The PMIP4 contribution to CMIP6 – Part 1: Overview and over-arching analyses plan*

Masa Kageyama [1], Pascale Braconnot[1], Sandy P. Harrison[2], Alan M. Haywood[3], Johann H. Jungclaus[4], Bette L. Otto-Bliesner[5], Jean-Yves Peterschmitt[1], Ayako Abe-Ouchi[6,7], Samuel Albani[8], Patrick J. Bartlein[9], Chris Brierley[10], Michel Crucifix[11], Aisling Dolan[3], Laura Fernandez-Donado[12], Hubertus Fischer[13], Peter O. Hopcroft[14], Ruza F. Ivanovic[3], Fabrice Lambert[15], Daniel J. Lunt[14], Natalie M. Mahowald[16], W. Richard Peltier[17], Steven J. Phipps[18], Didier M. Roche[1,19], Gavin A. Schmidt[20], Lev Tarasov[21], Paul J. Valdes[14], Qiong Zhang[22], Tianjun Zhou[23]

[1]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France
[2]Centre for Past Climate Change and School of Archaeology, Geography and Environmental Science (SAGES) University of Reading, Whiteknights, RG6 6AH, Reading, United Kingdom
[3]School of Earth and Environment, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, United Kingdom
[4]Max Planck Institute for Meteorology, Bundesstrasse 53, 20146 Hamburg, Germany
[5]National Center for Atmospheric Research, 1850 Table Mesa Drive, Boulder, Colorado 80305, United States of America
[6]Atmosphere Ocean Research Institute, University of Tokyo, 5-1-5, Kashiwanoha, Kashiwa-shi, Chiba 277-8564, Japan
[7]Japan Agency for Marine-Earth Science and Technology, 3173-25 Showamachi, Kanazawa, Yokohama, Kanagawa, 236-0001, Japan
[8]Institute for Geophysics and Meteorology, University of Cologne, Cologne, Germany
[9]Department of Geography, University of Oregon, Eugene, OR 97403-1251, United States of America
[10]University College London, Department of Geography, WC1E 6BT, United Kingdom
[11]Université catholique de Louvain, Earth and Life Institute, Louvain-la-Neuve, Belgium
[12]Dpto. Física de la Tierra, Astronomía y Astrofísica II, Instituto de Geociencias (CSIC-UCM), Universidad Complutense de Madrid, Spain
[13]Climate and Environmental Physics, Physics Institute & Oeschger Centre for Climate Change Research, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland
[14]School of Geographical Sciences, University of Bristol, Bristol, United Kingdom
[15]Catholic University of Chile, Department of Physical Geography, Santiago, Chile
[16]Department of Earth and Atmospheric Sciences, Bradfield 1112, Cornell University, Ithaca, NY 14850, United States of America
[17]Department of Physics, University of Toronto, 60 St. George Street, Toronto, Ontario M5S 1A7, Canada
[18]Institute for Marine and Antarctic Studies, University of Tasmania, Private Bag 129, Hobart, TAS 7001, Australia
[19]Earth and Climate Cluster, Faculty of Earth and Life Sciences, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
[20]NASA Goddard Institute for Space Studies and Center for Climate Systems Research, Columbia University 2880 Broadway, New York, NY 10025, United States of America
[21]Department of Physics and Physical Oceanography, Memorial University of Newfoundland and Labrador, St. John's, NL, A1B 3X7, Canada
[22]Department of Physical Geography, Stockholm University, Stockholm, Sweden
[23]LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, P.O. Box 9804, Beijing 100029, China

Correspondence to: Masa Kageyama (Masa.Kageyama@lsce.ipsl.fr)

* This paper is the first of a series of 4 GMD papers on the PMIP4-CMIP6 experiments. Part 2 (Otto-Bliesner et al. 2017) gives the details about the two PMIP4-CMIP6 interglacial experiments, Part 3 (Jungclaus et al., 2017) about the last millennium experiment, and Part 4 (Kageyama et al., 2017) about the Last Glacial Maximum experiment. The mid Pliocene Warm Period experiment is part of the Pliocene Model Intercomparison Project (PlioMIP) - Phase 2, detailed in Haywood et al. (2016).

Abstract.

The goal of the Paleoclimate Modelling Intercomparison Project (PMIP) is to understand the response of the climate system to different climate forcings and feedbacks. Through comparison with observations of the environmental impact of these climate changes, or with climate reconstructions based on physical, chemical or biological records, PMIP also addresses the issue of how well state-of-the-art numerical models simulate climate change. Paleoclimate states can be radically different from those of the recent past documented by the instrumental record, and thus provide an out-of-sample test of the models used for future climate projections and a way to assess whether they have the correct sensitivity to forcings and feedbacks. Five different periods have been designed to contribute to the objectives of the sixth phase of the Coupled Model Intercomparison Project (CMIP6): the millennium prior to the industrial epoch (CMIP6 name: *past1000*), the mid-Holocene, 6,000 years ago (*midHolocene*); the Last Glacial Maximum, 21,000 years ago (*lgm*); the Last Interglacial, 127,000 years ago (*lig127k*) and mPWP, the mid-Pliocene Warm Period, 3.2 million years ago (*midPliocene-eoi400*). These climatic periods are well documented by paleoclimatic and paleoenvironmental records, with climate and environmental changes relevant for the study and projections of future climate changes. This manuscript describes the motivation for the choice of these periods and the design of the numerical experiments and database requests, with a focus on their novel features compared to the experiments performed in previous phases of PMIP and CMIP. It also outlines the analysis plan that takes advantage of the individual periods, the comparisons of the results across periods and across CMIP6 in collaboration with other MIPs.

1 Introduction

Instrumental meteorological and oceanographic data show that the Earth has undergone a global warming of ~0.85°C since the beginning of the industrial revolution (Hartmann et al., 2013), largely in response to the increase of atmospheric greenhouse gases. Concentrations of atmospheric greenhouse gases are projected to increase significantly during the 21st century, reaching levels well outside the range of recent millennia. Thus, in making future projections, models are operating well outside the current conditions, for which they have been validated, and that do not provide a full sampling of how climate responds to various external factors. Paleoclimate states radically different from those of the recent past provide an out-of-sample test of the models and a way to assess whether they have the correct sensitivity to forcings and feedbacks.

We have to look back several million years to find a period of Earth's history when atmospheric CO_2 concentrations were similar to the present day (the mid-Pliocene Warm Period, ca. 3.3 to 3 million years ago) and several tens of million years (e.g. the early Eocene, ~55 to 50 million years ago) to find concentrations similar to those possible by the end of this century under current emissions trajectories. Although these periods are not perfectly analogous to the future (their geographies are different from the present one), they can offer key insights into climate processes that operate in a higher CO_2, warmer world (e.g. Lunt et al., 2010, 2012; Caballero and Huber, 2010). During the Quaternary (2.58 million years ago to present), the Earth's geography was very similar to today and the main external factors driving climatic changes are the astronomical parameters, which determine the seasonal and latitudinal distribution of incoming solar energy. Changes in greenhouse gas concentrations and in ice sheets effectively acted as additional forcing factors on the dynamics of the atmosphere and the ocean. In addition, rapid climate transitions, on human-relevant timescales (decades to centuries), have been documented for this most recent period (e.g. Marcott et al., 2014; Steffensen et al., 2008)

By combining several past periods, the credibility of climate projections can be assessed using information about longer-term paleoclimate changes that are as large as the anticipated future change. Replicating the totality of past climate changes with state-of-the-art climate models, driven by appropriate forcings (e.g. insolation, atmospheric composition) and boundary conditions (e.g. ice sheets), is a challenge (Braconnot et al., 2012; Harrison et al., 2015). It is challenging, for example, to represent the correct amplitude of past climate changes such as glacial-interglacial temperature differences (e.g. the temperatures at the Last Glacial Maximum, ~21,000 years ago, vs. the pre-industrial temperatures, cf. Harrison et al., 2014) or the correct spatial patterns such as the northward extension of the African monsoon during the mid-Holocene, ~ 6,000 years ago (Perez-Sanz et al., 2014). Interpreting paleo-environmental data can also be challenging, in particular if trying to disentangle the relationships between changes in large-scale atmospheric or oceanic circulation, broad-scale regional climates and local environmental responses to these changes during climate periods where the relative importance of various climate feedbacks might be different from today. This challenge is paralleled by concerns about future local or regional climate changes and their impact on the environment. Modelling paleoclimates is therefore a means to understand past climate and environmental changes better, using physically based tools, as well as a means to evaluate model skill in forecasting the responses to major drivers.

These challenges are at the heart of the Paleoclimate Modeling Intercomparison project (PMIP) and the new PMIP4 set of simulations (Otto-Bliesner et al., 2017, Jungclaus et al., 2017, Kageyama et al., 2017, Haywood et al., 2016) proposed has the ambition to tackle them. PMIP was established in the 1990s (Joussaume and Taylor, 1995) in order to understand the mechanisms of past climate changes, in particular the role of the different climate feedbacks, and to evaluate how well the climate models used for climate projections simulate well-documented climate states outside the range of present and recent climate variability. Since the beginning, PMIP has also closely followed developments in climate modelling, in parallel to the Atmospheric Model Intercomparison Project (AMIP) and the Coupled Model Intercomparison project (CMIP). Paleoclimate experiments for the Last Glacial Maximum, the mid Holocene, 6000 year ago and the last millennium were formally included in CMIP during its fifth phase (CMIP5, Taylor et al., 2012), equivalent to the third phase of PMIP (PMIP3, Braconnot et al, 2012). It has then become possible to compare the mechanisms causing past and future climate changes in a rigorous way and evaluate the models used for projections under climate states very different from the present one (e.g. Harrison et al., 2014, Harrison et al., 2015). A measure of the success of PMIP3 is provided by the number of participating modelling groups (more than 20) and the prominent role of PMIP results in the fifth IPCC assessment report (Masson-Delmotte et al., 2013; Flato et al., 2013). PMIP3 also identified significant knowledge gaps and areas where progress is needed. PMIP4 has been designed to address these issues.

The five periods proposed for PMIP4-CMIP6 represent climate states with different greenhouse gas concentrations, astronomical parameters, ice sheet extents, and volcanic and solar activities (Figure 1, Table 1), consistent with the need to provide a large sample of the climate response to different forcings. The periods are, (abbreviated name is provided before the full name, name of corresponding PMIP4-CMIP6 experiment is given in italics within parentheses at the end of each line):

- LM, the millennium before the start of the industrial revolution, from 850 to 1849 CE (*past1000*)
- MH, the mid-Holocene, 6,000 years ago (*midHolocene*)
- LGM, the Last Glacial Maximum, 21,000 years ago (*lgm*)
- LIG, the Last Interglacial, 127,000 years ago (*lig127k*)
- mPWP, the mid-Pliocene Warm Period, 3.2 million years ago (*midPliocene-eoi400*)

In section 2, we give more background on the periods chosen for the CMIP6 experiments and the associated forcings and boundary conditions. The experimental set-up of the experiments is described in section 3, with reference to the appropriate manuscript where details and the additional sensitivity experiments considered in PMIP4 can be found. The analysis plan is outlined in Section 4. A short conclusion is given in section 5.

2. PMIP4 experiments for CMIP6

2.1 Five contrasted time periods to answer the CMIP6 questions

The five PMIP4-CMIP6 experiments have been chosen to best contribute to the CMIP6 key questions (Section 4). Two climatic periods have been major foci throughout PMIP's history: the mid-Holocene and the Last Glacial Maximum (Table 1). These two periods represent reference points for assessing the sensitivity of the

climate system to changes in atmospheric CO_2 concentration and orbitally-induced changes in tropical circulation and the monsoons (Braconnot et al., 2012; Harrison et al., 2015). They are considered as entry cards in the PMIP4-CMIP6 set of experiments, so that a sufficient number of simulations are available to be able to trace model evolution and ability to represent these climate states since the first phase of PMIP.

The transient simulations of the millennium prior to the industrial epoch (Schmidt et al., 2011, 2012, Junglaus et al., 2017) allow the study of the mechanisms of decadal to centennial climate variability (natural variability vs. impact of solar, volcanic and anthropogenic forcings, cf. Table 1) and were included in PMIP3-CMIP5. In addition, a number of other time periods were included in PMIP3, in particular the mid-Pliocene Warm Period

10   (mPWP, cf. the PlioMIP project, Haywood et al., 2010, 2011), and the last interglacial period (130,000 to 115,000 years before present, Lunt et al., 2013). The latter simulations were used to examine whether climate models could produce a rate of ice-sheet melting in agreement with a global sea level at least 5m higher than now (Masson-Delmotte et al., 2013; Dutton et al., 2015). Discussions on transient simulations of climate behaviour, focusing on the last interglacial period and the last deglaciation (Ivanovic et al., 2016) were also

15   initiated, as were simulations of climates of deeper times, in particular the early Eocene, ~50 million years ago (Lunt et al, 2012; Lunt et al, 2017). Questions on climate sensitivity and polar amplification and on the relationships between climate-ice-sheet system and sea level led us to propose, for PMIP4-CMIP6, two additional periods for which there is good data coverage, the possibility to design simple, but realistic, simulations, and large working groups interested in the analyses and the collaboration with other MIP

20   participating in CMIP6. These additional periods are the last interglacial, 127 000 years ago and the mid Pliocene Warm Period (Table 1).

