are central relative to other SSPs for the same global average forcing pathway. However, given difficulties in identifying a central scenario (due for example to consideration of multiple variables and regions), in practice this goal implies avoiding SSPs with trends for land use or aerosols that are outliers relative to other SSPs.

(3) *ensure consistency with scenarios that are most relevant to the IAM/IAV community*. Not all scenarios for a given global average forcing pathway are anticipated to be equally relevant to IAM and IAV research. This goal implies choosing the SSP that we anticipate to be especially relevant, so that if the climate effects of land use and aerosols turn out to be larger than anticipated, climate simulations will still be consistent with that scenario.

Forcing category	Type of Scenario	Forcing in 2100[1] (W/m2)	SSP	Short name	Use by other MIPs[2]
Tier 1[3]					
High	SSP-based RCP	8.5	5	SSP5-8.5	C^4MIP, GeoMIP, ISMIP6, RFMIP
Medium-high	Gap: Baseline	7.0	3	SSP3-7	AerChemMIP, LUMIP
Medium	SSP-based RCP	4.5	2	SSP2-4.5	VIAAB, CORDEX, GeoMIP, DAMIP, DCPP
Low	SSP-based RCP	2.6	1	SSP1-2.6	LUMIP
Tier 2					
Additional 21st century scenarios					
Medium[4]	SSP-based RCP	6.0	1	SSP1-6.0	GeoMIP
Low	Gap: Mitigation	3.7	4	SSP4-3.7	
Overshoot scenario					
Overshoot[5]	Gap: Mitigation	2.6	X	SSPx-2.6 over	
Ensembles[6]					
SSP3-7.0, 9-member ensemble	Gap: Baseline	7.0	3	SSP3-7.0	AerChemMIP, LUMIP
Extensions					
SSP5-8.5, long-term extension	SSP-based RCP	8.5	5	SSP5-8.5 ext	C^4MIP, ISMIP6, GeoMIP
SSP5-8.5, long-term – overshoot	SSP-based RCP	8.5	5	SSP5-8.5 ext-over	C^4MIP, ISMIP6, GeoMIP
SSP1-2.6, long-term extension	SSP-based RCP	2.6	1	SSP1-2.6 ext	
Tier 3					
Additional 21st century scenarios					
Low[7]	Gap: Mitigation	<2.6	X	SSPX-Y	

Notes

1 Forcing levels are nominal identifiers. Actual forcing levels of the SSPs depend, for non-climate policy scenarios, on socio-economic developments while for scenarios that include climate policy, the objective was to replicate forcing in the RCPs run as part of CMIP5. These values differed somewhat from the nominal levels.

2 Current plans by other MIPs to use ScenarioMIP scenarios either directly or as a basis for a variant to be run as part of their own design are indicated here.

3 We strongly recommend that modeling groups participating in ScenarioMIP run at least the four scenarios in Tier 1, and as many additional scenarios as possible, guided by this prioritization. However, for any group running fewer than four scenarios, SSP5-8.5 should be considered the highest priority.

4 Due to uncertainty in the forcing level that would be achieved in the SSP1 baseline, the SSP choice for this scenario remains to be confirmed when final IAM scenario results are available in August 2015.

5 The details of the overshoot scenario remain to be defined, pending IAM scenarios to be carried out over the next several months. This scenario will have a forcing overshoot within the 21st century relative to the SSP1-2.6 scenario, but the specific SSP and the timing of the overshoot remain to be defined.

6 We request that models run 9 or more additional initial condition ensemble members for the SSP3-7.0 scenario (if not 9, then as many as possible).

7 This scenario is conditional on the development of such a scenario in the context of the SSPs by the IAM modelling teams involved in this exercise, and therefore details remain to be defined. See discussion below.

Overall, the proposal detailed in Table 1 has the following general features:

- A small number of scenarios (4 in Tier 1, with only one run per scenario being requested from each model) required for any model participating in this MIP, with model runs of additional scientific value in Tiers 2 and 3.
- Tier 1 spans a wide range of uncertainty in future forcing pathways important for research in climate science, IAM, and IAV studies, while also providing key scenarios to anchor experiments in a number of other MIPs (see last column in Table 1).
- Tier 2 builds on Tier 1 scenarios by further filling in the range of forcing pathways with additional scenarios of interest and also includes additional ensemble members and long-term extensions.
- Tier 3 would provide a scenario lower than the RCP 2.6 forcing pathway intended to inform policy interest in a global average temperature limit below 1.5 °C warming above pre-industrial levels.
- The new versions of the scenarios with similar forcing as the RCPs will continue to support scenario-based IAM and IAV research into the mid-2020s. These scenarios will be based not only on new (CMIP6) climate models, but also on updated forcing pathways generated by new IAM model runs based on the SSPs.
- The four new “gap” scenarios will explore forcing pathways beyond those of the existing RCPs: a scenario that is representative of the forcing level of many baseline scenarios (i.e., no mitigation) in the literature (around 7 W/m2), scenarios that addresses policy discussions of mitigation pathways that fall between RCPs 2.6 and 4.5 (3.7 W/m2) and below RCP 2.6, and an overshoot pathway that explores the climate science and policy implications of peak in forcing during the 21st century.
- Scenarios that can anchor experiments in a number of other MIPs (see below) to investigate targeted questions, including for example the influence of land use, aerosols and other NTCFs, and overshoot on climate outcomes; carbon cycle feedbacks; and ice sheet-climate interactions.

We list here more specific descriptions and justifications for each of the scenarios in the design, as well as for some over-arching features of the design. For each of the 21st century scenarios, we describe the relevance of the forcing pathway and also the rationale for the choice of the SSP.

Tier 1: 21st century scenarios

- **SSP5-8.5:** This scenario represents the high end of the range of plausible future pathways, updates the RCP8.5 pathway in CMIP5, and is planned to be used by a number of other MIPs to help address their scientific questions. SSP5 was chosen for this forcing pathway because it is the only SSP with emissions high enough to produce this level of forcing.
- **SSP3-7.0:** This scenario represents the medium to high end of the range of plausible future forcing pathways. It fills a gap in CMIP5 forcing pathways that is particularly important because it represents a forcing level common to several (unmitigated) SSP baselines. These baseline scenarios will be very important to IAV studies interested in quantifying “avoided impacts,” which requires comparing impacts in a mitigation scenario with those occurring in an unmitigated baseline scenario. SSP3 was chosen because SSP3-7.0 is a scenario with both substantial land use change (in particular decreased global forest cover) and high NTCF emissions (particularly SO_2) and therefore will play an important role in LUMIP and AerChemMIP, addressing scenario-relevant questions about the sensitivity of regional climate

to land use and aerosols. In addition, SSP3 (combined with this forcing pathway) is especially relevant to IAM/IAV studies because it combines relatively high societal vulnerability (SSP3) with relatively high forcing. This scenario would also be used as the basis for our requested large ensemble (discussed below).

- **SSP2-4.5:** This scenario represents the medium part of the range of plausible future forcing pathways and updates the RCP4.5 pathway in CMIP5. It will be used by CORDEX MIP (along with SSP5-8.5) for regional downscaling, a product that will be very valuable to the IAV community, by DCPP MIP for short-term predictions out to 2030, and by DAMIP as a continuation of the historical simulations to update regression-based estimates of the role of single forcings beyond 2015 and to run single forcing experiments into the future by using it as the reference scenario. SSP2 was chosen because its land use and aerosol pathways are not extreme relative to other SSPs (and therefore appear as central for the concerns of DAMIP and DCPP MIP), and also because it is relevant to IAM/IAV research as a scenario that combines intermediate societal vulnerability with an intermediate forcing level.
- **SSP1-2.6:** This scenario represents the low end of the range of plausible future forcing pathways and updates the RCP2.6 pathway in CMIP5. SSP1 was chosen because it has substantial land use change (in particular increased global forest cover) and will be used by LUMIP to help address their scientific questions. From the IAM/IAV perspective this scenario is highly relevant since it combines low vulnerability with low challenges for mitigation as well as a low forcing signal. The scenario depicts thus a “best case” future from the sustainability perspective.

Tier 2: 21st century scenarios

- **SSP1-6.0:** This scenario fills in the range of medium forcing pathways and updates the RCP6.0 pathway in CMIP5. SSP1 was chosen because it is relevant to IAM/IAV research as a scenario that defines the low end of the forcing range for unmitigated SSP baseline scenarios and because together with SSP1-2.6 it could be used to investigate differences in impacts across global average forcing pathways even if the regional climate effects of land use and aerosols turn out to be strong.
- **SSP4-3.7:** This scenario fills a gap at the low end of the range of plausible future forcing pathways. There is substantial mitigation policy interest in scenarios that reach 3.7 W/m2 by 2100, since mitigation costs differ substantially between forcing levels of 4.5 W/m2 and 2.6 W/m2 (depicted by the RCPs). Climate model simulations would allow for impacts of a 3.7 scenario to be compared to those occurring in the 4.5 or 2.6 scenarios, to evaluate relative costs and benefits of these scenarios. SSP4 was chosen because it is relevant to IAM/IAV research as a scenario with relatively low challenges to mitigation (SSP4) and therefore is a plausible pairing with a relatively low forcing pathway.

Tier 2: 21st century overshoot

This scenario fills a gap in existing climate simulations by investigating the implications of a substantial 21st century overshoot in radiative forcing relative to a longer-term target. There is substantial interest in the impact, mitigation and adaptation implications of such overshoot, which begins with understanding the climate consequences of such a pathway. Existing IAM scenarios within this category produce an overshoot of 1-1.5 W/m2 relative to the level in 2100, so we anticipate a scenario with approximately this degree of overshoot. The overshoot in forcing will occur relative to the Tier 1 SSP1-2.6 scenario, and possibly to its extension.

Tier 2: Large ensemble

It is important for scenario-based research to represent the influence of internal variability on climate outcomes. To accommodate this need, while also economizing on model runs, we request that models run multiple initial condition ensemble members only for one scenario, based on the assumption that variability estimated for one scenario can be applied to outcomes for others. We are requesting this ensemble be carried out for the SSP3-7.0 scenario (a Tier 1 scenario) for two reasons:

- The relatively high forcing level reached by this scenario by the end of the 21st century will enable the exploration of potential changes in internal variability over a substantial range of global average radiative forcing and temperature change, which could not be assessed if the large ensemble was run for a lower scenario, e.g. RCP4.5. Understanding potential changes in variability over a wide range of forcing levels is essential to support the possibility of transferring variability under the large ensemble to other scenarios for which we request only a single ensemble member.
- SSP3-7.0 has relatively strong land use change and high emissions of NTCFs (unlike the SSP5-8.5 scenario), and therefore has been identified as an important experiment on which variants will be conducted by LUMIP and AerChemMIP to investigate the climate implications of regional differences in land use and aerosol emissions. This topic is also very important to scenario-based studies. In those MIPs, the opportunity to conduct signal-to-noise studies made possible by multiple initial condition ensemble members will be critical.

We request that models run 9 additional ensemble members (if not 9, then as many as possible). These additional ensemble members would be considered Tier 2 scenarios (i.e., not required model runs for participation in ScenarioMIP). For all other scenarios, only a single ensemble member is requested.

Tier 2: Long-term extensions

There is strong interest from the climate and impacts communities in long-term extensions of scenarios beyond 2100. The ScenarioMIP long-term extensions will consist of three experiments.

- Two of these will provide low and high cases for long-term change, comprising extensions for SSP5-8.5 and SSP1-2.6 in a style similar to the extensions of these scenarios in CMIP5. For SSP5-8.5, this involves emissions that are eventually reduced to a level that is found to produce equilibrated radiative forcing at a relatively high level by 2300 in a simple climate model. For SSP1-2.6 this involves an extension of the negative carbon emissions reached in 2100, leading to slowly declining forcing.
- To complement these cases, an additional experiment will consider the implications of rapid decarbonization from SSP5-8.5 beginning in 2100, and will linearly reduce all emissions to SSP1-2.6 levels by 2200, an experiment which is expected to produce a significant overshoot of radiative forcing by 2300.

Tier 3: 21st century scenarios

- **Scenario below RCP2.6:** This scenario represents the very low end of the range of scenarios in the literature measured by their radiative forcing pathway. Plausible IAM scenarios more than incrementally below RCP2.6 in terms of radiative forcing are rare and have only recently become available in the peer reviewed literature. There is policy interest in scenarios that would inform a possible goal of limiting global mean warming to 1.5°C above pre-industrial levels. We include such a scenario in Tier 3 because such a scenario in the context of SSPs has not yet been developed by the IAM community and because the SSP1-2.6 scenario and its long-term extension can in part also inform analyses of the implications of 1.5 °C warming. Its inclusion in the ScenarioMIP design is conditional on the IAM community, as represented by the Integrated Assessment Modeling Consortium (IAMC) Scientific Working Group on

Scenarios, producing such a scenario. It is likely that such a scenario would become available for climate model simulations at some delay relative to scenarios in Tiers 1 and 2.

Emissions- vs concentration-driven simulations

We recommend that the scenarios specified in the ScenarioMIP design be run as concentration-driven experiments. Such scenarios are more consistent with the "integration" role that these scenarios will play in the broader research community. The conceptual framework for scenario-based research is based on investigating the implications of alternative climate futures. These climate futures will be more similar (for a given scenario) in concentration-driven runs than in emissions-driven runs (given uncertainties in the carbon cycle), and therefore will better serve this purpose of the overall scenario framework. Concentration driven scenarios still represent uncertainty in the carbon cycle and in climate-carbon cycle feedbacks, through their influence on the anthropogenic carbon emissions allowable for a given concentration pathway. Indeed, ESM results indicating the uncertainty across models in allowable emissions will be a very important outcome for the IAM community. We recognize that concentration-driven scenarios do not allow for assessing amplification effects of feedbacks (in which climate change influences the carbon cycle, producing more emissions and more climate change, and further influencing the carbon cycle, etc.). However, amplification could be investigated in other C^{4}MIP simulations.

CMIP5 vs CMIP6 models

CMIP6 climate projections will differ from those for CMIP5 due to both a new generation of climate models as well as a new set of scenarios of emissions and land use. For multiple research communities it will be useful to evaluate the difference in climate outcomes that is due to of the changes in climate models alone. For example, such an evaluation is valuable in order to determine whether CMIP5 and CMIP6 results could be used together in research on impacts and adaptation (and how), or whether IAM and IAV researchers should abandon CMIP5 runs in favor of CMIP6 runs when they become available. It is not part of the ScenarioMIP design being proposed to CMIP6 to carry out simulations that would inform this evaluation. However, we believe it would be interesting to the community if at least a few climate modeling teams investigated this question. Possible approaches include running the new SSP-based scenarios with the previous (CMIP5) generation of models, running the previous (RCP) scenarios using the new (CMIP6) generation of models, or carrying out relevant analyses with climate model emulators.

Connections to other MIPs

The ScenarioMIP design is intended to provide a basis for targeted scenarios to be run in other MIPs in order to address specific questions regarding the sensitivity of climate change outcomes to particular aspects of these scenarios, especially land use and emissions of NTCFs. We describe here current plans for coordinated experiments. A summary of the scenarios within the ScenarioMIP design that are currently part of plans for other MIPs is provided in the experimental design table above.

Aerosols and Chemistry MIP (AerChemMIP)

AerChemMIP has a Tier 1 experiment (with additional Tier 2 and 3 related studies) directed at the sensitivity of climate to near term climate forcers. This experiment will use the SSP3-7.0 scenario from ScenarioMIP as a starting point and devise a lower air pollutant variant of this scenario by assuming pollution controls, or maximum feasible reductions in air pollutants. In addition, AerChemMIP will make use of the LUMIP land-use variant on SSP3-7.0 (with land use from SSP1-2.6) to study couplings between land-use changes and atmospheric chemistry.

Coupled Climate Carbon Cycle MIP (C^{4}MIP)

ScenarioMIP will coordinate with C^{4}MIP on targeted scenarios regarding concentration vs emission driven simulations. While the ScenarioMIP protocol will recommend concentration-driven simulations (see above), C^4MIP/Tier 1 will recommend emission-driven simulations for the SSP5-8.5 in order to explore the implications of carbon cycle feedbacks on projected climate change. As mentioned before, C^4MIP also has an interest in the extensions of scenario beyond 2100 (e.g. up to 2300 as in CMIP5) in order to investigate climate change impacts on Earth System components that operate on longer time scales (vegetation, permafrost, oceanic circulation and carbon export, etc.). C^4MIP has expressed high interest in analyzing the ScenarioMIP overshoot scenario. Overshoot scenarios are also of potential interest to GeoMIP given that geoengineering may be an option for avoiding overshoot.