The true power of PMIP is the connection to the environmental observations and climate reconstructions. Uncertainties in the paleoenvironmental observations, or perhaps more broadly in the climate inferences made

25   from those observations, are a key part of PMIP analyses, as is the structural uncertainty across the model responses. Both of these factors have been part of the PMIP approach from the beginning. Improved reconstructions, increased complexity and realism of climate simulations require putting more emphasis on the understanding of impacts of the uncertainties on the drivers themselves. For each of the selected periods, this encompasses time-uncertainty in the reconstructions (e.g. are all data synchronous? what date should be used to

30   compute the astronomical parameters to compare with available data?) as well as structural uncertainty in the boundary conditions applied (e.g. in the continental reconstructions, ice sheet height and extent, vegetation cover), and in the transient forcings (for instance in the last millennium simulations for solar, volcanic aerosol or land use/land cover change). Differences between plausible reconstructions of boundary conditions and forcings can impact the assessment of model skill. In these cases, we have included alternative forcings and boundary

35   conditions for the PMIP4-CMIP6 experiment or to be used in PMIP4 sensitivity experiments (Jungclaus et al., 2017; Otto-Bliesner et al., 2017; Kageyama et al., 2017; Haywood et al., 2016).

More background is provided for each of the PMIP4-CMIP periods in next sections. All the experiments can be run independently and have value for comparison to the CMIP6 DECK (Diagnostic, Evaluation and

40   Characterization of Klima) and historical experiments (Eyring et al., 2016). They are therefore all considered as

Tier 1 within CMIP6 (Table **1**). It is not mandatory for groups wishing to take part in PMIP4-CMIP6 to run all five PMIP4-CMIP6 experiments. It is however mandatory to run at least one of the two entry cards, i.e. the *midHolocene* or the *lgm*.

Figure 1: Context of the PMIP4 experiments (from left to right: mPWP, mid-Pliocene Warm Period; LIG, last interglacial; LGM, last glacial maximum; MH, mid-Holocene; LM, last millennium; H, CMIP6 historical simulation): (a)-(d) insolation anomalies (differences from 1950 CE), for July at 65°N, calculated using the programs of Laskar et al. (2004, panel (a)) and Berger (1978, panels (b)-(d)); (e) $\delta^{18}O$ (magenta, Lisiecki and Raymo, 2005, scale at left), and sea level (blue line, Rohling et al., 2014; blue shading, a density plot of eleven mid-Pliocene sea level estimates (Dowsett and Cronin 1990; Wardlaw and Quinn, 1991; Krantz, 1991; Raymo et al., 2009; Dwyer and Chandler, 2009; Naish and Wilson, 2009; Masson-Delmotte et al., 2013; Rohling et al., 2014; Dowsett et al., 2016) scale at right); (f) and (g) $\delta^{18}O$ (magenta, Lisiecki and Raymo, 2005, $\delta^{18}O$ scale at left), and sea level (blue dots, with light-blue 2.5, 25, 75 and 97.5 percentile bootstrap confidence intervals, Spratt and Lisiecki, 2015; blue rectangle, LIG high-stand range, Dutton et al., 2015; dark blue lines, Lambeck et al., 2014, sea-level scale at right on panel (g)), (h) sea level (Kopp, et al., 2016, scale at right); (i) CO_2 for the interval 3.0-3.3 Ma shown as a density plot of eight mid-Pliocene estimates (Raymo et al., 1996; Stap et al., 2016; Pagani et al., 2010; Seki et al., 2010; Tripati et al., 2009; Bartoli et al., 2011; Seki et al., 2010; Kurschner et al., 1996); (j) and (k) CO_2 measurements (Bereiter et al., 2015, scale at left); (l) CO_2 measurements (Schmidt et al, 2011, scale at right); (m) and (n) CH_4 measurements (Loulergue et al., 2008, scale at left); (o) CH_4 measurements (Schmidt et al, 2011, scale at right); (p) volcanic radiative forcing (Schmidt et al., 2012, scale at right); (q) total solar irradiance (Schmidt et al., 2012, scale at right).

Table 1: Characteristics, purpose and CMIP6 priority of the five PMIP4-CMIP6 experiments

2.2 PMIP4-CMIP6 entry cards: the mid-Holocene (*midHolocene*) and last glacial maximum (*lgm*)

The MH and LGM periods provide examples of strongly contrasting climate states (Figure 1, Table **1**). The MH provides an opportunity to examine the response to orbitally-induced changes in the seasonal and latitudinal distribution of insolation. It is a period during which the northern hemisphere was characterised by enhanced northern hemisphere summer monsoons, extra-tropical continental aridity and much warmer summers. The LGM provides an opportunity to examine the impact of changes in ice sheets and continental extent (which increases due to the drop in sea level) and of the decrease in atmospheric greenhouse gases on climate. The LGM is particularly relevant because the forcing and temperature response from the LGM to the Holocene was as large as to that projected from present to the end of the 21st century (Braconnot et al., 2012).

Evaluation of the PMIP3-CMIP5 MH and LGM experiments has demonstrated that climate models simulate changes in large-scale features of climate that are governed by the energy and water balance reasonably well (Harrison et al., 2014, 2015), including changes in land-sea contrast and high-latitude amplification of temperature changes (Izumi et al., 2013; Izumi et al., 2015). They also simulate the scaling of precipitation changes with respect to temperature changes at the hemispheric scale realistically (Li et al., 2013). These results confirm that the simulated relationships between large-scale patterns of temperature and precipitation change in future projections are credible (Harrison et al., 2015). However, the PMIP3-CMIP5 simulations of MH and LGM climates show only moderate skill in reproducing reconstructed patterns of climate change overall (Hargreaves et al., 2013; Hargreaves and Annan, 2014; Harrison et al., 2014; Harrison et al., 2015). This arises because of persistent problems in simulating regional climates (e.g. Mauri et al., 2014; Perez-Sanz et al., 2014; Harrison et al., 2015). State-of-the-art models cannot fully reproduce the northward penetration of the African monsoon in response to MH orbital forcing (Perez-Sanz et al., 2014; Pausata et al., 2016), for example. This discrepancy was already noted in PMIP1 (Joussaume et al, 1999). While this likely reflects inadequate representation of feedbacks, model biases could also contribute to this mismatch (e.g. Zheng and Braconnot,

2013). Systematic benchmarking of the PMIP3-CMIP5 MH and LGM also show that better performance in paleoclimate simulations is not consistently related to better performance under modern conditions (Harrison et al., 2015). Hence the ability to simulate modern climate regimes and processes does not guarantee that a model will be good at simulating climate changes, emphasizing the importance of testing models against the paleoclimate record to increase confidence in projections of future climate (Braconnot et al., 2012; Hargreaves and Annan, 2014; Schmidt et al., 2014).

For PMIP4-CMIP6, we have modified the experimental design of the *midHolocene* and *lgm* experiments with the aim of obtaining more realistic representations of these climates (Table 2, Otto-Bliesner et al., 2017 for *midHolocene* and Kageyama et al., 2017 for *lgm*). One of these modifications is the inclusion of changes in atmospheric dust loading, which can have a large effect on regional climate changes. Dust has now been implemented in many CMIP6 models, either by using models with an interactive representation of dust or by prescribing atmospheric dust content. In PMIP3, the *midHolocene* CO_2 concentration was prescribed to be the same as in the pre-industrial control simulation, because the focus was on testing the impact of the insolation forcing on meridional climate gradients and seasonality. Realistic values of the concentration of atmospheric CO_2 and other trace gases will be used in PMIP4-CMIP6 (Table 2). This will allow the *midHolocene* experiment to be used as the initial state for transient simulations of the late Holocene planned as part of PMIP4, and ensure consistency of forcing between the *midHolocene* PMIP4-CMIP6 snapshot experiment and the transient simulations (Otto-Bliesner et al., 2017).

The PMIP3-CMIP5 LGM experiments considered a single ice sheet reconstruction (Abe-Ouchi et al., 2015). However, there is uncertainty about the geometry of the ice sheets at the Last Glacial Maximum. Thus the protocol for the PMIP4-CMIP6 *lgm* simulations includes a choice between the old PMIP3 ice sheet (Abe-Ouchi et al., 2015) or one of two new 21ky BP reconstructions based on somewhat different approaches: ICE-6G_C (Argus et al., 2014; Peltier et al., 2015) and GLAC-1D (Tarasov et al., 2012; Briggs et al., 2014, Ivanovic et al., 2016). Groups wishing to use the *lgm* equilibrium experiment to initialise PMIP4 transient simulations of the last deglaciation (Ivanovic et al., 2016) must use either ICE-6G_C or GLAC-1D because these are consistent with the ice sheet and meltwater forcings provided for the PMIP4 transient experiments. The impact of these different ice-sheet forcings will be a focus for sensitivity experiments in PMIP4 (Kageyama et al., 2017). There are uncertainties regarding other boundary conditions for the *midHolocene* and *lgm* experiments, including dust and vegetation (section 3.4.1), and these will also be investigated as part of the analysis of the entry-card simulations.

2.2 The last millennium (*past1000*)

The millennium prior to the industrial era, 850-1849 CE, provides a well-documented (e.g. PAGES2k-PMIP3 group, 2015) period of multi-decadal to multi-centennial changes in climate, with contrasting periods such as the Medieval Climate Anomaly and the Little Ice Age. This interval was characterised by variations in solar, volcanic and orbital forcings (Figure 1), which acted under climatic background conditions not too different from today. This interval provides a context for earlier anthropogenic impacts (e.g. land-use changes) and the current warming due to increased atmospheric greenhouse gas concentrations. It also helps constrain the uncertainty in the future climate response to a sustained anthropogenic forcing.

The PMIP3-CMIP5 *past1000* simulation provided an assessment of climate variability on decadal and longer scales and information on predictability under forced and unforced conditions (Table 1). The importance of forced variability on multi-decadal to centennial time scales was highlighted by comparing spectra from *past1000* simulations with those from control experiments (Fernández-Donado et al., 2013). Other studies focused on the temperature difference between the warmest and coldest centennial or multi-centennial periods and the relation to changes in external forcing, in particular variations in solar irradiance (Fernández-Donado et al., 2013; Hind and Moberg, 2013). Single-model ensembles have provided improved understanding of the importance of internal versus forced variability and the individual forcings when compared to reconstructions at both global and regional scales (Man et al., 2012; Phipps et al., 2013; Schurer et al., 2014; Man et al., 2014; Man and Zhou, 2014; Otto-Bliesner et al., 2016). The *past1000* simulations show relatively good agreement with regional climate reconstructions for the northern hemisphere, but less agreement with southern hemisphere records. The simulations exhibit more regional coherence than shown by southern hemisphere records, though it is not clear whether this is due to deficiencies in the southern hemisphere records, poor representation of internal variability and/or an overestimation of the forced response in the simulations (PAGES2k-PMIP3 group, 2015).

The PMIP4-CMIP6 *past1000* simulation (Jungclaus et al, 2017) builds on the DECK experiments, in particular the pre-industrial control (*piControl*) simulation as an unforced reference, and the *historical* simulations (Eyring et al., 2016). Moreover, the *past1000* simulations provide initial conditions for *historical* simulations starting in the 19th century that are considered superior to the *piControl* state, as they include integrated information from the forcing history (e.g. large volcanic eruptions in the early 19th century). It is therefore mandatory to continue the *past1000* simulation into the historical period using the same forcing as for the standard CMIP6 historical simulations. The PMIP4-CMIP6 *past1000* simulation will use a new, more comprehensive reconstruction of volcanic forcing (Sigl et al., 2015) and an experimental protocol that ensures a more continuous transition from the pre-industrial past to the future. The final choices result from strong interactions with the groups producing the different forcing fields for the historical simulations (Jungclaus et al, 2017). It is expected that more groups will be able to provide ensembles of past1000 runs and higher-resolution simulations, which will allow the analysis of a greater range of regional processes, such as the role of storm-tracks and blocking on regional precipitation.

2.3 The Last Interglacial (*lig127k*)

The Last Interglacial (ca 130-115 kyr BP) was characterised by a northern hemisphere insolation seasonal cycle even larger than for the mid-Holocene (Figure 1, Table **1**). This resulted in a strong amplification of high-latitude temperatures and reduced Arctic sea ice. Global sea level was at least 5 m higher than now for at least several thousand years (Masson-Delmotte et al., 2013; Dutton et al., 2015). Both the Greenland and Antarctic ice sheets contributed to this sea level rise, making it an important period for testing our knowledge of climate-ice sheet interactions in warm climates. The availability of quantitative climate reconstructions for the Last Interglacial (e.g. Capron et al., 2014) makes it feasible to evaluate these simulations and assess regional climate changes.

Climate model simulations of the Last Interglacial, reviewed and assessed in the AR5 (Masson-Delmotte et al., 2013), varied in their forcings and were not necessarily made with the same model or at the same resolution as the CMIP5 future projections. There are large differences between simulated and reconstructed mean annual surface temperature anomalies compared to present, particularly for Greenland and the Southern Ocean, and in the temperature trends in transient experiments run for the whole interglacial (Bakker et al., 2013; Lunt et al., 2013). Part of this discrepancy stems from the fact that the climate reconstructions were of the local maximum interglacial warming, and this was not globally synchronous, an issue which is addressed in the PMIP4-CMIP6 protocol.

The PMIP4-CMIP6 *lig127k* experiment will help to determine the interplay of warmer atmospheric and oceanic temperatures, changed precipitation, and changed surface mass and energy balance on ice sheet thermodynamics and dynamics (Table **1**). The major changes in the experimental protocol for *lig127k*, compared to the pre-industrial DECK experiment, are changes in the astronomical parameters and greenhouse gas concentrations (Table 2; Otto-Bliesner et al., 2017). Meaningful analyses of these simulations are now possible because of the concerted effort to synchronise the chronologies of individual records and thus provide a spatial-temporal picture of last interglacial temperature change (Capron et al., 2014, 2017), and also to document the timing of the contributions of Greenland and Antarctica to the global sea level (Winsor et al., 2012; Steig et al., 2015). Regional responses of tropical hydroclimate and of polar sea ice can be assessed and compared to the *mid-Holocene*. Outputs from the *lig127k* experiment will be used by ISMIP6 to force stand-alone ice sheet experiments (*lastIntergacialforcedism*) in order to quantify the potential sea level change associated with this climate. The *lig127k* experiment will also be the starting point for a transient experiment covering the interglacial to be run within PMIP4.

2.4 The mid-Pliocene Warm Period (*midPliocene-eoi400*)

The mPWP experiment focuses on a specific interglacial, dated ~3.2 Ma before present, during the wider mid-Pliocene Warm Period (mPWP, ca. 3.3 to 3 million years ago). This was the last time in Earth history when atmospheric CO_2 concentrations approached current values (~400 ppmv) with a continental configuration similar to today (Figure 1, Table **1**). Vegetation reconstructions (Salzmann et al., 2008) indicate that the area of deserts was smaller than today and boreal forests were present in high northern latitude regions that are covered by tundra today. Climate model simulations from the PlioMIP project produce global mean surface air temperature anomalies ranging from +1.9 °C to +3.6 °C (relative to each model's pre-industrial control) and an enhanced hydrological cycle (Haywood et al., 2013) with strengthened monsoons (Zhang et al., 2013). These simulations also show 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). Model–data comparisons provide high confidence that mean surface temperature was warmer than pre-industrial (Dowsett et al., 2012; Haywood et al., 2013; Masson-Delmotte et al., 2013). However, as is the case for the Last Interglacial, the mid-Pliocene simulations were not always derived from the same model at the same resolution as the CMIP5 future projections.

The PMIP4-CMIP6 *midPliocene-eoi400* experiment (Haywood et al., 2016) is designed to understand the long-term response of the climate system to a near modern concentration of atmospheric CO_2 (longer term climate sensitivity or Earth System Sensitivity). It will also be used to address the response of ocean circulation, Arctic sea-ice and modes of climate variability (e.g. El Niño Southern Oscillation), as well as the global response in the hydrological cycle and regional changes in monsoon systems (Table **1**). The simulation has the potential to be informative about required emission reduction scenarios designed to prevent an increase in global annual mean temperatures by more than 2 °C after 2100 CE. Boundary conditions include modifications to global ice distributions, topography/bathymetry, vegetation and CO_2 (Table 2, Section 3) and are provided by the US Geological Survey Pliocene Research and Synoptic Mapping Project (PRISM4: Dowsett et al., 2016).

3. Experimental set up and model configuration

The forcings and boundary conditions for each PMIP4-CMIP6 paleoclimate simulation are summarised in Table 2. The complete justification of the experimental protocols and analysis plans are given in a series of companion papers: Otto-Bliesner et al. (2017) for the *midHolocene* and *lig127ka* experiments, Kageyama et al. (2017) for the *lgm*, Jungclaus et al. (2017) for the *past1000* and Haywood et al. (2016) for the *midPliocene-eoi400* experiment. These papers also explain how the boundary conditions for each period have been designed and constitute key references for the experimental protocol for each of the PMIP4-CMIP6 simulations. Here we provide guidelines that are common to all of the experiments, focusing on the implementation of the boundary conditions where there is a need to ensure consistency between CMIP6 and PMIP4 experiments.

3.1 Model version and set-up

The climate models taking part in CMIP6 are very diverse: some represent solely the physics of the climate system, some include the carbon cycle and other biogeochemical cycles, and some include interactive natural vegetation and/or interactive dust cycle/aerosols. It is mandatory that the model version used for the PMIP4-CMIP6 experiments is exactly the same as for the other CMIP6 experiments, in particular the DECK and *historical* simulations. Except for the *past1000* simulation, all the other PMIP4-CMIP6 simulations are equilibrium experiments, in which the boundary conditions and forcings are constant from one year to another. The experimental set-up for each simulation is based on the DECK pre-industrial control (*piControl*) experiment (Eyring et al., 2016), i.e. the *piControl* forcings and boundary conditions are modified to obtain the forcings and boundary conditions necessary for each PMIP4-CMIP6 paleoclimate experiment (Table 2). No additional interactive component (such as vegetation or dust) should be included in the model unless it is already included in the DECK version. Such changes would affect the global energetics (Braconnot and Kageyama, 2015) and therefore prevent rigorous analyses integrating across multiple time periods or between MIPs (sections 4.2 and 4.3).

Table 2: summary of changes in boundary conditions w.r.t. *piControl* for each PMIP4-CMIP6 experiment

3.2 Summary of the forcings and boundary conditions: greenhouse gases, insolation, ice sheets

For each experiment, the greenhouse gases should be modified from the DECK *piControl* experiment according to Table 2. Astronomical parameters have to be adjusted for all experiments but *midPliocene-eoi400*. Two

experiments, *lgm* and *midPliocene-eoi400* require modifying ice sheets (Figure 2), which also implies modifying the coastlines, the ocean bathymetry (if feasible for *midPliocene-eoi400*), the topography and land surface types over the continents and to ensure rivers reach the ocean in order to close the global fresh water budget. The ocean initial salinity should be adjusted for these ice volume changes and modelling groups should ensure that the total mass of the atmosphere remains the same in all experiments. The details of the implementation of the ice sheets can be found in Haywood et al. (2016) for *midPliocene-eoi400* and Kageyama et al. (2017) for *lgm*.

Figure 2**: Changes in boundary conditions related to changes in ice sheets for the *midPliocene-eoi400* (top) and *lgm* (middle: ICE-6G_C and bottom: GLAC-1D) experiments. Coastlines for paleo-period shown as brown contours. Ice sheet boundaries for each period shown as red contour. Bright shading: changes in altitude over regions covered by ice sheets during the considered paleo-period. Faded shading: changes in altitude over ice-free regions.**

3.3 Vegetation and land use

Paleoenvironmental records show that natural vegetation patterns during each of the PMIP4-CMIP6 period were different from today. However, in order to ensure comparability between past, present and future climate simulations, the PMIP4-CMIP6 paleoclimate simulations should follow the same protocol as the DECK and historical simulations. If the DECK and historical simulations use dynamic vegetation, then the PMIP4-CMIP6 paleoclimate simulations should also. If the DECK and historical simulations use prescribed vegetation, then the same vegetation should be prescribed in the PMIP4-CMIP6 paleoclimate simulations. One exception to this is the *midPliocene-eoi400* experiment, where models which prescribe vegetation in the DECK and historical simulations should prescribe the mid-Pliocene vegetation (Haywood et al., 2016). The other exception is for models including interactive dust cycle for the LGM, which should impose vegetation that allows dust emissions over LGM dust emission regions.

Simulations to examine the impact of vegetation changes are of interest and can be evaluated using paleoclimate data. These can be made using prescribed vegetation changes, by running a model such as BIOME4 (https://pmip2.lsce.ipsl.fr/) off line to compute vegetation patterns compatible with a past climate state, or by running additional simulations with a non-standard version of the model with dynamic vegetation. Sensitivity experiments such as these will be encouraged within PMIP4 but are not part of the PMIP4-CMIP6 experiments.

For the *past1000* simulation, land-use changes have to be implemented in the same manner as for the *historical* simulation, using the land-use forcing provided by the Land Use Model Intercomparison Project (Lawrence et al., 2016) and the CMIP6 Land Use Harmonization dataset LUH2 (https://cmip.ucar.edu/lumip; for details see Jungclaus et al., 2017). This data set provides a seamless transition between the pre-industrial millennium and the historical period. It is derived from the HYDE3.2 (Klein Goldewijk, 2016) estimates of the area of cropland, managed pasture, rangeland, urban, and irrigated land. Different crop types are treated separately and estimates of wood harvest are also provided. LUH2 includes improved updates of shifting cultivation rates, management information and new estimates of wood consumption.

3.4 Natural aerosols

3.4.1 Mineral Dust

Natural aerosols show large variations on glacial-interglacial time scales, with glacial climates having higher dust loadings than interglacial climates (Kohfeld and Harrison, 2001; Maher et al., 2010). Dust emissions from northern Africa were significantly reduced during the MH (McGee et al., 2013). As is the case with vegetation, the treatment of dust in the *midHolocene, lig127k* and *lgm* simulations should parallel the treatment in the *piControl*. However, some of the models in CMIP6 include representations of interactive dust. For those models, maps of soil erodibility that account for changes in the extension of possible dust sources, will be provided for the *midHolocene, lig127k* and *lgm* experiments. The maps are the same for the interglacial experiments. Dust anomalies/ratios compared to the pre-industrial background should be used, for consistency with the DECK *piControl* simulation. As there have been instances of runaway climate-vegetation-dust feedback, leading to unrealistically cold LGM climates (Hopcroft and Valdes, 2015a), it is advisable to test model behaviour with an atmosphere-only model before running the entire *lgm* simulation.

To allow experiments with prescribed dust changes, three-dimensional monthly climatologies of dust atmospheric mass concentrations will be provided for the *piControl*, *midHolocene*, and *lgm*. These are based on two different models (Albani et al., 2014, 2015, 2016; Hopcroft et al., 2015, Figure 4) and modelling groups are free to choose between these data sets. Additional dust-related fields (dust emission flux, dust load, dust aerosol optical thickness, short- and long-wave, surface and top of the atmosphere dust radiative forcing) are also available from these simulations. Implementation should follow the same procedure as for the *historical* run. The implementation for *lig127k* experiment should use the same data set as for the *midHolocene*. Since dust plays an important role in ocean biogeochemistry (e.g. Kohfeld et al., 2005), three dust maps are provided for the *lgm* run. Two of these are consistent with the climatologies of dust atmospheric mass concentrations; the other is primarily derived from paleoenvironmental observations (Lambert et al., 2015, Figure 3). The modelling groups should use consistent data sets for the atmosphere and the ocean biogeochemistry. The Lambert et al. (2015) data set can therefore be used for models that cannot include the changes in atmospheric dust according to the other two data sets.

Figure 3: Maps of dust deposition (g m-2 a-1) simulated with the Community Earth System Model for a. PI (Albani et al., 2016), b. MH (Albani et al., 2015), and c. LGM (Albani et al., 2014). Maps of dust deposition (g m-2 a-1) for the LGM d. simulated with the Hadley Centre Global Environment Model 2-Atmosphere (Hopcroft et al, 2015), and e. reconstructed from a global interpolation of paleodust data (Lambert et al., 2015).

3.4.2 Volcanoes and stratospheric aerosols

The *past1000* experiment is the only one that requires imposing changes in volcanic aerosols. Modelling groups using interactive aerosol modules and sulphur injections in their historical simulations should follow the same method for the *past1000* experiment and use sulphur injection estimates by Toohey and Sigl. (2017) directly. For the other models, estimates of aerosol radiative properties as a function of latitude, height, and wavelength should be calculated using the Easy Volcanic Aerosol (EVA) module (Toohey et al., 2016). There are uncertainties associated with this approach. Additional sensitivity studies allowing the assessment of the impacts of these uncertainties on the *past1000* simulations will be made as part of the PMIP4 past1000 Tier 2 experiments (see Jungclaus et al., 2017). The sulphur injection time series and the EVA software package are provided via the PMIP4 web page (https://pmip4.lsce.ipsl.fr/doku.php/exp_design:lm).

3.5 Solar irradiance

For the *past1000* experiment, new reconstructions of TSI and SSI are provided that are based on recent estimates of cosmogenic isotopes and improved irradiance models (see Jungclaus et al., 2017 for details). The forcing prescribed for the Tier 1 past1000 experiment is constructed using a ^{14}C based reconstruction (Usoskin et al., 2016) of yearly sunspot numbers and an updated version of the Vieira et al. (2011) irradiance model. To achieve a smooth transition to the industrial period for historical experiments (1850 – 2015 CE) that start from the end of the *past1000* simulations, the forcing is scaled to match the CMIP6 historical forcing (Matthes et al., 2017). Alternative forcing reconstructions, reflecting uncertainty in the cosmogenic isotopes and the methods used in solar irradiance models, are provided as a basis for additional Tier 2 experiments (Jungclaus et al., 2017).