Detection and Attribution MIP (DAMIP)

DAMIP plans to use SSP2-4.5 as an anchoring scenario on the basis of which individual forcing simulations extended to the end of the century will be specified and then compared. These experiments are aimed at distinguishing the climate effects of different forcers and facilitating the identification of observational constraints and their use in future projections. SSP2-4.5 will also be used to extend the historical (all forcing) runs to 2020 for use in regression-based estimates of the role of individual forcings within the observational constraint provided by observational records up to the years beyond 2015 (by the time CMIP6 output will be available and the next IPCC assessment report will be written).

Decadal Predictions MIP (DCPP MIP)

DCPP MIP plans to use SSP2-4.5 forcings for its initialized short-term predictions out to 2030, and SSP2-4.5 runs as comparison to evaluate the prediction skills of those predictions.

Geoengineering MIP (GeoMIP)

GeoMIP has proposed several experiments that will use two scenarios from ScenarioMIP as a basis from which geoengineering measures would be implemented. Forcing pathways from other ScenarioMIP scenarios would serve as targets for those measures. In particular, SSP5-8.5 would be used as a basis for four experiments: using geoengineering to reduce forcing to a medium forcing (G6Sulfur and G6Solar experiments) or low forcing (G6Sulfur_SSP1-2.6) Tier 1 scenario, investigating the effect of cirrus cloud thinning (G7Cirrus experiment), and investigating the effect of fixed levels or stratospheric aerosol injections (GeoFixed10, 20, 50). For the G6Sulfur and G6Solar experiments will also be extended beyond 2100, with geoengineering applied to reduce forcing from the overshoot extension of SSP5-8.5 down to the forcing level of the medium forcing Tier 1 scenario. In addition, the medium forcing Tier 1 scenario would be used as a basis for a stratospheric aerosol injection experiment (G4SSA).

Ice Sheet MIP (ISMIP6)

ISMIP will be proposing two types of experiments that will draw on long-term extensions of a scenario from ScenarioMIP in order to investigate ice sheet response and ice-climate interactions on centennial timescales. In particular, an extension of SSP5-8.5 to 2300 would be used to provide climate model output for offline (uncoupled) ice sheet simulations, and to provide emissions/concentrations for fully coupled ice sheet-climate model experiments.

Land Use MIP (LUMIP)

LUMIP plans to design experiments that use two scenarios from ScenarioMIP as a basis for testing sensitivity to land use change. These two scenarios would differ both in forcing levels and in land use change. These two scenarios will be the SSP3-7.0 and the SSP1-2.6 scenarios. These two scenarios span a range of approximately 4.5 W/m2 (7.0 vs 2.6 W/m2 in 2100), and likely will differ substantially in land use change, with substantial deforestation in the SSP3-7.0 and net afforestation in SSP1-2.6.

Radiative Forcing MIP (RFMIP)

RFMIP has plans to estimate radiative forcing in different models for a plausible future scenario, preferably a high forcing pathway. At the moment the candidate is SSP5-8.5, whose forcings would be applied to current day fixed SSTs in the idealized setting of the RFMIP experiments.

Vulnerability, Impacts, Adaptation (VIA) Advisory Group

Researchers examining the consequences of climate change and potential adaptations are a key user group of CMIP outputs and products. ScenarioMIP will establish a close link with the impact community through the VIA Advisory Board and other relevant groups to facilitate integrated research that leads to a better understanding not only of the physical consequences of these scenarios on the climate system, but also of the climate impact on societies. In particular ScenarioMIP will link with the VIA Advisory Board to ensure that the climate model output from the scenarios allows for sector-specific indices being derived (e.g., heat damage degree days for ecosystems, consecutive dry days for agriculture and water resources).

- An overview of the proposed evaluation/analysis of the CMIP DECK and CMIP6 experiments*

- Proposed timing*

The following timeline for ScenarioMIP was developed at the August 2014 Aspen meeting on MIPs for scenarios, land use, and aerosols:

31 March 2015	Submission of final ScenarioMIP proposal to CMIP panel
Autumn 2015	Submit paper on ScenarioMIP design to the CMIP6 Special Issue
April 2015 – October 2016	Specification of future emissions and land use scenarios from IAMs, harmonization with historical emissions/land use, specification of future atmospheric concentrations
October 2016	Provision of IAM scenario information to ESMs; ScenarioMIP ESM runs begin
2018-2020	Co-author a paper with all participating modeling groups on outcomes of ScenarioMIP simulations

We note that the October 2016 date for completion of harmonized emissions and land use scenarios, as well as atmospheric concentrations, will need to be revisited when these activities begin and the scope of work can be better assessed. In particular, the timeline for a scenario below RCP2.6 may differ significantly from the plan for other scenarios.

- For each proposed experiment to be included in CMIP6**
    - the experimental design,
    - the science question and/or gap being addressed with this experiment,
    - possible synergies with other MIPs,
    - potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, and (D) policy makers.

See accompanying worksheet.

- If possible, a prioritization of the suggested experiments, including any rationale**

See above and accompanying worksheet.

- List of output and process diagnostics for the CMIP DECK/CMIP6 data request**

  - Please indicate whether the standard output archived from the *CMIP DECK* experiments needs to be complemented by additional diagnostics to be useful for the MIP or whether the output is only suggested for the additional experiments,
  - Some output might only be relevant if certain components or diagnostic tools are used interactively (e.g. carbon cycle, chemistry, simulators); please indicate clearly if this is the case,
  - If the data request is large, please indicate the importance of the various data to be archived via a tiered listing.

ScenarioMIP is not making a specific data request, but participating model groups should consult the data requests of other MIPs using scenarios (see Table 1), and also the recommended set of variables that will be requested for the DECK and the CMIP6 Historical Simulation. We believe that the ScenarioMIP data request is best represented by the union of the requests of other MIPs that will be explicitly drawing on ScenarioMIP simulations in their experimental designs. In addition, the VIA Advisory Board has indicated that they will produce a data request relevant to the impacts, adaptation, and vulnerability community. That data request, when available, will also apply to ScenarioMIP simulations. We do not think the process would be well served by generating, in addition, our own data request separate from these other sources.

- Proposed contributions and recommendations for**
  - model diagnostics and performance metrics for model evaluation,
  - observations/reanalysis that could be used to evaluate the proposed experiments. Status in obs4MIPs/ana4MIPs (in database / to be included).
  - when possible tools, code or scripts for model evaluation in open source languages (e.g., python, NCL, R).
- Any proposed changes from CMIP5 in NetCDF metadata (controlled vocabularies), file names, and data archive (ESGF) search terms.**
- Explanation of any proposed changes (relative to CMIP5) that will be required in CF, CMOR, and/or ESGF.**

References

Meehl, G.A., Moss, R., Taylor, K.E., Eyring, V., Stouffer, R.J, Bony, S. and B. Stevens (2014) Climate Model Intercomparisons: Preparing for the Next Phase. EOS Transactions of the American Geophysical Union 95 (9): 77-78.

Moss, R. H. et al. (2008) Towards New Scenarios for Analysis of Emissions, Climate Change, Impacts, and Response Strategies (IPCC Expert Meeting Report, IPCC, Geneva) .

Moss, R. H. et al. (2010). The next generation of scenarios for climate change research and assessment. Nature 463: 747-756. doi:10.1038/nature08823

O’Neill, B., Kriegler, E., Riahi, K., Ebi, K.L., Hallegatte, S., Carter, T.R., Mathur, R. and D.P. van Vuuren (2014) A new scenario framework for climate change research: the concept of shared socioeconomic pathways. Climatic Change, Special Issue, Nakicenovic N, Lempert R, Janetos A (eds) A Framework for the Development of New Socioeconomic Scenarios for Climate Change Research. DOI 10.1007/s10584-013-0906-1

van Vuuren DP, Kriegler E, O’Neill BC, Ebi KL, Riahi K, Carter TR, Edmonds J, Hallegatte S, Kram T, Mathur R, Winkler H (2014) A new scenario framework for Climate Change Research: scenario matrix architecture. Climatic Change, Special Issue, Nakicenovic N, Lempert R, Janetos A (eds) A Framework for the Development of New Socioeconomic Scenarios for Climate Change Research. DOI 10.1007/s10584-013-0906-1

Application for a Model Intercomparison Project on the climatic response to Volcanic forcing (VolMIP) as CMIP6-Endorsed MIP

Davide Zanchettin[1*], Claudia Timmreck[2], Myriam Khodri[3], Alan Robock[4], Gabi Hegerl[5], Anja Schmidt[6], Matthew Toohey[7], Francesco S.R. Pausata[8], Benjamin Black[9], Oliver Bothe[10], Jason M. English[11], Edwin Gerber[12], Hans F. Graf[13], Allegra N. LeGrande[14], Graham Mann[6], Timothy Osborn[15], Steven J. Phipps[16], Christoph C. Raible[17], Angelo Rubino[1], Björn Stevens[2], Didier Swingedouw[18], Kostas Tsigaridis[14,19], Qiong Zhang[8]

1	University of Venice, Italy
2	Max-Planck-Institute for Meteorology, Hamburg, Germany
3	IRD/IPSL/Laboratoire d'Océanographie et du Climat, France
4	Department of Environmental Sciences, Rutgers University, New Brunswick, USA
5	GeoScience, U. Edinburgh, UK
6	School of Earth and Environment, University of Leeds, UK
7	GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany
8	Department of Meteorology (MISU), Stockholm, Sweden
9	University of California , Berkley, U.S.A
10	HZG, Helmholtz Center Geesthacht, Geesthacht, Germany
11	Laboratory for Atmospheric and Space Physics University of Colorado, Boulder, USA
12	Courant Institute of Mathematical Sciences, New York University
13	University of Cambridge, UK
14	NASA GISS, Columbia University, USA
15	Climatic Research Unit, School of Environmental Sciences, University of East Anglia, UK
16	University of New South Wales, Sydney, Australia
17	Universität Bern, Swiss
18	Université de Bordeaux, France
19	Center for Climate Systems Research, Columbia University

* To whom the correspondence should be sent: University of Venice, Dept. of Environmental Sciences, Informatics and Statistics, Calle Larga Santa Marta, Dorsoduro 2137, Venice, Italy (davide.zanchettin@unive.it)

VolMIP

Name of MIP:

Model Intercomparison Project on the climatic response to Volcanic forcing (**VolMIP**)

Co-chairs of MIP (including email addresses):

Davide Zanchettin (davide.zanchettin@unive.it)
Claudia Timmreck (claudia.timmreck@mpimet.mpg.de)
Myriam Khodri (myriam.khodri@locean-ipsl.upmc.fr)

Members of the Scientific Steering Committee:

Gabi Hegerl (gabi.hegerl@ed.ac.uk)
Alan Robock (robock@envsci.rutgers.edu)
Anja Schmidt (A.Schmidt@leeds.ac.uk)
Matt Toohey (mtoohey@geomar.de)
Edwin Gerber (gerber@cims.nyu.edu)

Link to website (if available):

Official webpage:
http://www.volmip.org/
WCRP webpage:
http://www.wcrp-climate.org/index.php/modelling-wgcm-mip-catalogue/modelling-wgcm-mips/505-modelling-wgcm-volmip

Goal of the MIP and a brief overview

VolMIP is central to the three broad CMIP questions:

- How does the Earth system respond to external forcing?
- What are the origins and consequences of systematic model biases?
- How can we assess future climate changes given climate variability, predictability and uncertainties in scenarios?

VolMIP is motivated by the large uncertainties regarding the climatic responses to strong volcanic eruptions identified in CMIP5 simulations with respect to, e.g., the radiative forcing during periods of strong volcanic activity (e.g., Santer et al., 2014; Marotzke and Forster, 2015), the Northern Hemisphere's winter response (e.g., Driscoll et al., 2012, Charlton-Perez et al., 2013), the precipitation response (Iles and Hegerl, 2014) and the response of the oceanic thermohaline circulation (Ding et al., 2014), and by the apparent mismatch between simulated and reconstructed post-eruption surface cooling for volcanic eruptions during the last millennium (Mann et al., 2012, 2013; Anchukaitis et al., 2012; D'Arrigo et al., 2013; Schurer et al., 2013). Inter-model differences are likely related to differences in the prescribed volcanic aerosol forcing data used by different models, or variations in implementation, which create differences in the radiative forcing produced by the volcanic aerosol forcing. The use by some modeling groups of coupled aerosol modules for the CMIP6 historical experiments, with volcanic forcing thereby explicitly simulated based on estimates of SO_2 emissions (Gettelman et al., pers. comm., 2015), will increase inter-model spread in volcanic forcing.

Therefore, VolMIP fills the need for a coordinated model intercomparison with volcanic forcing – in terms of aerosols optical properties – constrained across participating models. Specifically, VolMIP will assess to what extent responses of the coupled ocean-atmosphere system to the same applied strong volcanic forcing are robustly simulated across state-of-the-art coupled climate models and identify the causes that limit robust simulated behavior, especially differences in their treatment of physical processes.

VolMIP is closely linked to the WCRP Grand Challenge on:

- “Clouds, circulation and climate sensitivity”, in particular through improved characterization of volcanic forcing and improved understanding of how the hydrological cycle and the large-scale circulation respond to volcanic forcing. VolMIP further contributes to the initiative on leveraging the past record through planned experiments describing the climate response to historical eruptions that are not (or not sufficiently) covered by CMIP6-DECK, -historical or other MIPs. VolMIP will contribute towards more reliable models through improved understanding of how model biases affect the response to a well-defined volcanic forcing.
- “Climate extremes” and “Regional climate information”, in particular through a more systematical assessment of regional climate variability – and associated predictability and prediction - during periods of strong volcanic forcing at both intraseasonal-to-seasonal (e.g., post-eruption Northern Hemisphere's winter warming) and interannual-to-decadal (e.g., post-eruption delayed winter warming) time scales.
- “Water Availability”, in particular through the assessment of how strong volcanic eruptions affect the monsoon systems and the occurrence of extensive and prolonged droughts.
- “CLIC and Cryosphere”, in particular concerning the onset of volcanically forced long-term feedbacks involving the cryosphere which is suggested by recent studies (e.g., Miller et al., 2014, Berdahl and Robock, 2013; Zanchettin et al., 2014). VolMIP encourages the interested modeling community to discuss sensitivity experiments focused on the climatic effects of aspects related to tephra deposition.

VolMIP addresses specific questions related to:

- The apparent mismatch between simulated and reconstructed post-eruption surface cooling for volcanic eruptions during the last millennium (Mann et al., 2012; Anchukaitis et al., 2012; D’Arrigo et al., 2013; Schurer et al., 2013). A possible reason for the mismatch are the large uncertainties in the volcanic forcing for eruptions that occurred during the pre-instrumental period and for which no direct observations are available. Therefore, VolMIP will be based on consensus forcing input data and related coupled climate simulations for some of the major volcanic eruptions that occurred during the pre-industrial period of the last millennium. Forcing data will be in the form of best estimates with uncertainties or of a range of estimates if a best estimate is not feasible with the given uncertainties.
- The mismatch between observed and modeled seasonal to interannual dynamical responses to volcanic eruptions during the instrumental period. Observations suggest that volcanic eruptions are followed by an anomalously strong Northern Hemisphere’s winter polar vortex, and significant positive anomalies in the North Atlantic Oscillation and Northern Annular Mode, but CMIP5 models do not robustly reproduce this behavior (e.g., Driscoll et al., 2012, Charlton-Perez et al., 2013). Observed volcanic events are, however, few and of limited magnitude, and their associated dynamical climate response is very noisy (e.g., Hegerl et al., 2011). The short-term dynamical response is now known to be sensitive to the particular structure of the applied forcing (Toohey et al., 2014). Using carefully constructed forcing fields and sufficiently large simulation ensembles, VolMIP will investigate the inter-model robustness of the short-term dynamical response to volcanic forcing, and elucidate the mechanisms through which volcanic forcing leads to changes in surface dynamics. Such improved understanding will be also beneficial for the predictability of interannual climate response to future eruptions.
- The large uncertainties in the interannual and decadal dynamical climatic responses to strong historical volcanic eruptions. As described above, coupled climate simulations produce a considerable range of atmospheric and oceanic dynamical responses to volcanic forcing, which likely depend on various aspects of model formulation, on the simulated background internal climate variability (e.g., Zanchettin et al., 2013), and also on eruption details including magnitude, latitude and season (e.g., Timmreck, 2012). VolMIP will help to identify the origins and consequences of systematic model biases affecting the dynamical climate response to volcanic forcing and to clarify how regional responses to volcanic forcing are affected by the background climate state, especially the phase of dominant modes of internal climate variability. As a consequence, VolMIP will improve

our confidence in the attribution and dynamical interpretation of reconstructed post-eruption regional features and provide insights into regional climate predictability during periods of strong volcanic forcing.