3.6 Spin-up and duration of experiments

The data stored in the CMIP6 database should be representative of the equilibrium climates of the MH, LGM, LIG and mPWP periods, and of the transient evolution of climate between 850-1849 CE for the *past1000* simulations. Spin-up procedures will differ according to the model and type of simulation, but the spin up should be long enough to avoid significant drift in the analysed data. Initial conditions for the spin-up can be taken from an existing simulation. In previous phases of PMIP, we recommended that the model should be run until the absolute value of the trend in global mean sea-surface temperature is less than 0.05 K per century and the Atlantic Meridional Overturning Circulation (AMOC) is stable. A parallel requirement for carbon-cycle models and/or models with dynamic vegetation is that the 100-year average global carbon uptake or release by the biosphere is <0.01 Pg C per year. We require the groups to document this spin-up by saving a limited set of variables during this phase (Table 3). A minimum of 100 years output is required for the equilibrium simulations but, given the increasing interest in analysing multi-decadal variability (e.g. Wittenberg, 2009), modelling groups are encouraged to provide outputs for 500 years or more if possible.

3.7 Documentation

Detailed documentation of the PMIP4-CMIP6 simulations is required. This should include:

- a description of the model and its components;
- information about the boundary conditions used, particularly when alternatives are allowed (Table 2);
- information on the implementation of boundary conditions and forcings. Figures showing the land-sea mask, land-ice mask, and topography as implemented in a given model are useful for the *lgm* and *midPliocene-eoi400* experiments, while figures showing insolation are particularly important for the *midHolocene* and *lig127k* experiments. Check lists for the implementation of simulations are provided in the PMIP4 papers which give detailed information for each experiment (*midHolocene*: Otto-Bliesner et al., 2017; *lgm*: Kageyama et al., 2017; *past1000*: Jungclaus et al., 2017; *lig127k*: Otto-Bliesner et al., 2017; *midPliocene-eoi400*: Haywood et al., 2016);
- information about the initial conditions and spin-up technique used (cf. section 3.6). A measure of the changes in key variables (Table 3) should be provided in order to assess remaining drift.

Documentation should be provided via the ESDOC website and tools provided by CMIP6 (http://es-doc.org/) to facilitate communication with other CMIP6 MIPs. This documentation should also be provided for the PMIP4

website to facilitate linkages with non-CMIP6 simulations to be carried out in PMIP4. The PMIP4 special issue, shared between *Geoscientific Model Development* and *Climate of the Past*, provides a further opportunity for modelling groups to document specific aspects of their simulations.

4. Overview of analyses plan and links to the required output

The five PMIP4-CMIP6 experiments naturally address the key CMIP6 question "How does the Earth System respond to forcing?" (Eyring et al, 2016), for multiple forcings and in climates states very different from the current or historical climates. For each target period, comparison with environmental observations and climate reconstructions enable us to determine whether the modelled responses are realistic. PMIP also addresses key CMIP6 question 2 "What are the origins and consequences of systematic model biases?" PMIP simulations and data-model comparisons will show whether the biases in the present-day simulations are found in other climate states. More importantly, analyses of PMIP simulations will show whether present-day biases have an impact on the magnitude of simulated climate changes. Finally, PMIP is also relevant to CMIP6 question 3 "How can we assess future climate changes given climate variability, predictability and uncertainties in scenarios?" through examination of these questions for documented past climate states and via the use of the last millennium simulations as reference state for natural variability.

The community using PMIP simulations is very broad, from climate modellers and palaeoclimatologists to biologists studying changes in biodiversity and archaeologists studying potential impacts of past climate changes on human populations. Because of this, we do not aim to give a comprehensive plan of PMIP analyses, but instead here we focus on topics closely related to the CMIP6 key questions. Each PMIP4-CMIP6 period has been selected for specific reasons (Table 1). Here, we list several analyses which are important for single periods as well as for the full PMIP4-CMIP6 ensemble, starting first by presenting examples of paleoclimate reconstructions available for comparison to the PMIP4-CMIP6 simulations. .

4.1 Paleoclimatic and paleoenvironmental reconstructions, model-data comparisons.

Past environmental and climatic changes are typically documented at specific sites, whether on land, in ocean sediments or in corals, or from ice cores. The evaluation of climate simulations requires these paleoclimatic and paleoenvironmental data to be synthesised for specific time periods. A major challenge in building such syntheses is to synchronise the chronologies of the different records. There are many syntheses of information on past climates and environments and, although Table 4 lists some of the sources of quantitative reconstructions for the PMIP4-CMIP6 time periods, it is not our goal here to provide an extensive review of these resources. Much of this information stems from the impact of climatic changes on the environment, such as on fires, dust, marine microfauna and vegetation (which can be documented by pollen records, or for the recent period, by tree rings). Past climatic information is also contained in isotopic ratios of oxygen and carbon, which can be found in ice sheets, speleothems, or in the shells of marine organisms. Furthermore, the ocean circulation can also be documented by geochemical tracers in marine sediments from the sea floor (e.g. $\Delta^{14}C$, $\delta^{13}C$, $^{231}Pa/^{230}Th$, ε_{Nd}). The fact that these physical, chemical or biological indicators are indirect records of the state of the climate system and can also be sensitive to other factors (such as atmospheric CO_2 concentrations for vegetation) has to be taken into account in model-data comparisons. Comparisons with climate model output can therefore be

performed from different points of view: either the climate model output can be directly compared to reconstructions of past climate variables, or the response of the climatic indicator itself can be simulated from climate model output and compared to the climate indicator. Such "forward" models include dynamical vegetation models, tree ring models, or models computing the growth of foraminifera (cf. Section 4.7). Some paleoclimatic indicators such as meteoric water isotopes have to be computed as the climate model is running, but are also examples of this forward modelling approach. Modelling the impacts of past climate changes on the environment is key to understand how climatic signals are transmitted to past climate records. It also provides an opportunity to test the types of models that are used in the assessment of the impacts of future climate changes on the environment.

Table 4: examples of data syntheses for the PMIP4-CMIP6 periods

Reconstructing paleoclimates and paleoenvironments, as well as building new syntheses of these reconstructions, are very active areas of research. We expect new data sets to become available, which will increase the number of possible model-data comparisons for the PMIP4-CMIP6 periods.

4.2. Role of forcings and feedbacks

Quantifying the role of forcings and feedbacks in creating climates different from today has been a focus of PMIP for many years. Compared to the PMIP3-CMIP5 models, many CMIP6 models will include new processes, such as dust, or improved representations of major radiative feedback processes, such as clouds. Improvements to the design of the *past1000*, *midHolocene* and *lgm* experiments are also proposed (section 2). We will evaluate the impact of these changes on the PMIP4-CMIP6 climates at global, large-scale (e.g. polar amplification, land-sea contrast) as well as regional scales, together with the mechanisms explaining these impacts. All the PMIP4-CMIP6 experiments will be run with the same model version, facilitating analyses across the five time periods to examine potential relationships between forcings of different nature and amplitude and the climate responses, and compare the processes involved in these responses (e.g. Izumi et al., 2013). For example, there are temperature thresholds that determine whether snow and ice can be present, and temperature thresholds also play a part in determining the distribution of specific vegetation types. Thus, a given change in climate could have different effects on snow/ice or vegetation feedback depending on the base climate state. Density thresholds also play a part in controlling the oceanic overturning circulation, again leading to the possibility that ocean changes may be modulated by background state (Swingedouw et al. 2009).

Multi-period analyses provide a way of determining whether systematic model biases affect the overall response and the strength of feedbacks independent of climate state. They will allow us to determine, for example, whether the persistent failure to reproduce the observed magnitude of change in monsoon precipitation and the relatively small impact of vegetation feedback during the MH is related to biases or base climate. Similarly, they will help to quantify whether simulated changes in ocean circulation at the LGM are affected by systematic model biases or threshold behaviour. Model-data comparisons (cf. Section 4.1) will be used to assess the realism of the simulated climate change and to detect key mechanisms affecting model behaviour independently of the base climate state. One challenge will be to develop new approaches to analyse the PMIP4-CMIP6 ensemble so

as to separate the impacts of model resolution, content, or complexity on the simulated climate. Similarly the uncertainties in boundary conditions will be addressed for periods for which alternative forcing is proposed.

Ice sheets represent strong changes in radiative forcing, as well as a direct forcing on atmosphere circulation. The PMIP4-CMIP6 ensemble will allow new analyses of the impact of smaller (mPWP) or larger (LGM) ice sheets. The ocean and sea-ice feedbacks will also be analysed. The representation of sea ice and Southern Ocean proved to be problematic in previous simulations of colder (LGM, Roche et al., 2012) and warmer climates (LIG, Bakker et al., 2013, Lunt et al., 2013). For the LGM, there is evidence of a shallower, and yet active overturning circulation in the North Atlantic (e.g. Lynch-Stieglitz et al., 2007, Böhm et al., 2015). Understanding this oceanic circulation as well as its links to surface climate is a topic of high importance since the Atlantic Meridional Overturning Circulation could modulate future climate changes at least in regions around the North Atlantic (IPCC 2013).

4.3 Benchmarking the PMIP4-CMIP6 simulations

The compatibility of past, historical and future climate simulations will allow benchmarking based on syntheses of paleoenvironmental data and paleoclimate reconstructions (Section 4.1) to be applied to same models used for future projections. We will make full use of the fact that modelling groups must also run the *piControl* and *historical* experiments, since the *piControl* and *historical* simulations provide two alternative reference states for paleoclimate simulations. Existing paleoclimate reconstructions have used different modern reference states, and this has been shown to have an impact on the magnitude of reconstructed changes (e.g. Hessler et al., 2014). Comparisons of the simulated *piControl* and the *historical* climates will provide a way of quantifying this source of reconstruction uncertainty, as will comparisons with present-day observations and reanalysis data sets (Obs4MIPS, Ferraro et al, 2015).

Systematic benchmarking of each of the PMIP4-CMIP6 simulations will benefit from the existing paleoclimatic data sets (Section 4.1, Table 4) and from the development of new data syntheses. Large-scale features, such as polar amplification, land-sea contrast, and the scaling between precipitation and temperature changes, as well as more regional features such as the monsoons and mid-continental climates, and climate responses over specific data-rich areas (such as Europe or North America) are prime targets for benchmarking. The ensemble of metrics developed for the PMIP3-CMIP5 *midHolocene* and *lgm* simulations (e.g. Harrison et al. 2014) will be expanded to include more process-oriented metrics. Benchmarking results from the PMIP4- *midHolocene* and *lgm* "entry card" simulations will be compared to benchmark metrics from previous generations of PMIP to provide a rigorous assessment of model improvements since the last phase of CMIP.

There are many aspects of the climate system which are difficult to measure directly, and which are therefore difficult to evaluate using traditional methods. The "emergent constraint" (e.g. Sherwood et al., 2014) concept, which is based on identifying a relationship to a more easily measurable variable, has been successfully used by the carbon-cycle and modern climate communities and holds great potential for the analysis of paleoclimate simulations. This could be particularly valuable to examine the realism of e.g. cloud feedbacks in the simulations

or the contribution of seasonal climate changes to hydrological budgets. Using multiple time periods to examine emergent constraints ensures that they are robust across climate states.

4.4 Relating past and future climate changes

Attempts to constrain climate sensitivity using information about the LGM period have been hampered by the fact that there were too few *lgm* experiments to draw statistically robust conclusions (Hargreaves et al., 2012; Harrison et al., 2014; Hopcroft and Valdes, 2015b). These attempts also ignored uncertainties in forcings and boundary conditions. PMIP4-CMIP6 is expected to result in a much larger ensemble of *lgm* experiments. The issue of climate sensitivity (*sensu stricto*) and earth-system sensitivity (PALEOSENS Project Members, 2012) will also be examined through joint analysis of multiple paleoclimate simulations and climate reconstructions from different archives. Our analyses will capitalise on the DECK *piControl* and *abrupt4xCO2* experiments, as well as on the CFMIP experiments *AMIP4K* and *AMIPminus4K*.

The relationship between radiative forcing and global temperature is not straightforward (Crucifix, 2006; Yoshimori et al., 2011), partly because the nature of the forcing that drives the Earth into different climate states preferentially triggers short wave or long wave radiative responses, that have different impacts on the energy or water exchanges, on the feedbacks between different climate system components, or have different large- or regional- scale patterns. Identification of similarities between past climates and future climate projections such as the one found for land-sea contrast or polar amplification (Izumi et al., 2013; Masson et al., 2013; Izumi et al., 2015) or for snow and cloud feedbacks for particular seasons (Braconnot and Kageyama, 2015) will be used to improve understanding of the relationship between patterns and time scales of external forcings and patterns and timing of the climate responses. Improved model-data comparisons should also provide new possibilities to link regional climate reconstructions to the Earth's global energetic and climate sensitivity. Additional constraints can be obtained by using perturbed-physics experiments, in which different versions of the same model are run using different values of key parameters (Annan et al., 2005: Yoshimori et al., 2011). The 'perturbed forcing' approach (Bounceur et al., 2015; Araya-Melo, 2015), using sensitivity experiments carried out in PMIP4, could provide a way to chart the sensitivity of the climate system in a multi-dimensional space of forcing conditions.