- The large uncertainties in the multidecadal and longer-term climate repercussions of prolonged periods of strong volcanic activity (e.g., Miller et al., 2012; Schleussner and Feulner, 2013; Zanchettin et al., 2013). VolMIP proposes an experiment describing the climate response to the close succession of strong volcanic eruptions that affected the early 19th century, whose long-term repercussions may be relevant for the initialization of CMIP6 historical simulations.

In summary, VolMIP will contribute towards advancing our understanding of the dominant mechanisms behind simulated post-eruption climate evolution, but also more generally of climate dynamics and of seasonal and decadal climate variability and predictability. Volcanic eruptions offer the opportunity to assess the climate system's dynamical response to changes in radiative forcing, a major uncertainty in future climate projections. Careful sampling of initial climate conditions and the possibility to consider volcanic eruptions of different strengths (e.g., Fröhlicher et al., 2012; Muthers et al., 2014, 2015; Zanchettin et al., 2014) will allow a better understanding of the relative role of internal and externally-forced climate variability during periods of strong volcanic activity, hence improving the evaluation of climate models and enhancing our ability to accurately simulate past, as well as future, climates.

For these purposes, VolMIP defines a common protocol to improve comparability of results across different Earth system models and coupled general circulation models, and accordingly subjects them to the same set of idealized volcanic perturbations – implemented through prescribed aerosols optical parameters - under similar background climate conditions (Zanchettin et al., in prep, 2015).

VolMIP experiments will be designed based on a twofold strategy.

- A first set of experiments is designed to systematically investigate inter-model differences in the long-term (up to the decadal time scale) dynamical climate response to the same idealized volcanic eruptions that are characterized by a high signal-to-noise ratio in the response of global-average surface temperature. The main goal of these experiments is to assess the signal propagation pathways of volcanic perturbations within the simulated climates, the associated determinant processes and their representation across models.
- A second set of experiments will be used to systematically investigate inter-model differences in the short-term dynamical response to the same idealized 1991 Pinatubo-like eruption and discriminate the parts that are due to internal variability and to model characteristics. The proposed set of experiments will include sensitivity experiments designed to determine the different contributions to such uncertainty that are due to the direct radiative (i.e., surface cooling) and to the dynamical (i.e., stratospheric warming) response. A joint experiment with the Decadal Climate Prediction Panel (DCPP) using the same idealized 1991 Pinatubo-like volcanic forcing will address the impact of volcanic forcing on seasonal and decadal climate predictability.

Identification of consensus forcing input data for both types of experiments is an integral part of VolMIP. Some of the participating modeling groups are currently testing the proposed methodologies through coordinated activities within VolMIP and in cooperation with the Stratospheric Sulfur and its Role in Climate Initiative (SSiRC) model intercomparison initiative, the SPARC DynVar activity and DCPP. In addition to the identification of consensus forcing input data in terms of aerosol optical parameters, the VolMIP protocol defines for all the experiments additional constraints about the implementation of the forcing (e.g., spectral interpolation).

An overview of the proposed experiments

An overview of the proposed experiments is provided in Tables 1, 2 and 3, where they are summarized according to their prioritization. VolMIP experiments are divided into two main branches: long-term volcanic forcing experiments and short-term volcanic forcing experiments.

Long-term volcanic forcing experiments

Experiments based on coupled climate simulations to assess inter-model differences in the climate response to *very strong* volcanic eruptions up to the decadal time scale.

- *VolLongS60EQ:* This Tier 1 experiment is designed to realistically reproduce the radiative forcing resulting from the 1815 eruption of Mt. Tambora, Indonesia. The experiment will not account for the actual climate conditions when the real event occurred (e.g., presence and strength of additional forcing factors). Instead, the experiment is designed to span very different initial climate states to systematically assess uncertainties in the post-eruption behavior that are related to background climate conditions.
- *VolLongS100HL:* An additional, non-mandatory experiment, which applies the same approach as *VolLongS60EQ* and extends the investigation to the most relevant historical high-latitude volcanic eruption (1783-1784 Laki, Iceland). The unique eruption style (large SO_2 mass releases: 100 Tg SO_2, and close temporal spacing: 5 active phases within 5 months) will substantially contribute to outstanding questions about the magnitude of the climatic impact of high-latitude eruptions. Due to the long emission period, results of this experiment may have implications for sulfate aerosol geo-engineering.
- *VolLongC19th*: A “volcanic cluster” experiment to investigate the climate response to a close succession of strong volcanic eruptions. The proposed experiment is designed to realistically reproduce the volcanic forcing generated by the early 19th century volcanic cluster (including the 1809 eruption of unknown location and the 1815 Tambora and 1835 Cosigüina eruptions). The early 19th century is the coldest period in the past 500 years (Cole-Dai et al., 2009) and therefore of special interest for multidecadal variability. In addition long-term repercussions may be relevant for the initialization of CMIP6 historical simulations.

Short-term volcanic forcing experiments

Experiments based on coupled climate simulations to assess uncertainty and inter-model differences in the seasonal-to-interannual climatic response to more frequent large volcanic eruptions over the recent observational period. Such eruptions are characterized by smaller magnitude compared to those used for the *VolLong* experiments, hence they are characterized by a rather low signal-to-noise ratio in the response of global-average surface temperature.

- *VolShort20EQfull*: This Tier 1 experiment uses the same volcanic forcing recommended for the 1991 Pinatubo eruption which is used in the CMIP6 historical simulation (Thomason et al., 2015), but produces a large ensemble of short-term simulations in order to accurately estimate simulated responses to volcanic forcing which may be comparable to the amplitude of internal interannual variability.
- *VolShort20EQsurf/strat:* Tier 1 simulations aimed at investigating the mechanism(s) connecting volcanic forcing and short-term climate anomalies. Specifically, these experiments will aim to disentangle dynamical responses to the two primary thermodynamic consequences of aerosol forcing: stratospheric heating and surface cooling.
- *VolShort20EQslab:* Non-mandatory slab-ocean experiment, which is proposed to clarify the role of coupled atmosphere-ocean processes (most prominently linked to the El Niño-Southern Oscillation) in determining the dynamical response.
- *VolShort20EQini:* Non-mandatory experiment to address the impact of volcanic forcing on seasonal and decadal climate predictability. The experiment will address the climate implication of a future Pinatubo-like eruption. The experiment is designed in cooperation with DCPP. It complies with the VolMIP protocol about the forcing and its implementation. VolMIP supports other DCPP decadal prediction experiments using idealized forcing from the 1963 Agung and 1982 El Chichón eruptions.

Experimental set-up:

Length of integration

- *LongS*: for each simulation: at least 20 years (mandatory), but preferably longer (30-40 years) to cover the multi-decadal oceanic response;

- *LongC*: at least 50 years to cover the multi-decadal oceanic response and to assess stationarity of post-cluster climate;
- *Short*: for each simulation: 3 years, since the experiment focuses on the short-term responses;
- *Short.ini:* a minimum of 5 years (up to 10 years) for each initialized run.

Initial conditions:

- *LongS*: predefined states describing different states of dominant modes of variability (see "ensemble size") sampled from an unperturbed control integration, under common constant boundary forcing across the different models (*PiControl* simulations from DECK). The VolMIP experiments should maintain the same constant boundary forcing as the control integration, except for the volcanic forcing;
- *LongC*: as *LongS,* but inclusion of background volcanic forcing and a dedicated spin-up procedure for this experiment are currently under discussion to account for possible implications of volcanic forcing on ocean heat content in long transient simulations (e.g., Gregory, 2010);
- *Short*: predefined states describing different states of dominant modes of variability (see "ensemble size") sampled from an unperturbed control integration, under common constant boundary forcing across the different models (PiControl simulations from CMIP6-DECK). The VolMIP experiments should maintain the same constant boundary forcing as the control integration, except for the volcanic forcing;
- *Short.ini*: initialized on 1st November 2015, or any other date in November or December for which initialized hindcasts are available (depending on the modelling Center).

Ensemble size:

- *LongS*: should be large to systematically account for the range of variability depicted by the dominant processes influencing interannual and decadal climate variability. VolMIP will accordingly identify a set of desired initial conditions. Nine simulations are planned for the Tier 1 experiment, which would allow spanning warm/cold/neutral and strong/weak/neutral states of El Niño-Southern Oscillation (ENSO) and of the Atlantic Meridional Overturning Circulation (AMOC), respectively;
- *LongC*: at least an ensemble of 3 simulations;
- *Short*: same rationale as for *LongS*, but further taking into account additional phenomena primarily contributing to internal atmospheric variability, such as the Quasi Biennial Oscillation (QBO), the characteristics of the polar vortex and the North Atlantic Oscillation (NAO). A core of 25 simulations is requested for the Tier 1 experiment, but a larger ensemble size is recommended;
- *Short.ini:* at least 5-member ensembles, but preferably 10-member ensembles.

Forcing input:

The applied radiative forcing should be consistent across the participating models for all events included in the protocol. Therefore, VolMIP will provide a self-consistent set of forcing parameters in terms of aerosol optical properties (e.g., aerosol optical depth, effective radius, single scattering albedo and asymmetry factor) that can be used by all models. In addition, VolMIP will define for all the experiments constraints about the implementation of the forcing.

- *Long:* These experiments are based on pre-industrial volcanic events for which no direct observation is available. VolMIP will collect candidate forcing sets from proxy-based reconstructions and simulations from coupled climate models including modules for stratospheric chemistry and aerosol microphysics, and aims to select a single, consensus forcing data set for the *Long* simulations. If ad-hoc forcing inputs cannot be generated for an event, VolMIP will indicate reference forcing data sets to be used that are already available to the community.
- *Short*: *VolShort20EQfull* will use the CMIP6 stratospheric aerosol data set (Thomason et al., 2015) for the volcanic forcing of the 1991 Pinatubo eruption which is set up for the CMIP6 historical simulation. The mechanistic experiments will not account for forcing based on imposed aerosol optical properties as the usual approach in VolMIP. Instead, *VolShort20EQsurf* will specify a prescribed perturbation to the shortwave flux to mimic the attenuation of solar radiation by volcanic aerosols, and therefore the cooling

of the surface. Two approaches are currently under discussion: first, the changes could be prescribed at the top of atmosphere under clear sky conditions (variable sw_toaflux_aero_cs of VIRF); alternatively, the changes could be prescribed as restoring of the surface albedo (see below for more details). The goal is to enforce a uniform shortwave perturbation across all models and to isolate the impact of shortwave reflection from the impact of aerosol heating in the stratosphere. Similarly, *VolShort20EQStrat* will specify a prescribed perturbation to the total (LW+SW) radiative heating rates, seeking to mimic the local impact of volcanic aerosol (i.e. zmlw_aero and zmsw_aero of VIRF). This could be implemented by changes within the radiation code, or even outside the code, by adding an additional temperature tendency. Concrete details about the forcing implementation will be prescribed in Zanchettin et al. (2015). This activity will be conducted in close collaboration with SPARC DynVar.

The observation-based volcanic-forcing to be used in the CMIP6 historical and VolMIP *VolShort20EQfull* experiments contains information about the real-world structure of the stratospheric circulation at the time of the eruptions, which does not necessarily match the states of individual free-running model realizations. To further investigate the impact of the forcing structure on the dynamical response, VolMIP will support the development of an idealized volcanic forcing dataset, where the spatial structure of the forcing is much more uniform than observation-based forcings. This work shares parallels with the WCRP Grand Challenge initiative “Easy Aerosol” and RFMIP, and we envision cooperation in the future months between the two groups. Additional dedicated sensitivity experiments will be carried out by individual model centers to contribute to this activity.

Surface albedo changes due to tephra deposition are neglected in all the experiment as well as indirect cloud radiative effects.

An overview of the proposed evaluation/analysis of the CMIP DECK and CMIP6 experiments

VolMIP experiments will provide context to CMIP6-DECK (AMIP) and -historical simulations where volcanic forcing is among the dominant sources of climate variability and inter-model spread. VolMIP will provide essential information for the interpretation of the CMIP6 historical experiments. VolMIP will provide a well-defined set of forcing parameters in terms of aerosol optical properties and is thus complementary to the Stratospheric Sulfur and its Role in Climate (SSiRC) coordinated multi-model initiative, which uses global aerosol models to investigate radiative forcing uncertainties associated to given SO_2 emissions. The importance of VolMIP experiments is enhanced as some climate modelling groups plan to perform the CMIP6 historical simulations with online calculation of volcanic radiative forcing based on SO_2 emissions (Gettelman et al., pers. Com, 2015). VolMIP closely cooperates with SSiRC and the different model groups as well as RFMIP to build the scientific basis to distinguish between differences in volcanic radiative forcing data and differences in the climate model response to volcanic forcing.

Time schedule

2015 31. March	Submission of final VolMIP proposal to CMIP Panel and WIP co-chairs
2015 April 8	VolMIP splinter meeting at Tambora conference in Bern (Switzerland) – discussion of forcing input data for the VolLongS60EQ experiment
2015 April 13	Submission of draft for invited VolMIP contribution to special issue of the PAGES magazine (http://www.pages-igbp.org/products/pages-magazine) focused on volcanoes and climate
2015 April 15	Comment to volcanic forcing data sets in CMIP6 by VolMIP SC
2015 April – December	Submission of draft to GMD (Zanchettin et al., 2015) documenting detailed experimental design
2015 June 7-12	DCPP Aspen workshop (participation of VolMIP co-chair)

2015 June	Invited talks from VolMIP Co-chairs at the 26th IUGG General Assembly in Prague
2015 July	VolMIP talk at "Our Common Future Under Climate Change" conference in Paris, France
2015 October 20-23	VolMIP contribution at the workshop on CMIP5 Model Analysis and Scientific Plans for CMIP6 (EMBRACECMIP2015), Dubrovnik, Croatia
2015 -2016	Work on idealized volcanic forcing fields
2016	Execution of Tier1 experiments
2016	VolMIP workshop for discussion of experiments
2017- 2019	Execution of Tier2 (Tier3) experiments
2017	Public sharing and analysis of model output

Possible synergies with other MIPs:

VolMIP is closely linked to and will co-operate with the following ongoing modeling activities and MIPs:

- **PMIP** (https://pmip3.lsce.ipsl.fr/) – PMIP and VolMIP provide complementary perspectives on one of the most important and less understood factors affecting climate variability during the last millennium. VolMIP systematically assesses uncertainties in the climatic response to volcanic forcing associated with initial conditions and structural model differences. In contrast, the PMIP last-millennium experiments, i.e., the past1000 simulations, describe the climatic response to volcanic forcing in long transient simulations where related uncertainties are due to the reconstruction of past volcanic forcing, the implementation of volcanic forcing within the models, initial conditions, the presence and strength of additional forcings, and structural model differences. VolMIP and PMIP are expected to tighten cooperation in the upcoming months to strengthen the synergies between the two MIPs.
- **GeoMIP** (http://climate.envsci.rutgers.edu/GeoMIP/) – GeoMIP and VolMIP share interest on the climatic effects of stratospheric aerosol loadings. The closest association between proposed experiments is between VolMIP *Long* and GeoMIP G6sulfate simulations.
- **RFMIP** (Radiative Forcing MIP) – Precise quantification of the forcing to which models are subject is central for both RFMIP and VolMIP. RFMIP has encouraged other MIPs to standardize as far as possible to the RFMIP methodology for computing radiative forcings. RFMIP has planned transient volcanic and solar forcing experiments with fixed preindustrial SST to diagnose volcanic and solar effective forcing, instantaneous forcing and adjustments, which is complementary to the *Short* experiments for VolMIP.
- **DAMIP** (Detection and Attribution MIP) – DAMIP and VolMIP share the common interest of assessing the relevance of volcanic forcing over the historical past. In particular, VolMIP can address the substantial uncertainty associated with the effects of volcanism on the historical periods. DAMIP's histALL, histNAT, histVLC and histALL_aerconc can provide context to the *Short* set of VolMIP simulations, since they include the 1991 Pinatubo eruption within transient climate situations.
- **DCPP** (Decadal climate prediction panel) - VolMIP and DCPP are closely working together on the impact of future volcanic eruptions on seasonal and decadal predictions, with a common experiment. The proposed VolMIP's *Short* experiment including 1991 Pinatubo-like volcanic forcing in decadal prediction runs *(Short20EQini)* and the DCPP experiment C3.4 are identical and will be jointly prepared/discussed at the DCPP workshop on June 7-12, 2015, in Aspen, CO (USA).
- **SPARC DYNVAR (**http://www.sparcdynvar.org/) – The SPARC DynVar group aims to assess the impact of uncertainty in atmospheric dynamics on climate projections and to understand the underlying physical processes. DynVar is therefore deeply involved in the setup and analysis of VolMIP's *Short* experiments.

- VolMIP is closely linked to with the ongoing modeling activities within **SPARC-SSiRC** (http://www.sparcssirc.org/). The Stratospheric Sulfur and its Role in Climate Initiative (SSiRC) model intercomparison uses global aerosol models to understand the radiative forcing of stratospheric aerosols (background, volcanic) and to assess related parameter uncertainties.

Potential benefits of the experiment to (A) climate modeling community, (B) Integrated Assessment Modelling (IAM) community, (C) Impacts Adaptation and Vulnerability (IAV) community, and (D) policy makers.

A. VolMIP will contribute towards identifying the causes that limit robust simulated behavior under strong volcanic forcing conditions. Uncertainty in simulated estimates of clear-sky radiative forcing is largest around strong volcanic eruptions, which poses VolMIP at the core of CMIP6. VolMIP will also clarify more general aspects of the dynamical climatic response to strong external forcing, especially differences in the models' treatment of physical processes. VolMIP will further evaluate the possibility of robustly identifying key climate feedbacks in coupled climate simulations following well-observed eruptions (e.g., Soden et al., 2002), and assess the role of model biases for simulations-observations discrepancies.

B. VolMIP will contribute towards advancing our understanding of the dominant mechanisms behind simulated post-eruption climate evolution, but also more generally of climate dynamics, decadal variability and of past transitions between different multi-centennial climate states, such as the transition between the so-called Medieval Climate Anomaly and Little Ice Age. Careful and systematic sampling of initial climate conditions and consideration of volcanic eruptions of different strength will help in better understanding the relative role of internal and externally-forced climate variability during periods of strong volcanic activity, hence improving the evaluation of climate models and advancing our understanding of past climates.

C. VolMIP will identify regions that are most robustly significantly affected by strong volcanic eruptions, and it will provide a framework for assessing the immediate as well as decadal climate repercussions of future volcanic events.

D. VolMIP will contribute towards advancing our understanding of the relative role of internal and volcanically-forced climate variability, therefore providing relevant information to policy makers concerning how the latter may contribute to the spread of future climate scenarios (where volcanic forcing is presently not accounted for).

All model output archived by CMIP6-Endorsed MIPs is expected to be made available under the same terms as CMIP output. Most modeling groups currently release their CMIP data for unrestricted use. If you object to open access to the output from your experiments, please explain the rationale.

No objection

List of output and process diagnostics for the CMIP DECK/CMIP6 data request:

VolMIP output is planned to be converted into the standard format using the CMOR package, following the same criteria adopted for past1000 and historical simulations. Additional output is needed for *Short* experiments, in particular for the DynVar diagnostic tool, which includes key diagnostics of parameterized and resolved wave forcings, radiative and latent heating rates. A daily temporal resolution of output data for the stratosphere is desirable.

Reply to WGCM Comments from the WGCM Synthesis of Comments on VolMIP Proposal for CMIP6

Original WGCM comments in italics.

Comments 1,2 and 4 are pointing out the same thing. Inclusion of the effect of ash deposition on snow and ice would be an interesting attempt, but, as is pointed out by Comment 4 itself, could complicate the experimental design for little scientific gain. Maybe it would be sufficient if the scientists involved bear in mind that the results may be slightly biased due to the lack of consideration on ash deposition. I do not think the lack somehow reduces the value of VolMIP.

We agree with WGCM that inclusion of volcanic ash deposition would complicate the design of VolMIP experiments for little scientific gain. We now specify in the description of the forcing input that “Tephra surface deposition is neglected in all the experiments.” Nonetheless, we propose VolMIP as an ideal framework for the modeling community to discuss sensitivity experiments focused on the climatic effects of tephra deposition.

Comment 5 can be addressed by either adding data assimilation procedure to the VolMIP protocol, or adding volcano experiments to the DCPP protocol, with the latter appearing to be simpler. Perhaps VolMIP and DCPP can communicate to discuss the best way to deal with volcano eruptions in a simple manner under the DCPP protocol. This will also enhance the presence of VolMIP community in CMIP6.

There are already ongoing coordinated activities between VolMIP and DCPP. Both groups started to discuss common experiments since a couple of months. VolMIP Tier 3 experiment *VolShort20EQini* (see Table 3) focuses on potential decadal climate predictability during periods of strong volcanic forcing. The experiment is designed as *VolShort20EQfull*, but as decadal prediction runs joint experiment with DCPP (C3.4). A first preliminary experimental set up was discussed at the MIKLIP/SPECS meeting in Offenbach and we expect the final design for this experiment to be defined in the Aspen workshop in June 2015. Claudia Timmreck will represent VolMIP there.

Comment 6 may require higher resolution for many of the models participating in VolMIP. Encouraging modeling groups capable of high resolution modeling to make analysis on this aspect would be constructive and ensure relevance of VolMIP to GC.

The model version used for the VolMIP experiments is the same used for the DECK experiments to ensure comparability between VolMIP results and *past1000* and *historical* simulations (for the latter, as long as volcanic forcing is prescribed through aerosol optical parameters). However, VolMIP would certainly benefit and welcome the use of high-resolution models in additional sensitivity experiments.

Attribution of regional climate changes during periods of strong volcanic activity is one of VolMIP’s specific foci. VolMIP’s Tier 1 *VolLongS60EQ* and Tier 2 *VolLongC19thC* experiments will contribute improving our understanding – also about its attribution - of one of the major regional climatic events occurred in Europe in the pre-industrial millennium: the year without a summer in the aftermath of the 1815 Tambora eruption.

Present research on changes in extreme events, such as frequency and intensity of hot and cold spells as well as heavy rainfall has been successfully conducted with present-generation models. Higher resolutions would certainly be beneficial, particularly to rainfall extremes, but is not essential.

Comments 3,7 do not require any direct response
Agreed

References

Anchukaitis K, Breitenmoser P , Briffa K, Buchwal A, Büntgen U, Cook E, D'Arrigo R, Esper J, Evans M, Frank D, Grudd H, Gunnarson B, Hughes M, Kirdyanov A, Körner C, Krusic P , Luckman B, Melvin T, Salzer M, Shashkin A, Timmreck C, Vaganov E, Wilson R. (2012) Tree-rings and volcanic cooling. Nature Geoscience, 5: 836-837doi:10.1038/ngeo1645

Arfeuille, F., B. P. Luo, P. Heckendorn, D. Weisenstein, J. X. Sheng, E. Rozanov, M. Schraner, S. Brönnimann, L. W. Thomason, and T. Peter (2013), Uncertainties in modelling the stratospheric warming following Mt. Pinatubo eruption,. Atmos. Chem. Phys., 13, 11221-11234, doi:10.5194/acp-13-11221-2013, 2013

Berdahl, M., and A. Robock (2013) Northern Hemispheric cryosphere response to volcanic eruptions in the Paleoclimate Modeling Intercomparison Project 3 last millennium simulations, J. Geophys. Res. Atmos., 118, 12,359–12,370, doi:10.1002/2013JD019914

Cole-Dai J, D. Ferris, A. Lanciki, J. Savarino , M. Baroni, MH Thiemens (2009) Cold decade (AD 1810–1819) caused by Tambora (1815) and another (1809) stratospheric volcanic eruption, Geophys. Res. Lett., 36,L22703 doi:10.1029/2009GL04088.

Driscoll, S., A. Bozzo, L. J. Gray, A. Robock, and G. Stenchikov (2012) Coupled Model Intercomparison Project 5 (CMIP5) simulations of climate following volcanic eruptions, J. Geophys. Res., 117, D17105, doi:10.1029/2012JD017607

D'Arrigo, R., Wilson, R., & Anchukaitis, K. J. (2013) Volcanic cooling signal in tree ring temperature records for the past millennium. Journal of Geophysical Research: Atmospheres, 118(16), 9000-9010

Ding, Y., J. A. Carton, G. A. Chepurin, G. Stenchikov, A. Robock, L. T. Sentman, and J. P. Krasting (2014) Ocean response to volcanic eruptions in Coupled Model Intercomparison Project 5 (CMIP5) simulations. J. Geophys. Res., 119, 5622-5637, doi:10.1002/2013JC009780.

Driscoll, S., Bozzo, A., Gray, L. J., Robock, A., & Stenchikov, G. (2012) Coupled Model Intercomparison Project 5 (CMIP5) simulations of climate following volcanic eruptions. Journal of Geophysical Research: Atmospheres, 117, D17105, doi:10.1029/2012JD017607.

Froelicher, T. L., F. Joos, C. C. Raible, J. L. Sarmiento (2013) Atmospheric CO2 response to volcanic eruptions: the role of ENSO, season, and variability. Global Biogeochemical Cycles, 27, 239-251

Gettelman, A. (2015) Volcanic SO2 and Forcing for CMIP6, personal communication (email 11/03/2015)

Gregory, J. M. (2010) Long-term effect of volcanic forcing on ocean heat content. Geophys. Res. Lett., 37, L22701, doi:10.1029/2010GL045507

Hegerl, G., J. Luterbacher, F. González-Rouco, S. F. B. Tett, T. Crowley and E. Xoplaki (2011) Influence of human and natural forcing on European seasonal temperatures. Nat. Geosc. 4:99-103, doi:10.1038/NGEO1057

Iles C. and Hegerl G.C. (2014) The global precipitation response to volcanic eruptions in the CMIP5 models. Environm. Res. Lett., 9, 104012

Mann, M.E., Fuentes, J.D., Rutherford, S. (2012) Underestimation of volcanic cooling in tree-ring based reconstructions of hemispheric temperatures. Nature Geosciences, doi 10.1038/ngeo1394

Mann, M. E., Rutherford, S., Schurer, A., Tett, S. F., & Fuentes, J. D. (2013) Discrepancies between the modeled and proxy-reconstructed response to volcanic forcing over the past millennium: Implications and possible mechanisms. Journal of Geophysical Research: Atmospheres, 118(14), 7617-7627

Marotzke, J., and Forster, P. M. (2015) Forcing, feedback and internal variability in global temperature trends. Nature, 517, 565-570. doi:10.1038/nature14117

Mignot, J., M. Khodri, C. Frankignoul, and J. Servonnat (2011), Volcanic impact on the Atlantic Ocean over the last millennium, Clim. Past, 7, 1439–1455, doi:10.5194/cp-7-1439-2011

Miller, G. H., Geirsdóttir, Á., Zhong, Y., Larsen, D. J., Otto-Bliesner, B. L., Holland, M. M., Bailey, D. A., Refsnider, K. A., Lehman, S. J., Southon, J. R., Anderson, C., Björnsson, H., and Thordarson, T. (2012) Abrupt onset of the Little Ice Age triggered by volcanism and sustained by sea-ice/ocean feedbacks, Geophys. Res. Lett., 39, L02708, doi:10.1029/2011GL050168

Muthers, S., J. G. Anet, E. Rozanov, C. C. Raible, T. Peter, A. Stenke, A. Shapiro, J. Beer, F. Steinhilber, S. Broennimann, F. Arfeuille, Y. Brugnara, and W. Schmutz (2014a) Sensitivity of the winter warming pattern following tropical volcanic eruptions to the background ozone climatology, Journal of Geophysical Research, 119, 1340-1355. DOI:10.1002/2013JD020138

Muthers, S., F. Arfeuille, and C. C. Raible (2015) Dynamical and chemical ozone perturbations after volcanic eruptions: Role of the climate state and the strength of the eruption. Atmos. Chem. Phys., submitted

Santer, B. D. et al. (2014) Volcanic contribution to decadal changes in tropospheric temperature. Nature Geosc., 7(3), 185–189, doi:10.1038/ngeo2098

Schurer, A., Hegerl, G.C., Mann, M., Tett, S.F.B., Phipps, S (2013) Separating forced from chaotic variability over the last millennium. J Climate, doi:10.1175/JCLI-D-12-00826.1

Schleussner, C. F. and Feulner, G. (2013) A volcanically triggered regime shift in the subpolar North Atlantic Ocean as a possible origin of the Little Ice Age, Clim. Past, 9, 1321–1330, doi:10.5194/cp-9-1321-2013

Soden, B. J., R. T. Wetherald, G. L. Stenchikov, and A. Robock (2002) Global cooling after the eruption of Mount Pinatubo: A test of climate feedback by water vapor. Science 296(5568): 727-730, doi:10.1126/science.296.5568.727

Thomason, L., J. P. Vernier, A. Bourassa, F. Arfeuille, C. Bingen, T. Peter, B. Luo (2015) Stratospheric Aerosol Data Set (SADS Version 2) Prospectus, In preparation for GMD

Timmreck C. (2012) Modeling the climatic effects of volcanic eruptions, invited review paper Wiley Interdisciplinary Reviews: Climate Change, doi: 10.1002/wcc.192

Toohey M, K. Krüger, M. Bittner, C. Timmreck, H. Schmidt (2014) The impact of volcanic aerosol on the Northern Hemisphere stratospheric polar vortex: mechanisms and sensitivity to forcing structure, Atmos. Chem. Phys., 14, 13063-13079, doi:10.5194/acp-14-13063-2014

Zanchettin, D., C. Timmreck, H.-F. Graf, A. Rubino, S. Lorenz, K. Lohmann, K. Krueger, and J. H. Jungclaus (2012) Bi-decadal variability excited in the coupled ocean–atmosphere system by strong tropical volcanic eruptions. Clim. Dyn., 39:1-2, 419-444, doi:10.1007/s00382-011-1167-1

Zanchettin, D., O. Bothe, H. F. Graf, S. J. Lorenz, J. Luterbacher, C. Timmreck and J. H. Jungclaus (2013) Background conditions influence the decadal climate response to strong volcanic eruptions, J. Geophys. Res. Atmos., 118, doi:10.1002/jgrd.50229

Zanchettin, D., et al. (2015) VolMIP - Model Intercomparison Project on the climate response to volcanic forcing. In preparation for GMD

Zanchettin, D., O. Bothe, C. Timmreck, J. Bader, A. Beitsch, H.-F. Graf, D. Notz and J. H. Jungclaus (2014) Inter-hemispheric asymmetry in the sea-ice response to volcanic forcing simulated by MPI-ESM (COSMOS-Mill). Earth Syst. Dynam., 5, 223–242, doi:10.5194/esd-5-223-2014

Table 1 – Tier 1 VolMIP experiments

Name	**Description**	**Start year**	**Configuration**	**Ens. Size**	**Years per simulation (minimum)**	**Total years**	**Gaps of knowledge being addressed with this experiment**
VolLongS60EQ	Idealized equatorial eruption corresponding to an initial emission of 60 Tg of SO_2. This eruption has a magnitude roughly corresponding to the 1815 Tambora eruption, the largest historical tropical eruption, which was linked to the so-called "year without a summer" in 1816	PID (from *PiControl*)	AOGCM/ESM	9	20	180	Uncertainty in the climate response to strong volcanic eruptions, with focus on coupled ocean -atmosphere feedbacks and interannual to decadal global as well as regional responses. The mismatch between reconstructed and simulated climate responses to historical strong volcanic eruptions, with focus on the role of simulated background internal climate variability.
VolShort20EQfull	1991 Pinatubo forcing as used in the CMIP6 *historical* simulations. Requires special diagnostics of parameterized and resolved wave forcings, radiative and latent heating rates. A large number of ensemble members is required to address internal atmospheric variability	PID	AOGCM/ESM	25	3	75	Uncertainty in the climate response to strong volcanic eruptions with focus on short-term response. Robustness of volcanic imprints on Northern Hemisphere's winter climate and of associated dynamics.
VolShort20EQsurf	As VolShort20EQfull, but with prescribed perturbation to the shortwave flux to mimic the attenuation of solar radiation by volcanic aerosols	PID	AOGCM/ESM	25	3	75	Mechanism(s) underlying the dynamical atmospheric response to large volcanic eruptions, in particular in Northern Hemisphere's winters. The experiment considers only the effect of volcanically induced surface cooling. Complimentary experiment to VolShort20EQstrat.
VolShort20EQstrat	As VolShort20EQfull, but with prescribed perturbation to the total (LW+SW) radiative heating rates	PID	AOGCM/ESM	25	3	75	Mechanism(s) underlying the dynamical atmospheric response to large volcanic eruptions, in particular in Northern Hemisphere's winter. The experiment considers only the effect of volcanically-induced stratospheric heating. Complimentary experiment to VolShort20EQsurf.