4.5 Changes in mean climate vs. changes in climate variability

Multi-period analyses will also be useful for understanding the relationship between mean climate state and modes of natural variability (e.g. Liu et al., 2014; Saint-Lu et al., 2015). Future changes in modes of climate variability, such as El Niño Southern Oscillation (ENSO), are poorly constrained (Christiansen et al., 2013) because model projections are insufficiently long to provide robust statistics for low frequency (multi-decadal and centennial) variations. Robust statistics of ENSO changes have been derived through analysis of high-resolution paleo-records (Emile-Geay et al., 2016). The equilibrium paleoclimate experiments in PMIP4-CMIP6 will provide an opportunity to sample simulations for long enough to obtain robust estimates of ENSO changes (Stevenson et al., 2010) and analyses of multiple long simulations with different forcings should provide a better understanding of changes in ENSO behaviour (Zheng et al., 2008; An and Choi, 2014) and help determine whether state-of-the-art climate models underestimate low frequency noise (Laepple and Huybers, 2014). The PMIP Paleovariability Working Group will develop diagnostics for climate variability (Philips et al., 2014) to be

applied to all the PMIP4-CMIP6 simulations. Analyses will focus on how models reproduce the relationship between changes in seasonality and interannual variability (Emile-Geay et al., 2016), the diversity of El-Niño events (Capotondi et al., 2015; Karamperidou et al., 2015; Luan et al., 2015), and the stability of teleconnections within the climate system (e.g. Gallant et al., 2013; Batehup et al., 2015).

For shorter time scales, the *past1000* simulations and corresponding high temporal resolution data are one of the only means to examine the mechanisms and realism of the relationships between events at the daily scale (e.g. weather extremes) and longer-term climatic changes.

4.6 Interactions with other CMIP6 MIPs and the WCRP Grand Challenges

PMIP has already developed strong links with several other CMIP6 MIPs (Table 5). CFMIP includes an idealized experiment mimicking the *lgm* simulation: *AMIPminus4K* is an atmosphere-only experiment in which the sea-surface temperatures are uniformly lowered by 4K (a mirror of the *AMIP4K* experiment in which sea-surface temperatures are increased by 4K). These experiments allow investigations of cloud feedbacks and associated circulation changes in a colder versus a warmer world and this will assist in disentangling the processes at work in the *lgm* climate. Some MIPs have designed experiments based on PMIP data, including VolMIP for the study of the impact of large past volcanic eruptions and ISMIP6 for the impact of the last interglacial climate on the Greenland ice sheet. Links with CFMIP and ISMIP6 mean that PMIP will also contribute to the World Climate Research Programme (WCRP) Grand Challenges "Clouds, Circulation and Climate Sensitivity" and "Cryosphere and Sea Level" respectively. Furthermore, PMIP will provide input to the WCRP Grand Challenge on "Regional Climate Information", through a focus on evaluating the mechanisms of regional climate change in the past, for example in the Arctic.

Interactions between PMIP and other CMIP6 MIPs have mutual benefits: PMIP provides simulations of large climate changes that have occurred in the past and evaluation tools which capitalise on extensive data syntheses, while other MIPs will employ diagnostics and analyses which will be useful for analyzing the PMIP4 experiments. This is the case for AerChemMIP (Collins et al., 2017) for the aerosol forcings, SIMIP (Notz et al., 2016) and OMIP (Griffies et al., 2016) for the sea-ice and ocean components, LS3MIP (van den Hurk et al., 2016) for the land surface, C4MIP (Jones et al., 2016) for the carbon cycle, ISMIP (Nowicki et al., 2016) for ice sheets, and CFMIP (Webb et al., 2017) for the cloud forcing and feedback analyses. The analytical tools developed in RFMIP (Pincus et al., 2016) will be useful for assessing the LGM GHG radiative forcing and those developed in VolMIP (Zanchettin et al., 2016) and LUMIP (Lawrence et al., 2016) will be relevant for the analyses of the impacts of volcanic and land use forcings in the *past1000* simulation. The *past1000* experiment also offers a long time series perturbed by natural forcings and reconstructed land use changes for detection and attribution exercises and it is therefore relevant for DAMIP (Gillett et al., 2016). We have ensured that all the outputs necessary for the application of common diagnostics across PMIP and other CMIP6 MIPs will be available (see section 4.7).

Table 5: interactions of PMIP with other CMIP6 MIPs

4.7 Implications: required model output for the PMIP4-CMIP6 database

The list of variables required to analyse the PMIP4-CMIP6 paleoclimate experiments (https://wiki.lsce.ipsl.fr/pmip3/doku.php/pmip3:wg:db:cmip6request) reflects plans for multi-time period analyses and for interactions with other CMIP6 MIPs. We have included relevant variables from the data requests of other MIPs, including the CFMIP-specific diagnostics on cloud forcing, as well as land surface, snow, ocean, sea ice, aerosol, carbon cycle and ice sheet variables from LS3MIP, OMIP, SIMIP, AerChemMIP, C4MIP, and ISMIP6 respectively. Some of these variables are also required to diagnose how climate signals are recorded by paleoclimatic sensors via models of e.g. tree growth (Li et al., 2014), vegetation dynamics (Prentice et al., 2011) or marine planktonic foraminifera (e.g. Lombard et al., 2011; Kageyama et al., 2013). The only set of variables defined specifically for PMIP are those describing oxygen isotopes in the climate system. Isotopes are widely used for paleoclimatic reconstruction and are explicitly simulated in several models. We have asked that average annual cycles of key variables are included in the PMIP4-CMIP6 data request for equilibrium simulations, as these proved exceptionally useful for analyses in PMIP3-CMIP5.

Variations in the shape of the Earth's orbit govern the latitudinal and seasonal distribution of insolation, and also produce variations in the lengths of individual "months" (where months are defined alternatively as either (a) the duration in days for the Earth to complete one-twelfth of its orbit (the "celestial" or "angular" calendar), or (b) a specific number of days, e.g. 31 days in January, 30 days in June (the "conventional" or "modern" calendar). For example, at 6 ka, perihelion occurs in August, aphelion in February and those months were approximately 1.5 days shorter and longer than at present, respectively (Fig. 5a). Variations in the lengths of months (or seasons) must therefore be taken into consideration when examining experiment minus control long-term mean differences, because the effect of the changing calendar on the calculation of long-term means can be as large as the potential differences among the means themselves (Joussaume and Braconnot, 1997; Pollard and Reusch, 2002; Timm et al., 2008; Chen et al., 2011). The size of the potential calendar effect or bias is illustrated in Fig. 4, and is even larger for *lig127k*, for which eccentricity is large. Figure 4 shows the difference between present-day long-term means for October temperature and precipitation, and those calculated using the appropriate celestial month lengths for 6 and 127 ka. Modifications to month length have not usually been taken into account in the model output post-treatment procedures (but see Harrison et al., 2014). The most straightforward way for dealing with the calendar effect is to save and use daily data for the calculation of monthly or seasonal means, and so we include those in the PMIP4-CMIP6 data request for some key variables. A second approach, less desirable, but probably adequate for our purposes, is to use a bias-correction approach, in particular, like that of Pollard and Reusch (2002), with the mean-preserving daily interpolation approach of Epstein (1991).

Figure 4: The calendar effect: (a) month-length anomalies, 140 ka to present, with the PMIP4 experiment times indicated by vertical lines. The month-length anomalies were calculated using the formulation in Kutzbach and Gallimore (1988). (b and c) The calendar effect on October temperature at 6 and 127 ka, calculated using Climate Forecast System Reanalysis near-surface air temperature (https://www.earthsystemcog.org/ projects/obs4mips/), 1981-2010 long-term means, and assuming the long-term mean differences in temperature are zero everywhere. (e and f) The calendar effect on October precipitation at 6 and 127 ka, calculated using the CPC Merged Analysis of Precipitation (CMAP) enhanced precipitation (http://www.esrl.noaa.gov/ psd/data/gridded/ data.cmap.html), 1981-2010 long-term means, and again assuming that the long-term mean differences in temperature are zero everywhere. Calendar effects were calculated by interpolating present-day monthly temperature or precipitation to a daily time

step as in Pollard and Reusch, 2002 (but using a mean-preserving algorithm for pseudo-daily interpolation for monthly values; Epstein, 1991), and then recalculating the monthly means using the appropriate paleo calendar (Bartlein and Shafer, 2016). Note that the 6 and 127 ka map patterns for both variables, while broadly similar, are not simply rescaled versions of one another.

Daily or 6-hourly values are also useful for running regional models. It is important to test the use of regional models for climate model projections at the regional scale. Regional models are also used to produce fine-scale palaeoclimate scenarios for use by the impact community, for example to study past climate impacts on biodiversity via ecological niche modelling. Paleoclimate indicators often respond to climate features not adequately captured with monthly data alone (such as growing season length). Daily weather variables are therefore required for some forward models, as well as to compute bioclimatic variables which are reconstructed e.g. based on pollen data (e.g. Bartlein et al., 2011).

5. Conclusions

PMIP4-CMIP6 simulations provide a framework to compare current and future anthropogenic climate change with past natural variations of the Earth's climate. PMIP4-CMIP6 is a unique opportunity to simulate past climates with exactly the same models as used for simulations of the future. This approach is only valid if the model versions and implementation of boundary conditions are consistent for all periods, and if these boundary conditions are seamless for overlapping periods.

PMIP4-CMIP6 simulations are important in terms of model evaluation for climate states significantly different from the present and historical climates. We have chosen climatic periods well documented by paleoclimatic and paleoenvironmental records, with climate and environmental changes relevant for the study and projections of future climate changes: the millennium prior to the industrial epoch (*past1000*), 6,000 years ago (*midHolocene*), the last glacial maximum (*lgm*), the last interglacial (*lig127k*) and the mid-Pliocene (*midPliocene-eoi400*). .

The PMIP4-CMIP6 experiments will also constitute reference simulations for projects developed in the broader PMIP4. The corresponding sensitivity experiments, or additional experiments, are embedded in the PMIP4 project and are described in the companion papers to this overview (Haywood et al., 2016, Otto-Bliesner et al., 2017, Jungclaus et al., 2017, Kageyama et al., 2017). They are essential for a deeper understanding of the drivers of past climate changes for the PMIP4-CMIP6 climates or as initial conditions for transient simulations (e.g. Ivanovic et al., 2016, for the last deglaciation, Otto-Bliesner et al., 2017 for the last interglacial and the Holocene), or for examining time periods from deeper with high atmospheric CO_2 concentrations (Lunt et al, 2017). Figure 5 summarises the position of the PMIP4-CMIP6 experiments with respect to the other PMIP4 experiments and projects on the right-hand-side. The left-hand-side shows how the PMIP4-CMIP6 experiments relate to the CMIP6 DECK and some other CMIP6 MIPs. PMIP4-CMIP6 experiments have been designed to be analyzed by both communities.

Figure 5: the PMIP4-CMIP6 experiments in the framework of CMIP6, with associated MIPs, and in the framework of PMIP4, with its working groups.

The PMIP community anticipates major benefits from analysis techniques developed by the other CMIP6 MIPs, in particular in terms of learning about the processes of past climate changes in response to forcings (e.g. greenhouse gases, astronomical parameters, ice sheet and sea level changes) as well as the role of feedbacks (e.g. clouds, ocean, sea-ice). Collaborations have already been developed with e.g. CFMIP, ISMIP6 and VolMIP, and the hope is to build additional collaborations with other CMIP6 MIPs. PMIP4-CMIP6 has the potential to be mutually beneficial for the paleoclimate and present/future climate scientists to learn about natural large climate changes and the mechanisms at work in the climate system for climates states as different from today as future climate is projected to be.

Data availability

All data mentioned in the present manuscript can be downloaded following the instructions given in the companion papers giving details on the PMIP4-CMIP6 experimental protocols (Otto-Bliesner et al., 2017, Jungclaus et al., 2017, Kageyama et al., 2017, Haywood et al., 2016).

Acknowledgements

MK and QZ acknowledge funding from the French-Swedish project GIWA. PB, JJ and SPH acknowledge funding from the JPI-Belmont project "PAleao-Constraints on Monsoon Evolution and Dynamics (PACMEDY)" through their respective national funding agencies. SPH also acknowledges funding from the Australian Research Council (DP1201100343) and from the European Research Council for "GC2.0: Unlocking the past for a clearer future". AMH and AMD acknowledge funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement no. 278636 and the EPSRC-supported Past Earth Network. RFI is funded by a NERC Independent Research Fellowship [#NE/K008536/1]. SJP's contribution is supported under the Australian Research Council's Special Research Initiative for the Antarctic Gateway Partnership (Project ID SR140300001). FL acknowledges support from CONICYT projects 15110009, 1151427, ACT1410, and NC120066.

References

Abe-Ouchi, A., Saito, F., Kageyama, M., Braconnot, P., Harrison, S. P., Lambeck, K., Otto-Bliesner, B. L., Peltier, W. R., Tarasov, L., Peterschmitt, J.-Y., and Takahashi, K.: Ice-sheet configuration in the CMIP5/PMIP3 Last Glacial Maximum experiments, Geosci. Model Dev., 8, 3621-3637, doi:10.5194/gmd-8-3621-2015, 2015.

Albani, S., Mahowald, N. M., Perry, A. T., Scanza, R. A., Zender, C. S., Heavens, N. G., Maggi, V., Kok, J. F., and Otto-Bliesner, B. L. .: Improved representation of dust size and optics in the CESM, Journal of Advances in Modeling of Earth Systems, 6, doi:10.1002/2013MS000279, 2014.