Vol = Volcano, Long = long-term simulation, Short = short-term simulation, S = Single (XXX = approx. amount of Tg of SO_2 release), C = Cluster (XXX = approx. period of the cluster), HL = high latitude, EQ = equator, full = full-forcing simulation, surf = short-wave forcing only, strato = stratospheric thermal (long-wave) forcing only, slab = slab ocean simulation, ini = simulation initialized for decadal prediction

Table 2 – Tier 2 VolMIP experiments

Name	**Description**	**Start year**	**Configuration**	**Ens. Size**	**Years per simulation**	**Total years**	**Connection with other MIPs**	**Gaps of knowledge being addressed with this experiment**
VolLongS100HL	Idealized high-latitude (60°N) eruption emitting 100 Tg of SO_2 over five months. The eruption's strength and length roughly correspond to that of the 1783-84 Laki eruption.	PID	AOGCM/ESM	9	20	180	PMIP, GeoMIP	Uncertainty in climate response to strong high-latitude volcanic eruptions (focus on coupled ocean-atmosphere). Laki has a unique eruption style (large SO_2 mass releases occurred at short temporal intervals). Outstanding questions about the magnitude of the climatic impact of high-latitude eruptions.
VolLongC19thC	Early 19th century cluster of strong tropical volcanic eruptions, including the 1809 event of unknown location, and the 1815 Tambora and 1835 Cosigüina eruptions.	PID (integration starts on year 1809)	AOGCM/ESM	3	50	150	PMIP, GeoMIP	Uncertainty in the multi-decadal climate response to strong volcanic eruptions (focus on long-term climatic implications). Contribution of volcanic forcing to the climate of the early 19th century, the coldest period in the past 500 years. Discrepancies between simulated and reconstructed climates of the early 19th century.

Vol = Volcano, Long = long-term simulation, Short = short-term simulation, S = Single (XXX = approx. amount of Tg of SO_2 release), C = Cluster (XXX = approx. period of the cluster), HL = high latitude, EQ = equator, full = full-forcing simulation, surf = short-wave forcing only, strato = stratospheric thermal (long-wave) forcing only, slab = slab ocean simulation, ini = simulation initialized for decadal prediction

Table 3 – Tier 3 VolMIP experiments

Name	**Description**	**Start year**	**Configuration**	**Ens. Size**	**Years per simulation**	**Total years**	**Connection with other MIPs**	**Gaps of knowledge being addressed with this experiment**
VolShort20EQslab	As VolShort20EQfull, but with a slab ocean	PID	AOGCM/ESM	25	3	75	ENSOMIP DCPP	Effects of volcanic eruptions on ENSO dynamics.
VolShort20EQini/ DCPP C3.4	As VolShort20EQfull, but as decadal prediction runs joint experiment with DCPP. Forcing input and implementation of the forcing fully comply with the VolMIP protocol	2015	AOGCM/ESM	10(5)	5	50	DCPP	Influence of large volcanic eruptions in future climate. Influence of large volcanic eruptions on seasonal and decadal climate predictability

Vol = Volcano, Long = long-term simulation, Short = short-term simulation, S = Single (XXX = approx. amount of Tg of SO_2 release), C = Cluster (XXX = approx. period of the cluster), HL = high latitude, EQ = equator, full = full-forcing simulation, surf = short-wave forcing only, strato = stratospheric thermal (long-wave) forcing only, slab = slab ocean simulation, ini = simulation initialized for decadal prediction

WCRP COORDINATED REGIONAL DOWNSCALING EXPERIMENT (CORDEX)

Application for CMIP6-Endorsed MIPs

Date: 31 March 2015

- Name of MIP: Coordinated Regional Downscaling Experiment (CORDEX)
- Co-chairs of MIP (including email-addresses)

  Filippo Giorgi <giorgi@ictp.it> and William Gutowski <gutowski@iastate.edu>
- Members of the Scientific Steering Committee

Anne Frigon <Frigon.Anne@ouranos.ca>
Daniela Jacob <daniela.jacob@hzg.de>
Hyung-Suk Kang <hyunskang@korea.kr>
R. Krishnan <krish@tropmet.res.in>
Chris Lennard <lennard@csag.uct.ac.za>
Grigory Nikulin <grigory.nikulin@smhi.se>
Silvina Solman <solman@cima.fcen.uba.ar>
Tannecia Stephenson <tannysyd@yahoo.com>
Bertrand Timbal <B.Timbal@bom.gov.au>
Fredolin Tangang <ftangang@gmail.com>
Director, International Project Office for CORDEX: Eleanor O’Rourke <eleanor.orourke@smhi.se>
WCRP liaison: Boram Lee <blee@wmo.int>

- Link to website: http://wcrp-cordex.ipsl.jussieu.fr/

- Goal of the MIP and a brief overview:

The overall vision of CORDEX is to advance and coordinate the science and application of regional climate downscaling through global partnerships. CORDEX has a set of four goals:

1. To better understand relevant regional/local climate phenomena, their variability and changes, through downscaling.
2. To evaluate and improve regional climate downscaling models and techniques
3. To produce coordinated sets of regional downscaled projections worldwide
4. To foster communication and knowledge exchange with users of regional climate information.

The RCD information samples uncertainties in regional climate change associated with varying forcing from GCM simulations and greenhouse gas concentration scenarios, natural climate variability and different downscaling methods. The CORDEX downscaling activities base themselves as much as possible on the latest sets of GCM climate simulations. For example, the CORDEX Phase I RCM experiments were based on driving GCMs participating to CMIP5, which was an invaluable resource for the design and implementation of CORDEX.

More generally, RCD techniques, including both dynamical and statistical approaches, are being increasingly used to provide higher-resolution climate information than is available directly from contemporary global climate models. The techniques available, their applications, and the community using them are broad and varied, and this is a growing area. These techniques, and the results they produce must be applied appropriately and their strengths and weaknesses need to be understood. This requires a better evaluation and quantification of the performance of the different techniques for application to specific problems, along with an understanding of physical processes and uncertainties underlying regional climate projections. Building on experience gained in the global modelling community, a coordinated, international effort to objectively assess and intercompare various RCD techniques provides a means to evaluate their performance, to illustrate benefits and shortcomings of different approaches, to produce multi-model, multi-method based information and to provide a more solid scientific basis for impact assessments and other uses of downscaled climate information.

The WCRP views regional downscaling as both an important research topic and an opportunity to engage a broader community of climate scientists in its activities. The Coordinated Regional Climate Downscaling Experiment (CORDEX) has served as a catalyst for achieving this goal.

- References:

Many papers have been published using simulations in the CORDEX framework; some are listed at http://wcrp-cordex.ipsl.jussieu.fr/index.php/cordex-peer-review-publications. Giorgi et al. (2009, "Addressing climate information needs at the regional level: the CORDEX framework", *WMO Bulletin*, **58,** 175-183) and Jones et al. (2011, "The Coordinated Regional Downscaling EXperiment CORDEX, an international downscaling link to CMIP5." *CLIVAR Exchanges*, 16, 34-40) give a brief overview of initial program plans. General updates appear in the *WCRP CORDEX Newsletter* (http://wcrp-cordex.ipsl.jussieu.fr/index.php/cordex-newsletters).

- An overview of the proposed experiments:

The anticipated CORDEX experiments are downscaling activities that will use CMIP DECK, CMIP6 Historical Simulation and ScenarioMIP output to provide input conditions for both statistical and dynamical downscaling under the CORDEX framework. CORDEX has a general framework of specified regions, resolutions and simulation periods that all regional CORDEX activities adhere to. Specific details of downscaling experiments are a function of plans generated by groups participating in each of the CORDEX regions. In particular, for each region a matrix of GCM-RCD experiments is designed based on the need to cover as much as possible different dimensions of the uncertainty space (different scenarios, GCMs, RCD models and techniques). The dimension of this matrix depends on the participation of groups in the different regional domain activities.

An optimal design of GCM-RCM matrices requires the availability of a broad range of driving GCM data (6 hourly meteorological fields), spanning high-end, mid-level and low-end GHG emission scenarios, and all or at least a large portion of GCMs participating in CMIP6. For the initial stages of the CORDEX activities, the focus will be on historical climate simulations for the 20th century and projections for 21st century, implying that data would be needed minimally for the period 1950-2100 (but ideally 1900-2100). Therefore, as for CMIP5, 6-hourly forcing data from one realization of each contributing GCM is a minimal requirement.

CORDEX activities provide a unique opportunity to deliver a full range of the uncertainties attached with regional climate change projections by creating GCM-RCD matrices. It is therefore important that the uncertainties attached to the human activities in the 21st century are encapsulated; multiple scenarios will allow us to evaluate some of the uncertainty due to human choices and are therefore an important additional request should they become available as part of the CMIP6 simulations. In addition, multiple realizations from some GCMs would allow us to explore also another dimension of the uncertainty space, GCM/RCM internal variability.

- An overview of the proposed evaluation/analysis of the CMIP DECK, CMIP6 Historical Simulation and ScenarioMIP experiments:

CORDEX experiments would use output from

1) 30 years of the pre-industrial simulation (CMIP DECK)

2) 1950-2014 from the historical climate simulation (CMIP6 Historical Simulation)

3) 2015-2100 from the transient scenario climate simulation that uses RCP8.5, 4.5 and 2.6 for one realization of future projection (ScenarioMIP).

We request output from these three RCPs to span the range of plausible climate change, maintain continuity with CMIP5-based downscaling and satisfy needs for climate change information ascertained from user communities. Although one realization is requested, providing output from more realizations and more scenarios (from ScenarioMIP) is very welcome.

The downscaling activities will contribute to answering all three of the key questions for CMIP6 through regional simulations with different climate forcings (key question 1), evaluation of physical processes affecting added value and biases in the downscaled results (key question 2) and characterization of the impact of unforced variability, both internally generated and via ensemble boundary conditions, on the ratio of regional climate change signals versus the noise of unforced variability (key question 3).

The downscaling activities will contribute primarily to the WCRP grand challenges of regional climate information and climate extremes. Some of the downscaling will include evaluation of regional feedbacks associated with land-use change and aerosols, along with regional rendition of GCM responses to different climatic forcings.

Downscaled results using CMIP output will be evaluated for their ability to provide added value to the CMIP simulations. This will occur in three ways:

1) Analysis during the historical period (1950-2014) will indicate where and when the downscaling provides regional detail of physical behavior that agrees better with observations than the driving GCM output and provides robust additional fine scale climate information.

2) Analysis of downscaled projections (2015-2100) will assess where and when the downscaling provides regional detail of physical behavior that exceeds noise levels of unforced internal variability.

3) Analysis of downscaled CMIP DECK simulations for the pre-industrial control compared to the transient forcing case will determine potential regional climate-change detection. Should additional CMIP6 simulations occur that specify changes in just one of the major forcings (e.g., solar output, greenhouse gases, volcanic aerosols), then item 3) above would include additional downscaling of those runs with an eye toward regional attribution.

➢ Proposed timing:

At least some regional modeling groups will be poised to use CMIP DECK, CMIP6 Historical Simulation and ScenarioMIP output suitable for RCM boundary conditions as it becomes available.

The statistical downscaling program under CORDEX is in development. However, some participants in the program have been using CMIP5 output and should be ready to use appropriate CMIP DECK, CMIP6 Historical Simulation and ScenarioMIP output as it becomes available.

➢ For each proposed experiment to be included in CMIP6: N/A

➢ All model output archived by CMIP6-Endorsed MIPs is expected to be made available under the same terms as CMIP output. Most modeling groups currently release their CMIP data for unrestricted use. If you object to open access to the output from your experiments, please explain the rationale:

CORDEX is preparing a Memorandum of Understanding for output produced by CORDEX modelers that will follow the availability terms of CMIP output.

➢ List of output and process diagnostics for the CMIP DECK/Historical/ScenarioMIP data request:

CORDEX requests output from the targeted CMIP DECK, CMIP6 Historical Simulation and ScenarioMIP simulations sufficient to allow the downscaling activities and associated analyses listed above:

1) Output sufficient for dynamical and empirical statistical downscaling (transient climate-change simulation)
2) Output from multiple realizations of the same GCM for both the pre-industrial and transient climate-change simulations, to bring unforced variability into downscaling boundary conditions
3) Output that could allow regional detection and attribution work. This would entail boundary conditions from pre-industrial control runs (CMIP DECK) and runs with changes in only one climate forcing (if part of CMIP6).

Output variables needed from CMIP DECK, CMIP6 Historical Simulation and ScenarioMIP runs:

- Preferred output period: 1951-2100 for transient climate change (RCP8.5, RCP4.5 and RCP2.6 for 2015-2100); 30 years of pre-industrial control.

- For dynamical downscaling:

- 6-hourly instantaneous surface pressure
- 6-hourly instantaneous three-dimensional fields of temperature, atmospheric specific humidity, zonal wind and meridional wind

We suggest saving these variables to files with the same time period (e.g., 6 months, one year), to ensure uniform time periods covered for a GCM’s files for all variables and to avoid very large files (many Gb) that are awkward to handle.

- For statistical downscaling, in addition to the 6-hourly three-dimensional fields listed above, values for integrated quantities will be required:

- maximum daily surface (2m) temperature
- minimum daily surface (2m) temperature
- daily surface temperature (2m)
- daily surface dewpoint temperature (2m)
- daily zonal wind (10m)
- daily meridional wind (10m)
- daily precipitation
- daily vertical atmospheric column of water (or precipitable water)
- monthly sea surface temperature

- Supplementary variables that are desirable:

- daily soil moisture (vertically integrated)
- daily snow density
- daily snow albedo
- daily low and medium cloud cover
- 6-hourly instantaneous geopotential height at 850, 700 and 500 hPa

➢ Any proposed contributions and recommendations for observations

Assessments of added value will seek fine resolution (25-50 km or less) observational datasets. The obs4MIPs and ana4MIPs efforts are potentially useful and there is already some CORDEX interaction with obs4MIPS. For some regions, fine resolution observational datasets are beings sought in all CORDEX regions, especially those that could support evaluation of higher resolution CORDEX runs. CORDEX will help with efforts to make new datasets accessible in standardized formats via the ESGF infrastructure.

➢ Any proposed changes from CMIP5 in NetCDF metadata (controlled vocabularies), file names, and data archive (ESGF) search terms: NONE

➢ Explanation of any proposed changes (relative to CMIP5) that will be required in CF, CMOR, and/or ESGF: NONE

DynVar – Diagnostic MIP

Dynamics and Variability of the Stratosphere ‐ Troposphere System

Co-Chairs:
Edwin Gerber (gerber@cims.nyu.edu)
Elisa Manzini (elisa.manzini@mpimet.mpg.de)

Members of the Scientific Steering Committee:
Amy Butler (amy.butler@noaa.gov)
Natalia Calvo (nataliac@fis.ucm.es)
Andrew Charlton-Perez (a.j.charlton-perez@reading.ac.uk)
Marco Giorgetta (marco.giorgetta@mpimet.mpg.de)
Adam Scaife (Adam.scaife@metoffice.gov.uk)
Tiffany Shaw (tas1@uchicago.edu)
Shingo Watanabe (wnabe@jamstec.go.jp)

Website: http://www.sparcdynvar.org/

Goal of the MIP

DynVar focuses on the interactions between atmospheric variability, dynamics and climate change, with a particular emphasis on the two-way coupling between the troposphere and the stratosphere. The key questions addressed by the activity are:

- How do dynamical processes contribute to persistent model biases in the mean state and variability of the atmosphere, including biases in the position, strength, and statistics of blocking events, storm tracks and the stratospheric polar vortex?
- How does the stratosphere affect climate variability at intra-seasonal, inter-annual and decadal time scales?
- What is the role of dynamics in shaping the climate response to anthropogenic forcings (e.g. global warming, ozone depletion) and how do dynamical processes contribute to uncertainty in future climate projections?