Albani, S., Mahowald, N. M., Winckler, G., Anderson, R. F., Bradtmiller, L. I., Delmonte, B., François, R., Goman, M., Heavens, N. G., Hesse, P. P., Hovan, S. A., Kang, S. G., Kohfeld, K. E., Lu, H., Maggi, V., Mason, J. A., Mayewski, P. A., McGee, D., Miao, X., Otto-Bliesner, B. L., Perry, A. T., Pourmand, A., Roberts, H. M., Rosenbloom, N., Stevens, T., and Sun, J.: Twelve thousand years of dust: the Holocene global dust cycle constrained by natural archives, Clim. Past, 11, 869-903, doi:10.5194/cp-11-869-2015, 2015.

Albani, S., Mahowald, N. M., Murphy, L. N., Raiswell, R., Moore, J. K., Anderson, R. F., McGee, D., Bradtmiller, L. I.,Delmonte, B., Hesse, P. P., and Mayewski, P. A.: Paleodust variability since the Last Glacial Maximum and implications for iron inputs to the ocean. Geophys. Res. Lett., 43, doi:10.1002/2016GL067911, 2016.

An, S.-I. and Choi, J.: Mid-Holocene tropical Pacific climate state, annual cycle, and ENSO in PMIP2 and PMIP3, Climate Dynamics, 43, 957-970, doi:10.1007/s00382-013-1880-z, 2014.

Annan, J. D., Hargreaves, J. C., Ohgaito, R., Abe-Ouchi, A., and Emori, S.: Efficiently constraining climate sensitivity with ensembles of paleoclimate experiments, SOLA, (1) 181-184 doi:10.2151/sola.2005-047, 2005.

Araya-Melo, P. A., Crucifix, M., and Bounceur, N.: Global sensitivity analysis of the Indian monsoon during the Pleistocene, Clim. Past, 11, 45-61, doi:10.5194/cp-11-45-2015, 2015.

Argus, D. F., Peltier, W. R., Drummond, R. and Moore, A. W.: The Antarctica component of postglacial rebound model ICE-6G_C (VM5a) based on GPS positioning, exposure age dating of ice thicknesses, and relative sea level histories, Geophys. J. Int., ggu140, doi:10.1093/gji/ggu140, 2014.

Bakker, P., Stone, E. J., Charbit, S., Gröger, M., Krebs-Kanzow, U., Ritz, S. P., Varma, V., Khon, V., Lunt, D. J., Mikolajewicz, U., Prange, M., Renssen, H., Schneider, B., and Schulz, M.: Last interglacial temperature evolution – a model inter-comparison, Clim. Past, 9, 605-619, doi:10.5194/cp-9-605-2013, 2013.

Bartlein, P. J., Harrison, S.P., Brewer, S., Connor, S., Davis B.A.S., Gajewski, K., Guiot, J., Harrison-Prentice, T.I., Henderson, A., Peyron, O., Prentice, I.C., Scholze, M., Seppä, H., Shuman, B., Sugita, S., Thompson, R.S., Viau, A., Williams, J., Wu, H.,: Pollen-based continental climate reconstructions at 6 and 21 ka: a global synthesis, Climate Dynamics 37, 775-802, 2011.

Bartlein, P.J. and Shafer, S.L.: The impact of the "calendar effect" and pseudo-daily interpolation algorithms on paleoclimatic data-model comparisons- https://agu.confex.com/ agu/fm16/meetingapp.cgi/Paper/187186 2016.

Bartoli, G., Honisch, B., and Zeebe, R. E.: Atmospheric CO_2 decline during the Pliocene intensification of Northern Hemisphere glaciations, Paleoceanography, 26, 2011.

Batehup, R., McGregor, S., and Gallant, A. J. E.: The influence of non-stationary teleconnections on palaeoclimate reconstructions of ENSO variance using a pseudoproxy framework, Clim. Past, 11, 1733-1749, doi:10.5194/cp-11-1733-2015, 2015.

Bereiter, B., Eggleston, S., Schmitt, J., Nehrbass-Ahles, C., Stocker, T. F., Fischer, H., Kipfstuhl, S. and Chappellaz, J.: Revision of the EPICA Dome C CO2 record from 800 to 600 kyr before present, Geophys. Res. Lett., 42(2), 2014GL061957, doi:10.1002/2014GL061957, 2015.

Berger, A.: Long-term variations of daily insolation and quaternary climatic changes, J. Atmos. Sci., 35, 2362-2367, 1978.

Bounceur, N., M. Crucifix, and R. Wilkinson: Global sensitivity analysis of the climate-vegetation system to astronomical forcing: an emulator-based approach, Earth System Dynamics, 6(1), 205. 2015

Briggs, R. D., Pollard, D. and Tarasov, L.: A data-constrained large ensemble analysis of Antarctic evolution since the Eemian, Quat. Sci. Rev., 103, 91–115, doi:10.1016/j.quascirev.2014.09.003, 2014.

Braconnot, P., Kageyama, M.: Shortwave forcing and feedbacks in Last Glacial Maximum and Mid-Holocene PMIP3 simulations, Phil. Trans. R. Soc. A 2015 373 20140424; DOI: 10.1098/rsta.2014.0424, 2015.

Braconnot, P., Harrison, S.P., Kageyama, M., Bartlein, P.J., Masson-Delmotte, V., Abe-Ouchi, A., Otto-Bliesner, B. and Zhao, Y.: Evaluation of climate models using palaeoclimatic data, Nature Climate Change, 2, 417-424, 2012.

Caballero, R., Huber, M.: Spontaneous transition to superrotation in warm climates simulated by CAM3, Geophys. Res. Lett., 37, L11701, doi:10.1029/2010GL043468, 2010.

Capotondi, A., Wittenberg, A., T., Newman, M., Di Lorenzo, E., Yu, J.-Y., Braconnot, P., Cole, J., Dewitte, B., Giese, B., Guilyardi, E., Jin, F.-F., Karnauskas, K., Kirtman, B., Lee, T., Schneider, N., Xue, Y., Yeh, S.-W.: Understanding ENSO Diversity, Bull. Am. Met. Soc., 96, 921-938, doi:10.1175/Bams-D-13-00117.1., 2015.

Capron, E., Govin, A., Feng, R., Otto-Bliesner, B., and Wolff, E. W.: Critical evaluation of climate syntheses to benchmark CMIP6/PMIP4 127 ka Last Interglacial simulations in the high-latitude regions, Quaternary Science Reviews, 168, 137-150, doi:10.1016/j.quascirev.2017.04.019, 2017.

Capron, E., Govin, A., Stone, E.J., Masson-Delmotte, V., Mulitza, S., Otto-Bliesner, B., Rasmussen, T.L., Sime, L.C., Waelbroeck, C. and Wolff, E.W.: Temporal and spatial structure of multi-millennial temperature changes at high latitudes during the Last Interglacial. Quat. Sci. Rev., 103: 116-133, 2014.

Chen, G.-S., Kutzbach, J. E., Gallimore, R., and Liu, Z.: Calendar effect on phase study in paleoclimate transient simulation with orbital forcing, Climate Dynamics, 37, 1949-1960, 2011.

Collins, W. J., Lamarque, J.-F., Schulz, M., Boucher, O., Eyring, V., Hegglin, M. I., Maycock, A., Myhre, G., Prather, M., Shindell, D., and Smith, S. J.: AerChemMIP: quantifying the effects of chemistry and aerosols in CMIP6, Geosci. Model Dev., 10, 585-607, https://doi.org/10.5194/gmd-10-585-2017, 2017.

Crucifix, M.: Does the Last Glacial Maximum constrain climate sensitivity?, Geophysical Research Letters, 33, L18701, doi:10.1029/2006GL027137, 2006.

Dowsett, H.J., Cronin, T. M.: High eustatic sea level during the middle Pliocene: Evidence from the southeastern U.S. Atlantic Coastal Plain, Geology. 18(5), 435-438, 1990.

Dowsett, H. J., Robinson, M. M., Haywood, A. M., Hill, D. J., Dolan, A. M., Stoll, D. K., Chan, W.-L., Abe-Ouchi, A., Chandler, M. A., and Rosenbloom, N. A.: Assessing confidence in Pliocene sea surface temperatures to evaluate predictive models, Nature Climate Change, 2, 365–371, 2012.

Dowsett, H., Dolan, A., Rowley, D., Moucha, R., Forte, A. M., Mitrovica, J. X., Pound, M., Salzmann, U., Robinson, M., Chandler, M., Foley, K., and Haywood, A.: The PRISM4 (mid-Piacenzian) paleoenvironmental reconstruction, Clim. Past, 12, 1519-1538, doi:10.5194/cp-12-1519-2016, 2016.

Dutton, A., Carlson, A.E., Long, A.J., Milne, G.A., Clark, P.U., DeConto, R., Horton, B.P., Rahmstorf, S. and Raymo, M.E.: Sea-level rise due to polar ice-sheet mass loss during past warm periods, Science, 349, aaa4019, doi: 10.1126/science.aaa4019, 2015.

Dwyer, G.S. and Chandler, M.A.: Mid-Pliocene sea level and continental ice volume based on coupled benthic Mg/Ca palaeotemperatures and oxygen isotopes, Philosophical Transactions of the Royal Society, A. 367, 157-168, 2009.

Emile-Geay, J., Cobb, K. M., Carré, M., Braconnot, P., Leloup, J., Zhou, Y., Harrison, S. P., Corrège, T., McGregor, H. V., Collins, M., Driscoll, R., Elliot, M., Schneider, B., and Tudhope, A.: Links between tropical Pacific seasonal, interannual and orbital variability during the Holocene, Nature Geoscience, 9, 168, doi:10.1038/ngeo2608. 2016.

Epstein, E. S.: On obtaining daily climatological values from monthly means, Journal of Climate, 4, 365-368, 1991.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937-1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fernández-Donado, L., Gonzalez-Rouco, J.F., Raible, C.C., Amman, C.M., Barriopedro, D. Garcia-Bustamante, E., Jungclaus, J.H., Lorenz, S.J., Luterbacher, J. Phipps, S.J., Servonnat, J., Swingedouw, D., Tett, S.F.B., Wagner, S., Yiou, P., and Zorita, E.: Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium, Clim. Past, 9, 393-421, doi:10.5194/cp-9-393-2013, 2013.

Ferraro, R., Waliser, D. E., Gleckler, P., Taylor, K. E., and Eyring, V.: Evolving Obs4MIPs to Support Phase 6 of the Coupled Model Intercomparison Project (CMIP6), Bulletin of the American Meteorological Society, 96, ES131-ES133, 2015.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C. and Rummukainen, M.: Evaluation of climate models. In: Climatic Change 2013: The Physical 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.

Gallant, A. J. E., S. J. Phipps, D. J. Karoly, A. B. Mullan and A. M. Lorrey: Non-stationary Australasian teleconnections and implications for paleoclimate research, J. Clim., 26, 8827-8849, doi:10.1175/JCLI-D-12-00338.1, 2013.

Gillett, N. P., Shiogama, H., Funke, B., Hegerl, G., Knutti, R., Matthes, K., Santer, B. D., Stone, D., and Tebaldi, C.: The Detection and Attribution Model Intercomparison Project (DAMIP v1.0) contribution to CMIP6, Geosci. Model Dev., 9, 3685-3697, https://doi.org/10.5194/gmd-9-3685-2016, 2016.

Griffies, S. M., Danabasoglu, G., Durack, P. J., Adcroft, A. J., Balaji, V., Böning, C. W., Chassignet, E. P., Curchitser, E., Deshayes, J., Drange, H., Fox-Kemper, B., Gleckler, P. J., Gregory, J. M., Haak, H., Hallberg, R. W., Heimbach, P., Hewitt, H. T., Holland, D. M., Ilyina, T., Jungclaus, J. H., Komuro, Y., Krasting, J. P., Large, W. G., Marsland, S. J., Masina, S., McDougall, T. J., Nurser, A. J. G., Orr, J. C., Pirani, A., Qiao, F., Stouffer, R. J., Taylor, K. E., Treguier, A. M., Tsujino, H., Uotila, P., Valdivieso, M., Wang, Q., Winton, M., and Yeager, S. G.: OMIP contribution to CMIP6: experimental and diagnostic protocol for the physical component of the Ocean Model Intercomparison Project, Geosci. Model Dev., 9, 3231-3296, https://doi.org/10.5194/gmd-9-3231-2016, 2016.

Hargreaves, J. C., Annan, J. D.: Can we trust climate models?, invited opinion article for WIREs Climate Change, WIREs Clim Change, 5, 435–440, doi: 10.1002/wcc.288, 2014.

Hargreaves, J. C., Annan, J. D., Yoshimori, M., & Abe-Ouchi, A.: Can the Last Glacial Maximum constrain climate sensitivity?, Geophysical Research Letters, 39, http://doi.org/10.1029/2012GL053872, 2012.

Hargreaves, J. C., Annan, J. D., Ohgaito, R., Paul, A., and Abe-Ouchi, A.: Skill and reliability of climate model ensembles at the Last Glacial Maximum and mid-Holocene, Clim. Past, 9, 811-823, doi:10.5194/cp-9-811-2013, 2013.