Rather then proposing new experiments, we are ***requesting additional output***, critical for understanding the role of atmospheric dynamics in both present and past climate, and future climate projections. ***Without this output, we will not be able to fully assess the dynamics of mass, momentum, and heat transport - essential ingredients in projected circulation changes - nor take advantage of the increasingly accurate representation of the stratosphere in coupled climate models.*** Our rational is that by simply extending the standard output relative to that in CMIP5, there is potential for significantly expanding our research capabilities in atmospheric dynamics.

Proposed evaluation/analysis of the CMIP DECK and CMIP6 experiments

Understanding circulation changes in the atmosphere, particularly of the mid-latitude storm tracks, has been identified by the World Climate Research Programme (WCRP) as one of the grand challenges in climate research. Changes in the storm tracks are significantly coupled with lower atmosphere processes such as surface temperature gradients and moisture availability (e.g. Booth et al. 2013) as well as with processes in the stratosphere, from natural variability on synoptic to intraseasonal timescales (e.g. Baldwin and Dunkerton 2001) to the response to changes in stratospheric ozone (e.g. Son et al. 2008) and other anthropogenic forcings (e.g. Scaife et al. 2012). The storm tracks depend critically on the transport of momentum, heat and chemical constituents throughout the whole atmosphere. Both resolved (primarily Rossby) and parameterized (gravity) waves play the key roles in these transports, and it is important that the standard output of the DECK experiments, the CMIP6 Historical Simulation and (in principle) any MIP experiment allow proper diagnosis of these wave fluxes.

The lack of output is particularly acute in the stratosphere, where daily means of standard variables (e.g., zonal and meridional winds, height and temperature) and parameterized gravity wave forcings (a key driver of the circulation) were not well documented in CMIP5, and resolved waves could at best be coarsely assessed, given the importance of the vertical structure to momentum and mass transport. As detailed by Hardiman et al. (2013), the stratospheric community had to rely on direct collaboration to obtain necessary diagnostics to assess the Brewer-Dobson circulation, the first order circulation of mass and momentum in the stratosphere. Daily means of standard variables in both the troposphere and stratosphere would expand our ability to assess the synoptic dynamics of the atmosphere.

Investigation of the impact of solar variability and volcanic eruptions on climate also relies heavily on atmospheric wave forcing diagnostics, as well as radiative heating rates (particularly in the short wave). By extending our request to the energy budget and including diagnostics such as diabatic heating from cloud-precipitation processes, research on the links between moist processes and atmospheric dynamics will be enabled as well. The interplay between moist processes and circulation is central to the WCRP Grand Challenge on Clouds, Circulation and Climate Sensitivity (Bony et al. 2015).

The CMIP5 saw a significant upward expansion of models with a more fully resolved stratosphere (e.g. Gerber et al. 2012), and several multi-model studies have investigated the role of the stratosphere in present climate and in projections of future climate (e.g., Anstey et al. 2013; Charlton-Perez et al. 2013; Gerber and Son, 2014; Hardiman et al. 2013; Lott et al. 2014; Manzini et al. 2014; Min and Son 2013; Shaw et al. 2014; Wilcox and Charlton-Perez 2013) in addition to many other single model studies. These studies document a growing interest in the role of middle and upper atmosphere in climate, research that would take full advantage of these diagnostics.

Key science questions of CMIP6:

DynVar primarily addresses CMIP6 key science questions on the origin and consequences on systematic models biases in the context of atmospheric dynamics and on the storm track theme of the Clouds, Circulation and Climate Sensitivity Grand Challenge, by further enabling and stimulating research on atmospheric dynamics and storm tracks with CMIP models. The DynVar focus on daily fields and diagnostics of the atmospheric flow is extremely relevant to the Grand Challenge on Climate Extremes. We envision as well contributions to the questions on how the Earth System responds to forcing, assessments of future climate changes, and on the Grand Challenges on Regional Climate Information and on the Biospheric Forcings and Feedbacks theme.

Scientific Analysis Plan:
Analyses of the atmospheric circulation within the DECK experiments (AMIP and 40-year periods of control, abrupt4xCO2 and 1pctCO2), the CMIP6 historical and the ScenarioMIP RCP8.5 experiments (40 years, each) are the highest priority for DynVar. (Models need only commit to providing diagnostics to the DECK and the CMIP6 historical experiments, however, to participate in our MIP.) DynVar will be holding a workshop in June 2016 to organize our efforts. The goal of the workshop is to coordinate analysis of the CMIP6 simulations, avoiding duplicate efforts and ensuring that all key areas are investigated. To enhance participation and collaboration with the modeling centers, representatives will be invited to attend both the workshop and to participate in the scientific analysis and papers. DynVar has established a tradition of bringing modelers and theoreticians together.

In the past we have found that research on a mechanistic understanding of the atmosphere and on rectifying model biases is often best organized organically, rather than from a top down approach. The TEM diagnostics, for example, have been used in a number of CMIP5 studies, in addition to many mechanistic studies. DynVar is seeking to facilitate this research by making the key diagnostics available, more than "claiming" research topics as our own.

We appreciate, however, that modeling centers need a firm commitment that these new diagnostics will be used. As expressed above, the key scientific question addressed by the DynVar data request is "***What are the origins and consequences of systematic model biases in atmospheric dynamics?***" DynVar is prepared to take the lead on inter-model comparison paper(s) that investigate the momentum and heat balances of the historical climate (where it can be compared with observations and reanalysis) and how biases there relate to differences in the models's atmospheric circulation response to external forcing, both in the idealized DECK perturbation experiments and in the RCP8.5. The importance and challenge of addressing the atmospheric circulation response to global warming are highlighted in Shepherd (2014) and Vallis et al (2014). We note also that formerly organized DynVar activity around the CMIP5 lead to three papers (Gerber et al. 2012, Charlton-Perez et al. 2013, and Manzini et al. 2014). Many more papers arose independently.

Synergy with other MIPs:
We are actively coordinating our efforts with several proposed MIPs. Transport plays a key role in the AerChemMIP experiments with ozone depleting substances, making the TEM diagnostics particularly relevant. The short-term VolMIP experiments and the SolarMIP experiments focus in large part on stratosphere-troposphere coupling, where the momentum and heat budget diagnostics are directly relevant. Lastly, gravity wave effects are a focus of the HiResMIP. We are particular interested in how the resolved circulation changes when the reliance on gravity wave parameterization is reduced. The availability of dynamically oriented diagnostics within the DECK and the CMIP6 historical will provide the benchmark for other MIPs as well. We envision

fruitful potential collaboration with the following proposed MIPs: DCPP, DAMIP, GMMIP, and ScenarioMIP.

List of output and process diagnostics for the CMIP DECK/CMIP6 data request

We stress the need of archiving standard variables (e.g. zonal and meridional winds, temperature, and geopotential height) as daily means in the troposphere and stratosphere. We expect that the location and total number of vertical pressure levels for daily mean fields will be discussed during the definition of the standard output. It is important to archive finer vertical resolution on daily time scales. While model centers have been saving increasingly fine horizontal resolution (close to the native model grid), vertical sampling has been limited to standard levels that changed little from CMIP3 to 5. Quasi-geostrophic theory suggests that motions of horizontal scale Δx are linked with vertical scale Δz by rotation and stratification: $\Delta x \sim N/f \, \Delta z$, where stratification is quantified by the buoyancy frequency N and rotation by the Coriolis parameter f. weak rotation in the tropics implies that even finer vertical scales may be required (Lindzen and Fox-Rabinovitz, 1989). Modelers are mindful of these constraints with the native grids, but coarse vertical sampling in the archived output similarly limits the effectiveness at which we can make use of finer horizontal scales to diagnose atmospheric variability and the transport of heat and momentum across scales: the analysis will be dominated by truncation errors associated with the vertical grid. We therefore propose that model centers archive daily mean of key dynamical variables on the standard 17 vertical levels, as done for the monthly mean fields.

We request archival of the Transformed Eulerian Mean (TEM) atmospheric circulation, which allows diagnosis of resolved wave driving and transport, and of parameterized atmospheric gravity wave driving. These diagnostics are also widely used in the analysis of chemistry climate models (e.g. CCMI). The TEM diagnostics are particularly sensitive to vertical resolution and model formulation (Hardiman et al. 2010), and so ideally computed following the model's dynamical core assumptions and on the native grid of the model, before coarsened for archival. In addition, we request the archival of heating rates. *Note that a number of the requested diagnostics are 2-D fields (zonal means) on an atmospheric grid defined by latitudes and pressure levels. We are targeting both daily and monthly diagnostics.*

We have subdivided the DynVar proposed variables into three groups, serving **three main objectives and science questions**:

(1) Atmospheric variability across scales (short name: variability). Evaluation of atmospheric variability across time and spacial scales, including model biases in the position, strength, and statistics of blocking events, storm tracks and the stratospheric polar vortices. Estimate of changes in atmospheric variability. Collateral objectives: Evaluation of the atmospheric mean state and its changes.

Variability: Standard CMIP5, daily / monthly mean frequency, *but more vertical levels for 3D daily*		
psl	Sea Level Pressure [Pa]	2D
pr	Precipitation [kg m^{-2} s^{-1}]	2D
tas	Near-Surface Air Temperature [K]	2D
uas	Eastward Near-Surface Wind [m s^{-1}]	2D

vas	Northward Near-Surface Wind [m s^{-1}]	2D
ta	Air Temperature [K]	3D
ua	Eastward Wind [m s^{-1}]	3D
va	Northward Wind [m s^{-1}]	3D
wap	omega (=dp/dt) [Pa s^{-1}]	3D
zg	Geopotential Height [m]	3D
hus	Specific Humidity [1]	3D

(2) Atmospheric (TEM) zonal momentum budget (short name: momentum budget). Evaluation of momentum transport by the free atmosphere and at the surface, to understand how dynamical processes contribute to persistent model biases in the mean state and variability of the atmosphere. In the free atmosphere, limited to the Transformed Eulerian Mean (TEM) zonal momentum balance [2D, daily and monthly]. The TEM diagnostics must be calculated from high frequency (6hr or shorter time intervals) atmospheric fields. The TEM diagnostics requested are derived from the TEM primitive equations in spherical, log-pressure coordinates (Andrews et al 1987). To close the momentum balance, additional tendencies are requested, from the deposition of momentum flux carried by parameterized gravity waves. At the surface, momentum balance is diagnosed by requesting total stress as well as the stress by the planetary boundary layer.

The priority of each variable is denoted by (1) for key variables necessary to complete the analysis and (2) for variables that will also be analyzed, but could be omitted.

Momentum Budget:		
2D monthly / daily and zonal mean variables (Grid: YZT)		
vstar (1)	Residual Northward Wind [m s^{-1}]	Standard CCMI
wstar (1)	Residual Upward Wind [m s^{-1}]	Standard CCMI
fy (1)	Upward EP-flux [N m^{-1}]	Standard CCMI
fz (1)	Northward EP-flux [N m-1]	Standard CCMI
utenddivf (1)	u-Tendency by EP-flux Divergence [m s^{-2}]	Renamed from CCMI: acceldivf
2D daily and zonal mean variables (Grid: YZT)		
utendogw (1)	u-Tendency by orographic gravity waves [m s^{-2}]	New: DYVR
utendnogw (1)	u-Tendency by non-orographic gravity waves [m s^{-2}]	New: DYVR
psistar (2)	Residual Stream Function [kg s^{-1}]	New: DYVR
utendvstarad (2)	u-Tendency by Residual Northward Wind Advection [m s^{-2}]	New: DYVR
utendwstarad(2)	u-Tendency by Residual Upward Wind Advection [m s^{-2}]	New: DYVR
3D monthly mean variables (Grid: XYZT)		
utendogw (1)	u-Tendency by Orographic Gravity Waves [m s^{-2}]	New: DYVR
utendnogw (1)	u-Tendency by Non-orographic Gravity Waves [m s^{-2}]	New: DYVR
vtendogw (2)	v-Tendency by Orographic Gravity Waves [m s^{-2}]	New: DYVR
vtendnogw (2)	v-Tendency by Non-orographic Gravity Waves [m s^{-2}]	New: DYVR
2D daiy mean variables (Grid: XYT)		
tauu (1)	Surface Downward Eastward Wind Stress, Total [Pa]	New: DYVR
tauv (1)	Surface Downward Northward Wind Stress, total [Pa]	New: DYVR
tauu_pbl (2)	Surface Downward Eastward Wind Stress by Planetary	New: DYVR

	Boundary Layer [Pa]	
tauv_pbl (2)	Surface Downward Northward Wind Stress by Planetary Boundary Layer [Pa]	New: DYVR

(3) Atmospheric heat budget (short name: heat budget). Evaluation of the dynamics of heat transfer in the free atmosphere, characterize impacts of external radiative forcing (solar, volcanic, anthropogenic) and enable analysis to break down feedbacks in Earth System models. Additional tendencies are requested for gravity wave diagnostics.

Heat Budget: 2D monthly and zonal mean variables (Grid: YZT)		
zmtnt (1)	Diabatic Heating Rate [K s^{-1}]	Standard CCMI
tntlw (1)	Longwave Heating Rate [K s^{-1}]	Standard CCMI
tntsw (1)	Shortwave Heating Rate [K s^{-1}]	Standard CCMI
tntogw (2)	Temperature Tendency by Orographic Gravity Wave Dissipation [K s^{-1}]	New: DYVR
tntnogw (2)	Temperature Tendency by Non-orographic Gravity Wave Dissipation [K s^{-1}]	New: DYVR

Prioritization

Objective **(1)** focuses on standard output variables and is a question of resolution. A compromise to balance storage constraints would be to provide daily data on a standard horizontal grid. Reducing horizontal resolution by a factor of $2^{1/2}$ to allow double the vertical resolution would provide a great payout, i.e. 17 vertical levels with 200 km resolution would be more valuable than 9 vertical levels on a 150 km grid.

We have prioritized the output variables for objectives **(2)** and **(3)** in the tables above. Gravity wave tendencies and long and short wave heating rates cannot be inferred indirectly, and are critical for assessing the momentum and heat budget. The TEM diagnostics could be assessed (albeit with significantly diminished accuracy) from daily output provided it contains finer vertical resolution. And additional advantage of computing them on the native grids is the ability to capture information about wave and heat fluxes more efficiently, exchanging 2D fields for 3D fields.

Technical Notes on Averaging and the TEM Calculations

The final CMIP6 archived output should be on the standard pressure levels. As detailed below, however, we request that the calculations of the TEM diagnostics be done as close to the model levels as possible before being interpolated to the standard pressure levels at the end.

Time means are calculated by averaging over the day or month periods, either “offline” from outputs at 6-hour or higher frequency or directly computed over all time steps (i.e., “inline”). Similarly, zonal means are calculated averaging over all available longitudes, either offline (more commonly) or online (possible, but seldom done). For flux quantities with multiplying factors (e.g. the heat flux v’θ’), the zonal mean anomalies (e.g. v’ and θ’) should be computed from high frequency data (e.g. 6-

hourly or higher frequency) and their products then computed before averaging to daily or monthly mean. Before the calculation of its vertical derivative, the potential temperature must be averaged, zonally and temporally (daily or monthly).

Although the residual velocities are usually displayed in a log-p vertical coordinate system, we recommend calculating them on grids very close or identical to those of the dynamical core of the atmospheric model. For models using pressure levels (or hybrid coordinates based on pressure), Andrews et al (1983) details the calculation in pressure coordinates, and shows how the results can be shifted to log pressure coordinates in the final step. We recommend interpolating the fields of interest to pressure levels prior to taking the zonal and temporal averages (for both inline and offline calculations). Ideally, the pressure levels should be as close as possible to the average position of the model levels, to minimize the impact of the interpolation.

The calculation of the TEM diagnostics and the residual winds are also defined in the textbook: Andrews et al (1987), as well as in the review by Butchart (2014), among other publications. We appreciate that interpolation to pressure levels (or equivalently, log pressure coordinates) is problematic for non-hydrostatic models. Hardiman et al. (2010) discuss how the TEM can be computed in this context.