Harrison, S. P., Bartlein, P. J., Brewer, S., Prentice, I. C., Boyd, M., Hessler, I., Holmgren, K., Izumi, K., and Willis, K.: Climate model benchmarking with glacial and mid-Holocene climates, Climate Dynamics 43, 671-688, 2014.

Harrison, S. P., Bartlein, P. J., Izumi, K., Li, G., Annan, J., Hargreaves, J., Braconnot, P., and Kageyama, M.: Evaluation of CMIP5 palaeo-simulations to improve climate projections, Nature Clim. Change 5, 735-743, doi:10.1038/nclimate2649, 2015.

Hartmann, D.L., Klein Tank, A.M.G., Rusticucci, M., Alexander, L. V., Brönnimann, S., Charabi, Y., Dentener, F. J., Dlugokencky, E. J., Easterling, D. R., Kaplan, A., Soden, B. J., Thorne, P. W., Wild, M. and Zhai, P. M.: Observations: Atmosphere and Surface. 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, 2013.

Haywood, A. M., Dowsett, H. J., Otto-Bliesner, B., Chandler, M. A., Dolan, A. M., Hill, D. J., Lunt, D. J., Robinson, M. M., Rosenbloom, N., Salzmann, U., and Sohl, L. E.: Pliocene Model Intercomparison Project (PlioMIP): experimental design and boundary conditions (Experiment 1), Geosci. Model Dev., 3, 227-242, doi:10.5194/gmd-3-227-2010, 2010.

Haywood, A. M., Dowsett, H. J., Robinson, M. M., Stoll, D. K., Dolan, A. M., Lunt, D. J., Otto-Bliesner, B., and Chandler, M. A.: Pliocene Model Intercomparison Project (PlioMIP): experimental design and boundary conditions (Experiment 2), Geosci. Model Dev., 4, 571-577, doi:10.5194/gmd-4-571-2011, 2011.

Haywood, A. M., Hill, D. J., Dolan, A. M., Otto-Bliesner, B. L., Bragg, F., Chan, W.-L., Chandler, M. A., Contoux, C., Dowsett, H. J., Jost, A., Kamae, Y., Lohmann, G., Lunt, D. J., Abe-Ouchi, A., Pickering, S. J., Ramstein, G., Rosenbloom, N. A., Salzmann, U., Sohl, L., Stepanek, C., Ueda, H., Yan, Q., and Zhang, Z.: Large-scale features of Pliocene climate: results from the Pliocene Model Intercomparison Project, Clim. Past, 9, 191-209, doi:10.5194/cp-9-191-2013, 2013.

Haywood, A. M., Dowsett, H. J., Dolan, A. M., Rowley, D., Abe-Ouchi, A., Otto-Bliesner, B., Chandler, M. A., Hunter, S. J., Lunt, D. J., Pound, M., and Salzmann, U.: The Pliocene Model Intercomparison Project (PlioMIP) Phase 2: scientific objectives and experimental design, Clim. Past, 12, 663-675, doi:10.5194/cp-12-663-2016, 2016.

Hessler, I., Harrison, S.P., Kuchera, M., Waelbroeck, C., Chen, M-T., Anderson, C., de Vernal, A., Fréchette, B., Cloke-Hayes, A. and Londeix, L.: Implication of methodological uncertainties for mid-Holocene sea surface temperature reconstructions.. *Climate of the Past* 10: 2237-2252, 2014.

Hind, A., and Moberg, A.: Past millennium solar forcing magnitude. A statistical hemispheric-scale climate model versus proxy data comparison. Clim. Dyn., 41, 2527-2537, doi:10.1007/s00382-012-1526-6, 2013.

Hopcroft, P.O., and Valdes, P. J.: Last Glacial Maximum constraints on the Earth System Model HadGEM2-ES, Climate Dynamics, 45(5), 1657-1672, doi:10.1007/s00382-014-2421-0, 2015a.

Hopcroft, P.O., and Valdes, P. J.: How well do simulated last glacial maximum tropical temperatures constrain equilibrium climate sensitivity?, *Geophysical Research Letters*, 42(13), 5533-5539, doi:10.1002/2015GL064903, 2015b.

Hopcroft, P.O., Valdes, P.J., Woodward. S. and Joshi, M.: Last glacial maximum radiative forcing from mineral dust aerosols in an Earth System model, Journal of Geophysical Research, 120, 8186-8205, doi:10.1002/2015JD023742, 2015.

Ivanovic, R. F., Gregoire, L. J., Kageyama, M., Roche, D. M., Valdes, P. J., Burke, A., Drummond, R., Peltier, W. R. and Tarasov, L.: Transient climate simulations of the deglaciation 21–9 thousand years before present (version 1) – PMIP4 Core experiment design and boundary conditions, Geosci Model Dev, 9(7), 2563–2587, doi:10.5194/gmd-9-2563-2016, 2016.

Izumi, K., Bartlein, P. J., and Harrison, S. P.: Consistent large-scale temperature responses in warm and cold climates, Geophysical Research Letters, 40, 1817-1823, doi: 10.1002/grl.50350, 2013.

Izumi, K., Bartlein, P. J. and Harrison, S. P.: Energy-balance mechanisms underlying consistent large-scale temperature responses in warm and cold climates, Climate Dynamics, 44, 3111-3127, doi: 10.1007/s00382-014-2189-2, 2015.

Jones, C. D., Arora, V., Friedlingstein, P., Bopp, L., Brovkin, V., Dunne, J., Graven, H., Hoffman, F., Ilyina, T., John, J. G., Jung, M., Kawamiya, M., Koven, C., Pongratz, J., Raddatz, T., Randerson, J. T., and Zaehle, S.: C4MIP – The Coupled Climate–Carbon Cycle Model Intercomparison Project: experimental protocol for CMIP6, Geosci. Model Dev., 9, 2853-2880, https://doi.org/10.5194/gmd-9-2853-2016, 2016.

Joussaume, S. and Braconnot, P.: Sensitivity of paleoclimate simulation results to season definitions, Journal of Geophysical Research-Atmospheres, 102, 1943-1956, 1997.

Joussaume, S., Taylor, K. E., Braconnot, P., Mitchell, J. F. B., Kutzbach, J. E., Harrison, S. P., Prentice, I. C., Broccoli, A. J., Abe-Ouchi, A., Bartlein, P. J., Bonfils, C., Dong, B., Guiot, J., Herterich, K., Hewitt, C. D., Jolly, D., Kim, J. W., Kislov, A., Kitoh, A., Loutre, M. F., Masson, V., McAvaney, B., McFarlane, N., de Noblet, N., Peltier, W. R., Peterschmitt, J. Y., Pollard, D., Rind, D., Royer, J. F., Schlesinger, M. E., Syktus, J., Thompson, S., Valdes, P., Vettoretti, G., Webb, R. S., and Wyputta, U.: Monsoon Changes for 6000 Years Ago: Results of 18 Simulations from the Paleoclimate Modeling Intercomparison Project (PMIP), Geophys. Res. Lett., 26, 859–862, 1999.

Joussaume, S., Taylor, K. E.: Status of the Paleoclimate Modeling Intercomparison Project (PMIP). Proceedings of the first international AMIP scientific conference. WCRP Report, 425-430, 1995.

Jungclaus, J.H., Lohmann, K., and Zanchettin, D.: Enhanced 20th century heat transfer to the Arctic simulated in the context of climate variations over the last millennium. Climate of the Past, 10, 2201-2213, doi:10.5194/cp-10-2201-2014, 2014.

Jungclaus, J. H., Bard, E., Baroni, M., Braconnot, P., Cao, J., Chini, L. P., Egorova, T., Evans, M., González-Rouco, J. F., Goosse, H., Hurtt, G. C., Joos, F., Kaplan, J. O., Khodri, M., Klein Goldewijk, K., Krivova, N., LeGrande, A. N., Lorenz, S. J., Luterbacher, J., Man, W., Meinshausen, M., Moberg, A., Nehrbass-Ahles, C., Otto-Bliesner, B. I., Phipps, S. J., Pongratz, J., Rozanov, E., Schmidt, G. A., Schmidt, H., Schmutz, W., Schurer, A., Shapiro, A. I., Sigl, M., Smerdon, J. E., Solanki, S. K., Timmreck, C., Toohey, M., Usoskin, I. G., Wagner, S., Wu, C.-Y., Yeo, K. L., Zanchettin, D., Zhang, Q., and Zorita, E.: The PMIP4 contribution to CMIP6 – Part 3: the Last Millennium, Scientific Objective and Experimental Design for the PMIP4 *past1000* simulations, Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-278, in press, 2017.

Kageyama, M., Albani, S., Braconnot, P., Harrison, S. P., Hopcroft, P. O., Ivanovic, R. F., Lambert, F., Marti, O., Peltier, W. R., Peterschmitt, J.-Y., Roche, D. M., Tarasov, L., Zhang, X., Brady, E. C., Haywood, A. M., LeGrande, A. N., Lunt, D. J., Mahowald, N. M., Mikolajewicz, U., Nisancioglu, K. H., Otto-Bliesner, B. L., Renssen, H., Tomas, R. A., Zhang, Q., Abe-Ouchi, A., Bartlein, P. J., Cao, J., Lohmann, G., Ohgaito, R., Shi, X., Volodin, E., Yoshida, K., Zhang, X., and Zheng, W.: The PMIP4 contribution to CMIP6 – Part 4: Scientific objectives and experimental design of the PMIP4-CMIP6 Last Glacial Maximum experiments and

Commentaire [mk1]: to be updated with final ref

PMIP4 sensitivity experiments, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-18, in press, 2017.

Karamperidou, C., P. N. Di Nezio, A. Timmermann, F. F. Jin, and K. M. Cobb, The response of ENSO flavors to mid-Holocene climate: implications for proxy interpretation, Paleoceanography, 30, 527-547. 2015

Klein Goldewijk, K.: A historical land use data set for the Holocene; HYDE 3.2, DANS, http://dx.doi.org/10.17026/dans-znk-cfy3, 2016.

Kohfeld, K.E. and Harrison, S.P.: DIRTMAP: The geological record of dust. *Earth Science Reviews* 54: 81-114, 2001.

Kohfeld, K., Le Quéré, C., Harrison, S. P., Anderson, R. F.: Role of Marine Biology in Glacial-Interglacial CO_2 Cycles, Science, 308, 74-78, 2005.

Kopp, R. E., Kemp, A. C., Bittermann, K., Horton, B. P., Donnelly, J. P., Gehrels, W. R., Hay, C. C., Mitrovica, J. X., Morrow, E. D., and Rahmstorf, S.: Temperature-driven global sea-level variability in the Common Era, Proceedings of the National Academy of Sciences of the United States of America, 113, E1434-E1441, 2016.

Krantz, D.E.: A chronology of Pliocene sea-level fluctuations: The U.S. Middle Atlantic Coastal Plain record, Quaternary Science Reviews, 10(2-3),163-174, 1991.

Kurschner, W. M., vanderBurgh, J., Visscher, H., and Dilcher, D. L.: Oak leaves as biosensors of late Neogene and early Pleistocene paleoatmospheric CO_2 concentrations, Marine Micropaleontology, 27, 299-312, 1996.

Kutzbach, J. E. and Gallimore, R. G.: Sensitivity of a coupled atmposphere/mixed layer ocean model to changes in orbital forcing at 9000 years BP, Journal of Geophysical Research, 93 (D1), 803-821, 1988.

Laepple, T., and Huybers, P. J.: Global and regional variability in marine surface temperatures. Geophysical Research Letters, 41(7), 2528–2534. http://doi.org/10.1002/2014GL059345, 2014.

Lambeck, K., Rouby, H., Purcell, A., Sun, Y., and Sambridge, M.: Sea level and global ice volumes from the Last Glacial Maximum to the Holocene, PNAS, 111, 15296-15303, 2014.

Lambert, F., Tagliabue, A., Shaffer, G., Lamy, F., Winckler, G., Farias, L., Gallardo, L., and De Pol-Holz, R.: Dust fluxes and iron fertilization in Holocene and Last Glacial Maximum climates, Geophys. Res. Lett., 42(14), 6014-6023, doi:10.1002/2015GL064250, 2015.

Laskar, J., Robutel, P., Joutel, F., Gastineau, M., Correia, A. C. M., and Levrard, B.: A long-term numerical solution for the insolation quantities of the Earth, Astronomy & Astrophysics, 428, 261-285, 2004.

Lawrence, D. M., Hurtt, George C., Arneth, A., Brovkin, V., Calvin, K. V., Jones, Andrew D., Jones, C. D., Lawrence, P. J., de Noblet-Ducoudré, N., Pongratz, J., Seneviratne, S. I., and Shevliakova, E.: The Land Use Model Intercomparison Project (LUMIP): Rationale and experimental design, Geosci. Model Dev., 9, 2973-2998, doi:10.5194/gmd-9-2973-2016, 2016.

Li, G., Harrison, S. P., Izumi, K. and Prentice, I. C.: Precipitation scaling with temperature in warm and cold climates: an analysis of CMIP5 simulations, Geophysical Research Letters, 40, 4018-4024, doi: 10.1002/grl.50730, 2013.