References

Andrews, D. G., J.D. Mahlman, and R.W Sinclair (1983) Eliassen-Palm Diagnostics of wave-mean flow interaction in the GFDL SKYHI general circulation model, J. Atmos. Sci., 40, 2768-2784.

Andrews, D. G., J. R. Holton, and C. B. Leovy (1987) Middle Atmospheric Dynamics, 489 pp., Accademic Press

Anstey, J. A. and Coauthors (2013), Multi- model analysis of Northern Hemisphere winter blocking: Model biases and the role of resolution, J. Geophys. Res. Atmos., 118, 3956–3971, doi: 10.1002/jgrd.50231

Baldwin, M. P., and T. J. Dunkerton (2001), Stratospheric harbingers of anomalous weather regimes. Science, 294, 581–584.

Bony, S., and Coauthors (2015) Clouds, circulation and climate sensitivity, Nature Geoscience, DOI: DOI: 10.1038/NGEO2398.

Booth, J. F., S. Wang, L. Polvani (2013), Midlatitude storms in a moister world: lessons from idealized baroclinic life cycle experiments. Climate Dynamics 41, 787–802, doi: 10.1007/s00382-012-1472-3

Butchart, N. (2014), The Brewer-Dobson circulation, Rev. Geophys., 52, 157–184, doi:10.1002/2013RG000448.

Charlton-Perez, A. J. and Coauthors (2013), On the lack of stratospheric dynamical variability in low-top versions of the CMIP5 models, J. Geophys. Res. Atmos., 118, 2494–2505, doi: 10.1002/jgrd.50125

Gerber, E. P. and Coauthors (2012), Assessing and Understanding the Impact of Stratospheric Dynamics and Variability on the Earth System, Bull. Amer. Meteor. Soc., 93, 845-859, doi: 10.1175/BAMS-D-11- 00145.1

Gerber, E. P. and S.-W. Son, (2014), Quantifying the Summertime Response of the Austral Jet Stream and Hadley Cell to Stratospheric Ozone and Greenhouse Gases. J. Climate, 27, 5538-5559, doi: 10.1175/JCLI-D-13-00539.1

Hardiman, S. C., N. Butchart, N. Calvo (2013), The morphology of the Brewer-Dobson circulation and its response to climate change in CMIP5 simulations, Q. J. R. Meteorol. Soc., doi: 10.1002/qj.2258

Hardiman, S.C. et al. (2010), Using Different Formulations of the Transformed Eulerian Mean Equations and Eliassen–Palm Diagnostics in General Circulation Models, J. Atmos. Sci., 67, 1983-1995. DOI: 10.1175/2010JAS3355.1

Lindzen, R. S. and M. Fox-Rabinovitz (1989), Consistent Vertical and Horizontal Resolution, Mon. Wea. Rev., 117, 2575-2583.

Lott, F. and Coauthors (2014), Kelvin and Rossby-gravity wave packets in the lower stratosphere of some high-top CMIP5 models, J. Geo- phys. Res. Atmos., 119, 2156–2173, doi: 10.1002/2013JD020797

Manzini, E. and Coauthors (2014), Northern winter climate change: Assessment of uncertainty in CMIP5 projections related to stratosphere-troposphere coupling, J. Geophys. Res. Atmos., 119, doi: 10.1002/2013JD021403

Min, S.-K. and S.-W. Son (2013), Multi-model attribution of the Southern Hemisphere Hadley cell widening: major role of ozone depletion, J. Geophys. Res. Atmos., 118, 3007-3015.

Scaife, A. A. and Coauthors (2012) Climate change projections and stratosphere–troposphere interaction. Climate Dyn. doi: 10.1007/s00382-011-1080-7

Shaw, T. A., J. Perlwitz, O. Weiner (2014), Troposphere-stratosphere coupling: Links to North Atlantic weather and climate, including their representation in CMIP5 models. J. Geophys. Res., 10.1002/2013JD021191

Shepherd, T.G. (2014) Atmospheric circulation as a source of uncertainty in climate change projections, Nature Geoscience, DOI: 10.1038/NGEO2253.

Son, S.-W. and Coauthors (2008) The impact of stratospheric ozone recovery on the Southern Hemisphere westerly jet. Science, 320, 1486–1489.

Vallis, G. K., P. Zurita-Gotor, C. Cairns, and J. Kidston (2014) Response of the large-scale structure of the atmosphere to global warming, Q. J. R. Meteorol. Soc., DOI:10.1002/qj.2456

Wilcox, L. and A. Charlton-Perez (2013), Final warming of the Southern Hemisphere polar vortex in high- and low-top CMIP5 models. J. Geophys. Res. Atmos., 118, doi: 10.1002/jgrd.50254

SEA-ICE MODEL INTERCOMPARISON PROJECT (SIMIP)

CO-CHAIRS OF MIP (INCLUDING EMAIL-ADDRESSES)*

Alexandra Jahn (NCAR, US, ajahn@ucar.edu)
Dirk Notz (Max Planck Institute for Meteorology, Germany, dirk.notz@mpimet.mpg.de)

Members of the Scientific Steering Committee*

Marika Holland (NCAR, US)
Elizabeth Hunke (Los Alamos National Laboratory, US)
François Massonnet (Université catholique de Louvain, Belgium)
Julienne Stroeve (NSIDC, US)
Bruno Tremblay (McGill University, Canada)
Martin Vancoppenolle (Laboratoire d'Océanographie et du Climat, France)

LINK TO WEBSITE (IF AVAILABLE)*

http://www.climate-cryosphere.org/activities/groups/seaicemodeling/simip

GOAL OF THE MIP AND A BRIEF OVERVIEW*

Defines variables that are necessary to analyze sea-ice evolution in any CMIP6 experiment

This diagnostic MIP defines a list of variables that capture the evolution of sea ice in any experiment carried out as part of CMIP6. Given the importance of sea ice both as a driver and as an indicator of climatic changes, the analysis of the changing sea-ice cover in CMIP6 experiments provides insight into the time-integrated evolution of the climate system. To obtain all necessary information for such analysis for any given CMIP6 experiment is the overarching goal of this MIP.

To achieve this aim, we propose a list of those variables that are required to close the three budgets that govern the evolution of sea ice and its impact on the Earth's climate system. These are the conservation of heat, the momentum balance and tracer (salt) conservation. In addition, we provide a list of variables that allow for the analysis of the sea-ice state on a daily basis. We aim for the best possible compromise of output frequency and necessity of high-resolution sampling for closing the budgets. To achieve this aim, we group the variables according to their priorities, with the variables of the highest priority being necessary for a basic analysis of the sea-ice evolution in any CMIP6 experiment. By making sure that budgets can be closed, the analysis of sea ice in CMIP6 simulations has the potential to focus on processes rather than only on the sea ice state, leading to improved understanding of the biases in sea ice and the fidelity of projections of sea ice.

AN OVERVIEW OF THE PROPOSED EXPERIMENTS*

None

We do not propose any sea-ice specific experiments. Instead, we clearly define a list of variables that allow any scientist to analyze and intercompare the sea-ice state in any experiment that is carried out as part of CMIP6. The list of variables is accompanied by guidance as to how a standardized analysis of sea-ice evolution can be carried out that will allow for the straight-forward comparison of sea-ice evolution across different MIPs.

AN OVERVIEW OF THE PROPOSED EVALUATION/ANALYSIS OF THE CMIP DECK AND CMIP6 EXPERIMENTS*

Variables defining sea-ice state as well as atmospheric and oceanic forcing

The variables that are proposed in this MIP can be divided into those that determine the sea-ice state and those that determine the atmospheric and oceanic forcing that change this sea-ice state. State variables include standard output such as sea-ice area fraction and thickness, but also more advanced variables with lower priority, such as melt-pond coverage and information on the ice-thickness distribution for those models that have such information available.

Regarding the forcing, the proposed variables both on the atmospheric and on the oceanic side allow for a closure of the main budgets. Hence, they include a description of all heat fluxes that affect the ice, the momentum forcing and the transport of tracers (usually just salt) into and out of the ice. Also these forcing variables are split into routine output and more advanced measures.

Analyses carried out within SIMIP

As its primary aim, SIMIP establishes a protocol that allows for a unified analysis of the evolution of sea-ice across all CMIP6 activities. To foster such analysis, and to better understand the current and future evolution of sea ice, the researchers involved in SIMIP will analyse the sea-ice evolution of all DECK simulations that follow the SIMIP protocol. Through such analysis, we will determine what aspects of the large-scale sea-ice evolution show the most agreement and most disagreement across models. This will include an assessment of the main forcing that contributes to sea-ice retreat in the 1% increase simulations and an analysis of the main reason for different sensitivities of the sea-ice response in the 4x CO_2 experiments. We will also analyse the magnitude of internal variability of the sea-ice state in the control runs, and establish the main drivers of its differences across models.

Based on such analysis, which only becomes possible through the new set of variables that we define here, we will be able to identify the main drivers of sea-ice evolution in CMIP6 models, identify uncertainties in the formulation of these drivers in the models, and thus establish uncertainties in the evolution of sea ice that are caused by model uncertainties. This analysis will hence identify the most important gaps in our understanding of sea ice, will establish how these gaps limit our ability to project the future evolution of sea ice, and guide the planning of future observational campaigns.

This analysis will be presented in a joint publication with the modeling groups that implement the SIMIP protocol.

PROPOSED TIMING*

In parallel with all CMIP6 experiments

At least the standard variables to which we assign priority 1 in this MIP should be saved from any experiment carried out as part of CMIP6.

FOR EACH PROPOSED EXPERIMENT TO BE INCLUDED IN CMIP6**

We do not propose individual experiments. We therefore here only summarize why the variables we propose should be saved as output from any CMIP6 experiment.

THE EXPERIMENTAL DESIGN

The variables of this MIP should be saved independent of the experimental design

THE SCIENCE QUESTION AND/OR GAP BEING ADDRESSED WITH THIS EXPERIMENT;

The variables of this MIP will allow scientists to answer, for example, the following science questions:

- How sensitive is the sea-ice cover to changes in the external forcing?
- Are these changes primarily driven by changes in the atmosphere or in the ocean?
- What causes modelled biases in the simulation of the sea-ice state?
- What are the (coupled) processes and feedbacks that contribute most to a spread between model simulations?
- How much do simulations of the Earth's climate profit from improvements in the sea-ice model component?
- What's the internal variability of the Earth's sea-ice cover?
- How predictable is the sea-ice cover on time scales ranging from daily to decadal?
- What are the most pressing needs for observations?

POSSIBLE SYNERGIES WITH OTHER MIPS;

Sea ice integrates changes in the atmospheric and oceanic forcing

The variables that we propose will allow the users of any MIP to analyse and to understand the temporal evolution of the sea-ice cover in their simulations. Since the sea-ice state reflects changes in the climate system of the Earth on decadal time scales, changes in the sea-ice cover usually provide direct insight into Earth-System response to climate changes on time scales that are between the atmospheric and the oceanic response time. Such analysis is hence helpful in understanding the temporal evolution of changes in these other compartments of the Earth System.

POTENTIAL BENEFITS OF THE EXPERIMENT TO (A) CLIMATE MODELING COMMUNITY, (B) INTEGRATED ASSESSMENT MODELLING (IAM) COMMUNITY, (C) IMPACTS ADAPTATION AND VULNERABILITY (IAV) COMMUNITY, AND (D) POLICY MAKERS.

(A) Sea ice is both an integrator and a driver of changes in the climate system. The SIMIP protocol will allow the climate modeling community to understand and to compare the underlying budgets and to hence quantify the role of sea ice for any given experimental setup.

(B) Changes in the polar sea-ice cover, in particular in its spatial coverage, are among the most directly observable changes that occur in the Earth`s climate system. By allowing for a better understanding of the ongoing changes through the protocol defined here, the IAM community will be able to better estimate the reliability of their models during a period of already observed, significant changes in a specific climate variable.

(C) Since changes in the sea-ice cover are already ongoing, understanding these changes through the protocol defined here and assessing the reliability of modeled changes allows for a direct assessment of the quality of IAV models.

(D) Changes in sea ice are one of the most direct measures of ongoing changes in the Earth`s climate system. Earth System Models´ capability to simulate these changes is a key aspect to underpin the models´ credibility for policy makers. Hence, understanding any mismatch between models and observations will be central for understanding the robustness of these simulations for policy decisions. Such understanding will be possible for the modeled sea-ice cover through the protocol that we suggest here.

LIST OF OUTPUT AND PROCESS DIAGNOSTICS FOR THE CMIP DECK/CMIP6 DATA REQUEST**

WHETHER THE VARIABLE SHOULD BE COLLECTED FOR ALL CMIP6 EXPERIMENTS, OR ONLY SOME SPECIFIED SUBSET AND WHETHER THE OUTPUT IS NEEDED FROM THE ENTIRE LENGTH OF EACH EXPERIMENT OR SOME SHORTER PERIOD OR PERIODS;

The variables should be collected for all CMIP6 experiments when possible, ranked by priority if the full set cannot be provided. In addition to the normal monthly data for most variables, we will define certain short periods where daily data should be saved for all variables, in order to allow for a more in-depth analysis.

WHETHER THE OUTPUT MIGHT ONLY BE RELEVANT IF CERTAIN COMPONENTS OR DIAGNOSTIC TOOLS ARE USED INTERACTIVELY (E.G. INTERACTIVE CARBON CYCLE OR ATMOSPHERIC CHEMISTRY, OR ONLY IF THE COSP SIMULATOR HAS BEEN INSTALLED);

The variables of priority 1 are always relevant. Some variables of lower priority (ice-thickness distribution, tracer transport, melt-pond coverage) are only relevant if respective model components are being used

WHETHER THIS VARIABLE IS OF INTEREST TO DOWNSTREAM USERS (SUCH AS IMPACTS RESEARCHERS, WG2 USERS) OR WHETHER ITS PRINCIPAL PURPOSE IS FOR UNDERSTANDING AND ANALYSIS OF THE CLIMATE SYSTEM ITSELF. BE AS SPECIFIC AS POSSIBLE IN IDENTIFYING WHY THE VARIABLE IS NEEDED.

The principle purpose of the variable request is to analyze the climate system.

WHETHER THE VARIABLES CAN BE REGRIDDED TO A COMMON GRID, OR WHETHER THERE IS ESSENTIAL INFORMATION THAT WOULD BE COMPROMISED BY DOING THIS;

Variables should not be regridded, since this will not allow an a posteriori closure of any budget. However, we provide a short list of very basic variables that should additionally be provided on a common grid or as integrated quantity (i.e., hemispheric sea ice extent, area, and volume) to simplify a superficial analysis of model output

THE RELATIVE IMPORTANCE OF THE VARIOUS VARIABLES REQUESTED (INDICATED BY A TIERED LISTING) IS REQUIRED IF THE DATA REQUEST IS LARGE.

We group the variables in 3 priority levels.

ANY PROPOSED CONTRIBUTIONS AND RECOMMENDATIONS FOR**

MODEL DIAGNOSTICS AND PERFORMANCE METRICS FOR MODEL EVALUATION;

OBSERVATIONS/REANALYSIS DATA PRODUCTS THAT COULD BE USED TO EVALUATE THE PROPOSED EXPERIMENTS. INDICATE WHETHER THESE ARE AVAILABLE IN THE OBS4MIPS/ANA4MIPS DATABASE OR IF THERE ARE PLANS TO INCLUDE THEM;

TOOLS, CODE OR SCRIPTS FOR MODEL BENCHMARKING AND EVALUATION IN OPEN SOURCE LANGUAGES (E.G., PYTHON, NCL, R).

These usually depend on the overarching experiments that save sea-ice variables as part of their output. For the last few decades, satellite observations with varying accuracy are available of sea-ice drift, sea-ice concentration, sea-ice thickness, sea-ice age, sea-ice area, sea-ice volume and sea-ice extent. We will provide guidance to using these products in our variable description.

ANY PROPOSED CHANGES FROM CMIP5 IN NETCDF METADATA (CONTROLLED VOCABULARIES), FILE NAMES, AND DATA ARCHIVE (ESGF) SEARCH TERMS

Updated list of variables, saved in netCDF4 format.

EXPLANATION OF ANY PROPOSED CHANGES (RELATIVE TO CMIP5) THAT WILL BE REQUIRED IN CF, CMOR, AND/OR ESGF

Updated list of variables, saved in netCDF4 format

June 30th, 2015

PROPOSAL FOR A VIACS ADVISORY BOARD FOR CMIP6

- Name of Proposed Activity*: VIACS Advisory Board
- Co-chairs of MIP (including email-addresses)*: Alex Ruane (alexander.c.ruane@nasa.gov), Claas Teichmann (claas.teichmann@hzg.de)
- Proposed Members of the Advisory Board (which serves as a CMIP6 User Group and Scientific Steering Committee)*: Leaders of major Vulnerability, Impacts and Adaptation Sectors (potentially including: Jean Palutikof, Rob Swart, Dennis Lettenmaier, Dennis Ojima, Jerry Melillo, Almut Arneth, Shari Kovats, John Porter, John Shellenhuber/Katja Frieler, Richard Moss, Nigel Arnell, Tim Carter, Linda Mearns, Martin Parry, Claas Teichmann), with balance to ensure representation of regions and key international partners (e.g., TGICA, ICONICS, WGRC, PROVIA, etc.). Upon endorsement by CMIP members of the Advisory Board will be recruited as soon as possible.
- Link to website (if available)*: A website for the VIACS Advisory Board will be created at http://www.unep.org/provia/ including a description of the Board and links to CMIP6, data holdings at the Earth Systems Grid, and related online metrics and visualizations.
- Goal of the Activity and a brief overview*:

  Overview:

  To help form a more coherent and productive interaction between the climate modelers in CMIP6 and the Vulnerability, Impacts, Adaptation and Climate Services (VIACS) community, as well as contributing to the design of CMIP6-endorsed MIPs and online analysis capabilities that would enhance the benefit of CMIP simulations, we propose the creation of a VIACS Advisory Board for CMIP6. This will be anchored in the Programme of Research on Climate Change Vulnerability, Impacts, and Adaptation (PROVIA) and include direct participation of representatives from numerous regions, impacts sectors, and prominent international groups. The VIACS Advisory Board would not propose new climate model experiments, but would serve as an advisory body for planning and evaluation of existing CMIP6 MIPs, related outputs, and online metrics intended for use by the VIACS community. Climate modeling groups that are interested in building stronger engagement with the climate change applications community are encouraged to interact with the VIACS Advisory Board through CMIP6.

  The VIACS Advisory Board would facilitate efforts to address all three key science questions of CMIP6. The VIACS community has an acute interest in and high reliance on the best possible information about how the Earth System (in particular the impacted elements crucial to societal sectors) will respond to forcing, how model biases potentially influence decision-making in impacted sectors, and how climate variability, predictability, and uncertainty may be handled in preparing climate change adaptation and mitigation strategies that benefit impacted sectors. The VIACS Advisory Board will provide input in particular on the development and application of the representative concentration pathways and shared socioeconomic pathways (RCPs and SSPs, e.g., interacting with ScenarioMIP) which, along with the historical period DECK simulations, are the basis for most state-of-the-art projections of climate changes utilized in VIACS assessments (e.g., IPCC Working Group II).

The VIACS Advisory Board will also allow for closer engagement by the CMIP modeling groups as climate services are being established worldwide. Engagement with the CMIP modeling groups will help ensure that model output fits the climate service application needs, and also allows the modeling groups to provide synthesized input into this development process. Here, a close connection to the ScenarioMIP and CORDEX is needed. The VIACS Advisory Board will advise on the establishment of common evaluation concepts for global and regional climate data, best practices for the creation of individual climate service products. Another goal of the VIACS AB is to advise on the ways that climate services present information (e.g., vocabulary, uncertainties, information content, product consistency, the delivery and perception of messages). This can also benefit from networks of social science within the VIACS community.

Background on PROVIA:

The Global Programme of Research on Climate Change Vulnerability, Impacts and Adaptation (PROVIA) represents an interface between the research community and decision makers and other stakeholders for promoting improved policy-relevant research on vulnerability, impacts and adaptation (VIA), allowing scientists to coordinate and facilitate the dissemination and practical application of their research results. PROVIA is recognized within the World Climate Programme as offering a meeting point between researchers, stakeholders, and decision-makers interested in VIACS that can play the vital role of representing the perspectives of this highly diverse, transdisciplinary community. PROVIA helps international communities share practical experiences and research findings by improving the availability and accessibility of knowledge to the people that need it most. PROVIA aims to do so together with collaborative partners, knowledge networks, and the larger VIACS community, by identifying research needs and gaps, helping the scientific community to mobilize and communicate the growing knowledge based on VIACS assessment so that governments and other key stakeholders are able to consider this knowledge in their decision making processes.

Participants Representing Major Partners:

The VIACS Advisory Board will include representatives of regions, sectors, and major international groups. The participation of women and developing country representatives is crucial to providing an authentic VIACS community viewpoint, and the recruitment of Board members will seek a balance of regional perspectives. The VIACS Advisory Board would include leaders of established VIACS projects (e.g., AgMIP, WaterMIP, ISI-MIP), as well as senior scientists in impacts sectors (e.g., Agriculture, Urban, Biomes/Ecology, Forestry, Oceans/Fisheries, Coastal, Water Resources, Health, Economics, Energy, Infrastructure/Transportation) who have established credibility and networks within their impacts community as well as the wider climate change community. Members of PROVIA, the IPCC Task Group on Data and Scenario Support for Impact and Climate Analysis (TGICA), the World Climate Research Programme (WCRP) Working Group on

Regional Climate (WGRC), the International Committee On New Integrated Climate change assessment Scenarios (ICONICS), and the Climate Services community have contributed to the design of the VIACS Advisory Board and will help establish strong connections with these major international groups by serving on the Board.

Role of Board Members:

Board Members will serve two-year terms with rotating chairs to ensure new perspectives and a reasonable time commitment. Members of the VIACS Advisory Board will have a mandate to coordinate with other experts within their region/sector/group to provide community-based guidance that can be integrated at the VIACS Advisory Board level and then presented to CMIP6. The VIACS Advisory Board members would survey their respective communities, coordinate activities, and provide comprehensive feedback for CMIP6 to consider in designing and prioritizing scenarios and metrics for analysis and benchmarking that would be relevant for VIACS. The Board will help guide the development of online metrics and visualizations that will appeal to the VIACS community or researchers, stakeholders, and decision-makers. These include sector-specific indices (e.g., heat damage degree days for ecosystems, consecutive dry days for agriculture and water resources) and requirements for documentation and online guidance that will facilitate the use of CMIP6 products by the lay public. The Board would also continue the interactions already underway that bring the VIACS perspective into the development of new CMIP6-endorsed MIPs, and advise on better integration of the VIACS community model processes and results into the Earth Systems Grid where CMIP6 outputs are archived. The VIACS Advisory Board will also help raise awareness and establish good practices within the VIACS community.

- An overview of the proposed experiments*: The VIACS Advisory Board would not propose new experiments, but would serve as a Diagnostic activity for planning and evaluation of existing CMIP6 experiments. In this sense the VIACS Advisory Board will not produce a set of experiments that GCM groups can sign on to produce; instead, the Board will provide a forum for coordination interaction with the VIACS communities. GCM groups can participate in VIACS activities by expressing an interest in engaging with these climate change applications communities, for example helping to co-develop more relevant variable and experiment design as well as producing societally-relevant metrics and visualizations as CMIP6 products. We have thus adjusted the below sections to better illustrate the design and outcomes of this activity.
- An overview of the proposed evaluation/analysis of the CMIP DECK and CMIP6 experiments*: The VIACS Advisory Board will provide a forum for engagement between the GCM groups and the VIACS community using CMIP6 experiment results for the calculation of indices, as input to impact models, or directly for individual climate service related products (e.g., vulnerability maps). A common evaluation of the CMIP DECK historical and CMIP6 experiments is planned in order to investigate whether the models provide the data in a quality needed for climate services. Based on the outcome of this analysis and on projected climate changes, the VIACS Advisory Board will advise on the development of a common strategy to include CMIP6 projections in climate service products. The Board will help guide the development of online metrics and visualizations that will appeal to the VIACS community of researchers, stakeholders, and decision-makers. These include sector-specific

indices (e.g., heat damage degree days for ecosystems, consecutive dry days for agriculture and water resources) and requirements for documentation and online guidance that will facilitate the use of CMIP6 products by the lay public. The VIACS Advisory Board will also help evaluate and transfer good practices in CMIP output application within the VIACS community.

- Proposed timing*: Members of the VIACS Advisory Board and corresponding points of contact at CMIP6 will be recruited upon approval as a formal element of CMIP6. It is expected that the Board would be able convene for the first time about one month later, and would serve (through rotating terms) until the close of CMIP6 with the potential to continue into further CMIP activities.

For each proposed activity to be included in CMIP6**

- the activity design:

  The VIACS Advisory Board is designed to be a bridge between the VIACS community (~ IPCC Working Group II) and the climate modeling community (~ IPCC Working Group I), as well as to further interaction between the VIACS community and the integrated assessment modeling community (~ IPCC Working Group III). As such, the Board contributes by establishing regular communication between representatives of the CMIP6 MIPs and the VIACS community.

  Each consultation of the VIACS Advisory Board will comprise three steps. First, the co-Chairs and Board Coordinators will reach out to CMIP6 representatives to solicit input, requests, or questions to pose to the VIACS Advisory Board, either via email or teleconference. Second, the VIACS Advisory Board will hold a teleconference to discuss CMIP6 concerns and raise concerns from the VIACS community. Finally, a brief summary of the CMIP6/VIACS community interactions and action items identified by the Board will be prepared and shared with CMIP6 points of contact and all VIACS Advisory Board members. The VIACS Advisory Board would most often consult on a quarterly basis, however in the early stages of the project it is expected that it will be required to consult monthly to respond to CMIP6 design questions.

  The VIACS Advisory Board will also be active in the periods between teleconferences. Activity may include the solicitation of further input from communities represented by Board members (e.g., surveying members of AgMIP on an issue related to agriculture). Representatives of the VIACS Advisory Board will also participate in major CMIP6 meetings to give voice to the VIACS perspective. Although the Board will be tasked with providing feedback and ideas regarding the use of CMIP6 outputs for VIACS assessments, the assessments themselves are beyond the mandate of the VIACS Advisory Board itself.

- the science question and/or gap being addressed with this activity:

  The VIACS Advisory Board will help form a two-way bridge between CMIP and the users of its outputs in the VIACS community. The VIACS Advisory Board will provide inputs from the VIACS community on experiment and data design for CMIP6, guidelines for good practices in the use of CMIP6 outputs in VIACS assessments, and advice on the development of online metrics and visualizations for the Earth Systems Grid.

- possible synergies with other MIPs:
- The VIACS Advisory Board will provide VIACS perspective to MIPs with societal implications, for example including the development of RCPs and SSPs with ScenarioMIP, the use of ecosystem and agricultural models in conjunction with LUMIP, the health impacts of pollution policies in AerChemMIP, the role of water resource management in LandMIP.

- potential benefits of the activity to
    - (A) climate modeling community: The VIACS Advisory Board can advise on the most important climate and related variables to be requested from climate modelers, including downscaled information, for use in VIACS analysis. The Board can improve the relevance of climate model outputs to society through the development of more creative, robust, and efficient applications of climate model outputs. The Board will also facilitate dissemination of important scientific findings and caveats that need to be recognized in the design and communication of climate impact assessments.

    - (B) Integrated Assessment Modelling (IAM) community:
      The VIACS Advisory Board can advise on the most important socioeconomic and related quantitative variables to be requested from global IAMs that are consistent with climate projections generated in the CMIP6 process. The Board will provide important feedback on the implications of various policies and economic trajectories projected by the IAM community, potentially leading to shifts in the magnitude of feedbacks or the extent of plausible outcomes (e.g. land use, water resource availability, agricultural prices).

    - (C) Impacts, Adaptation, and Vulnerability (IAV; or VIA) community:
      The VIACS Advisory Board will seek to enhance significantly the level of communication between CMIP and the VIACS community, with mutual benefits. In particular, the Board will communicate and disseminate information to the VIACS community regarding access to, and understanding of, key climate model and related scenario outputs for VIACS research and wider societal applications. The Board will also help improve linkages between the activities of IPCC Working Group II and those of Working Groups I and III.

    - (D) policy makers:
      The VIACS Advisory Board will help CMIP6 incorporate the experience of the VIACS community interactions with policy makers around the world, leading to online metrics tailored toward policy makers and a greater

translation of climate model output toward social outcomes that are at the heart of policy maker interests.

- If possible, a prioritization of the suggested experiments, including any rationale**:

  Of particular interest for analysis and application will be the outputs from the historical DECK experiment (1900-present) and RCPs and associated SSPs (present-2100) developed within ScenarioMIP. Together, output from these experiments form the basis for most state-of-the-art projections of climate changes utilized in VIACS assessments. Additional interest also exists for simulation output from MIPs including DCPP (present-2040), LUMIP, AerChemMIP, and GEOMIP as well as interaction with Diagnostic MIPS such as CORDEX and GDDEX. The VIACS Advisory Board has already provided CMIP6 with a summary of variable and experiment priorities from a wide sample of leading VIACS researchers, projects, and programs, finding a large diversity in needs and priorities but allowing the GCM groups to determine how their efforts can lead to the most societally-relevant applications.

- All model output archived by CMIP6-Endorsed MIPs is expected to be made available under the same terms as CMIP output. Most modeling groups currently release their CMIP data for unrestricted use. If you object to open access to the output from your experiments, please explain the rationale.**: The VIACS Advisory Board will support open access to data and metrics created under its auspices.

  - The Activity addresses at least one of the key science questions of CMIP6: The VIACS Advisory Board would facilitate efforts to address all three key science questions of CMIP6. The VIACS community would be better able to determine how the Earth System (in particular the impacted elements relevant to society) will respond to forcing, how model biases potentially influence decision-making in impacted sectors, and how climate variability, predictability, and uncertainty may be handled in preparing climate change adaptation and mitigation strategies that benefit impacted sectors.
  - A sufficient number of modeling groups have agreed to participate in the MIP: There are a large number of impacts modeling groups across various sectors that will use CMIP6 outputs and potentially guidance from the VIACS Advisory Board. These include AgMIP, WaterMIP, ISI-MIP, and other community projects as well as smaller modeling groups not necessarily attached to large projects. It is the participation of these groups that will ensure the VIACS Advisory Board will make a substantial contribution toward CMIP output applications.
  - The MIP builds on the shared *CMIP DECK* experiments: The MIP will build in particular on the historical 20th Century simulations in the *DECK*, which will also form a basis for comparison against future scenarios and experiments from other MIPs.
  - A commitment to contribute to the creation of the CMIP6 data request and to analyze the data: The VIACS Advisory Board will work with CMIP6 to help identify and create metrics and visualizations of relevance to the VIACS community. More time is needed to coordinate and understand contributions to the CMIP6 data request from various communities already engaged by the preliminary organization of the VIACS Advisory

Board. Activity has been reported by the Climate Services Community, TGICA, AgMIP, and WGRC (which has conducted an extensive survey and is currently analyzing results).

- A commitment to identify observations needed for model evaluation and improved process understanding, and to contribute directly or indirectly to making such datasets available as part of obs4MIPs: The VIACS Advisory Board will collect information about the observational datasets utilized by various VIACS sectors and encourage the addition of those datasets to obs4MIPs to improve model evaluation and process understanding and encourage the addition of those datasets to obs4MIPs to improve model evaluation and process understanding.
- The proposed activity is of central importance to CMIP6: The VIACS Advisory Board will enhance the relevance of CMIP6 to society through all impact sectors.
- The proposed activity has been run at least by two modeling groups already: The VIACS Advisory Board will survey the best practices of impacts modelers from groups like AgMIP, ISI-MIP, WATERMIP, and others with a long history of VIACS contributions including the application of CMIP results.
- The proposed activity is useful in a multi-model context and to a number of climate researchers: The VIACS Advisory Board will encourage the use of multi-model ensembles both in the driving climate data and in the impacts models utilized. The Board will expedite the transfer of knowledge and practices from the climate modeling community's long use of ensemble approaches into the VIACS community which has only emphasized ensemble approaches in recent years.
- A commitment to scientifically analyze, evaluate and exploit the proposed experiment: The VIACS Advisory Board will enable a large number of researchers, stakeholders, decision-makers, and policy-makers to better integrate climate information into climate impact assessments across a number of sectors, with results also feeding back into the design and implications of climate modeling experiments. Initial discussions and planning has revealed wide support for an advisory board for CMIP6 from the VIACS community.