Commentaire [mk2]: to be updated with final ref

Li, G., Harrison, S.P., Prentice, I.C., Falster, D.: Interpretation of tree-ring data with a model for primary production, carbon allocation and growth. *Biogeosciences* 11: 6711-6724, 2014.

Lisiecki, L. E. and Raymo, M. E.: A Pliocene-Pleistocene stack of 57 globally distributed benthic delta O-18 records, Paleoceanography, 20, 2005.

Liu, Z., Lu, Z., Wen, X., Otto-Bliesner, B. L., Timmermann, A., & Cobb, K. M.: Evolution and forcing mechanisms of El Niño over the past 21,000 years. Nature, 515(7528), 550–553. http://doi.org/10.1038/nature13963, 2014.

Lombard, F., Labeyrie, L., Michel, E., Bopp, L., Cortijo, E., Retailleau, S., Howa, H., and Jorissen, F.: Modelling planktic foraminifer growth and distribution using an ecophysiological multi-species approach, Biogeosciences, 8, 853-873, doi:10.5194/bg-8-853-2011, 2011.

Loulergue, L., Schilt, A., Spahni, R., Masson-Delmotte, V., Blunier, T., Lemieux, B., Barnola, J.-M., Raynaud, D., Stocker, T. F. and Chappellaz, J.: Orbital and millennial-scale features of atmospheric CH4 over the past 800,000 years, Nature, 453(7193), 383–386, doi:10.1038/nature06950, 2008.

Lunt, D. J., Haywood, A., Schmidt, G., Salzmann, U., Valdes, P. and Dowsett, H.: Earth system sensitivity inferred from Pliocene modelling and data, *Nature Geoscience*, 3, 60-64, 2010.

Lunt, D., J., Dunkley Jones, T., Heinemann, M., Huber, M., LeGrande, A., Winguth, A., Loptson, C., Marotzke, J., Roberts, C. D., Tindall, J., Valdes, P., and Winguth, C.: A model–data comparison for a multi-model ensemble of early Eocene atmosphere–ocean simulations: EoMIP, Clim. Past, 8, 1717-1736, doi:10.5194/cp-8-1717-2012, 2012.

Lunt, D. J., Abe-Ouchi, A., Bakker, P., Berger, A., Braconnot, P., Charbit, S., Fischer, N., Herold, N., Jungclaus, J. H., Khon, V. C., Krebs-Kanzow, U., Langebroek, P. M., Lohmann, G., Nisancioglu, K. H., Otto-Bliesner, B. L., Park, W., Pfeiffer, M., Phipps, S. J., Prange, M., Rachmayani, R., Renssen, H., Rosenbloom, N., Schneider, B., Stone, E. J., Takahashi, K., Wei, W., Yin, Q., and Zhang, Z. S.: A multi-model assessment of last interglacial temperatures, Clim. Past, 9, 699-717, doi:10.5194/cp-9-699-2013, 2013.

Lunt, D. J., Huber, M., Anagnostou, E., Baatsen, M. L. J., Caballero, R., DeConto, R., Dijkstra, H. A., Donnadieu, Y., Evans, D., Feng, R., Foster, G. L., Gasson, E., von der Heydt, A. S., Hollis, C. J., Inglis, G. N., Jones, S. M., Kiehl, J., Kirtland Turner, S., Korty, R. L., Kozdon, R., Krishnan, S., Ladant, J.-B., Langebroek, P., Lear, C. H., LeGrande, A. N., Littler, K., Markwick, P., Otto-Bliesner, B., Pearson, P., Poulsen, C. J., Salzmann, U., Shields, C., Snell, K., Stärz, M., Super, J., Tabor, C., Tierney, J. E., Tourte, G. J. L., Tripati, A., Upchurch, G. R., Wade, B. S., Wing, S. L., Winguth, A. M. E., Wright, N. M., Zachos, J. C., and Zeebe, R. E.: The DeepMIP contribution to PMIP4: experimental design for model simulations of the EECO, PETM, and pre-PETM (version 1.0), Geosci. Model Dev., 10, 889-901, https://doi.org/10.5194/gmd-10-889-2017, 2017.

Maher, B., Prospero, J. M., Mackie, D., Gaiero, D., Hesse, P. P., Balkanski, Y.: Global connections between aeolian dust, climate and ocean biogeochemistry at the present day and at the last glacial maximum, Earth-Science Reviews, 99, 61-97, 2010.

Man, W., Zhou, T., and Jungclaus, J.H.: Simulation of the East Asian Summer Monsoon during the Last Millennium with the MPI Earth System Model, Journal of Climate, 25, 7852-7866, 2012.

Man, W., Zhou, T. and Jungclaus, J.H.: Effects of large volcanic eruptions on global summer climate and East Asian Monsoon changes during the Last Millennium: Analysis of MPI-ESM simulations, Journal of Climate, 27, 7394-7409, 2014.

Man, W. M. and Zhou, T.J.: Regional-scale surface air temperature and East Asian summer monsoon changes during the last millennium simulated by the FGOALS-gl climate system model. Adv. Atmos. Sci., 31(4), 765-778, 2014.

Marcott, S. A., Bauska, T. K., Buizert, C., Steig, E. J., Rosen, J. L., Cuffey, K. M., Fudge, T. J., Severinghaus, J. P., Ahn, J., Kalk, M. L., McConnell, J. R., Sowers, T., Taylor, K. C., White, J. W. C., and Brook, E. J.: Centennial-scale changes in the global carbon cycle during the last deglaciation, Nature, 514, 616–619, doi:10.1038/nature13799, 2014.

MARGO Project Members: Constraints on the magnitude and patterns of ocean cooling at the Last Glacial Maximum, Nature Geosci., 2, 127-132, 2009.

Masson-Delmotte, V., Schulz, M., Abe-Ouchi, A., Beer, J., Ganopolski, A., González Rouco, J., F., Jansen, E., Lambeck, K., Luterbacher, J., Naish, T., Osborn, T., Otto-Bliesner, B., Quinn, T., Ramesh, R., Rojas, M., Shao, X., and Timmermann, A.: Information from Paleoclimatic archives. In: Climate Change 2013: The physical 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.

Matthes, K., Funke, B., Andersson, M. E., Barnard, L., Beer, J., Charbonneau, P., Clilverd, M. A., Dudok de Wit, T., Haberreiter, M., Hendry, A., Jackman, C. H., Kretzschmar, M., Kruschke, T., Kunze, M., Langematz, U., Marsh, D. R., Maycock, A. C., Misios, S., Rodger, C. J., Scaife, A. A., Seppälä, A., Shangguan, M., Sinnhuber, M., Tourpali, K., Usoskin, I., van de Kamp, M., Verronen, P. T., and Versick, S.: Solar forcing for CMIP6 (v3.2), Geosci. Model Dev., 10, 2247-2302, https://doi.org/10.5194/gmd-10-2247-2017, 2017.

Mauri, A., Davis, B. A. S., Collins, P. M., and Kaplan, J. O.: The influence of atmospheric circulation on the mid-Holocene climate of Europe: a data–model comparison, Clim. Past, 10, 1925-1938, doi:10.5194/cp-10-1925-2014, 2014.

McGee, D., deMenocal, P. B., Winckler, G., Stuut, J. B. W., and Bradtmiller, L. I.: The magnitude, timing and abrupt-ness of changes in North African dust deposition over the last 20,000yr, Earth Planet. Sci. Lett., 371–372, 163–176, doi:10.1016/j.epsl.2013.03.054, 2013.

Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., Fraser, P. J., Montzka, S. A., Rayner, P. J., Trudinger, C. M., Krummel, P. B., Beyerle, U., Canadell, J. G., Daniel, J. S., Enting, I. G., Law, R. M., Lunder, C. R., O'Doherty, S., Prinn, R. G., Reimann, S., Rubino, M., Velders, G. J. M., Vollmer, M. K., Wang, R. H. J., and Weiss, R.: Historical greenhouse gas concentrations for climate modelling (CMIP6), Geosci. Model Dev., 10, 2057-2116, https://doi.org/10.5194/gmd-10-2057-2017, 2017.

Miller, K. G., Wright, J. D., Browning, J. V., Kulpecz, A., Kominz, M., Naish, T. R., Cramer, B. S., Rosenthal, Y., Peltier, W. R., and Sosdian, S.: High tide of the warm Pliocene: Implications of global sea level for Antarctic deglaciation, Geology, G32869, doi:10.1130/G32869.1, 2012.

Naish, T.R., Wilson, G.S.: Constraints on the amplitude of Mid-Pliocene (3.6–2.4Ma) eustatic sea-level fluctuations from the New Zealand shallow-marine sediment record. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 367(1886), 169-187, 2009.

Notz, D., Jahn, A., Holland, M., Hunke, E., Massonnet, F., Stroeve, J., Tremblay, B., and Vancoppenolle, M.: The CMIP6 Sea-Ice Model Intercomparison Project (SIMIP): understanding sea ice through climate-model simulations, Geosci. Model Dev., 9, 3427-3446, https://doi.org/10.5194/gmd-9-3427-2016, 2016.

Nowicki, S. M. J., Payne, A., Larour, E., Seroussi, H., Goelzer, H., Lipscomb, W., Gregory, J., Abe-Ouchi, A., and Shepherd, A.: Ice Sheet Model Intercomparison Project (ISMIP6) contribution to CMIP6, Geosci. Model Dev., 9, 4521-4545, https://doi.org/10.5194/gmd-9-4521-2016, 2016.

Otto-Bliesner, B., E. Brady, J. Fasullo, A. Jahn, L. Landrum, S. Stevenson, N. Rosenbloom, A. Mai, and G. Strand: Climate Variability and Change since 850 C.E.: An Ensemble Approach with the Community Earth System Model (CESM). Bull. Amer. Meteor. Soc., 97, 735-754, doi:10.1175/BAMS-D-14-00233.1, 2016.

Otto-Bliesner, B. L., Braconnot, P., Harrison, S. P., Lunt, D. J., Abe-Ouchi, A., Albani, S., Bartlein, P. J., Capron, E., Carlson, A. E., Dutton, A., Fischer, H., Goelzer, H., Govin, A., Haywood, A., Joos, F., Legrande, A. N., Lipscomb, W. H., Lohmann, G., Mahowald, N., Nehrbass-Ahles, C., Pausata, F. S.-R., Peterschmitt, J.-Y., Phipps, S., and Renssen, H.: The PMIP4 contribution to CMIP6 – Part 2: Two Interglacials, Scientific Objective and Experimental Design for Holocene and Last Interglacial Simulations, Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-279, in review, 2016.

Pagani, M., Liu, Z. H., LaRiviere, J., and Ravelo, A. C.: High Earth-system climate sensitivity determined from Pliocene carbon dioxide concentrations, Nature Geoscience, 3, 27-30, 2010.

PAGES2k – PMIP3 group: Continental-scale temperature variability in PMIP3 simulations and PAGES 2k regional temperature reconstructions over the past millennium. Clim. Past, 11, 1-27, 2015.

Pausata, F.S.R., G. Messori and Q. Zhang: Impacts of dust reduction on the northward expansion of the African monsoon during the Green Sahara period, Earth and Planetary Science Letters, 434, 298-307, doi:10.1016/j.epsl.2015.11.04, 2016.

Peltier, W. R., Argus, D. F. and Drummond, R.: Space geodesy constrains ice age terminal deglaciation: The global ICE-6G_C (VM5a) model, J. Geophys. Res. Solid Earth, 2014JB011176, doi:10.1002/2014JB011176, 2015.

Perez-Sanz, A., Li, G., González-Sampériz, P., and Harrison, S. P.: Evaluation of modern and mid-Holocene seasonal precipitation of the Mediterranean and northern Africa in the CMIP5 simulations, Clim. Past, 10, 551-568, doi:10.5194/cp-10-551-2014, 2014.

Phillips, A. S., Deser, C., and Fasullo, J.: Evaluating Modes of Variability in Climate Models. Eos, Transactions American Geophysical Union, 95(49), 453–455. http://doi.org/10.1002/2014EO490002, 2014.

Commentaire [mk3]: to be updated with final ref

Phipps, S.J., McGregor, H,V., Gergis, J., Gallant, A.J., Neukom, R., Stevensons, S., and Van Ommen, T.D.: Paleoclimate data-model comparison and the role of climate forcings over the past 1500 years. J. Climate, 26, 6915-6936, doi:10.1175/JCLI-D-12-00108.1, 2013.

Pincus, R., Forster, P. M., and Stevens, B.: The Radiative Forcing Model Intercomparison Project (RFMIP): experimental protocol for CMIP6, Geosci. Model Dev., 9, 3447-3460, doi:10.5194/gmd-9-3447-2016, 2016.

Pollard, D. and Reusch, D. B.: A calendar conversion method for monthly mean paleoclimate model output with orbital forcing, Journal of Geophysical Research: Atmospheres, 107, ACL 3-1-ACL 3-7, 2002.

Prentice, I.C., Kelley, D.I., Foster, P.N. Friedlingstein, P., Harrison, S.P. and Bartlein P.J.: Modeling fire and the terrestrial carbon balance. *Global Biogeochemical Cycles* 25 GB3005, doi:10.1029/2010GB003906, 2011.

Raymo, M. E., Hearty, P., De Conto, R., O' Leary, M., Dowsett, H. J., Robinson, M. M., and Mitrovica, J. X.: PLIOMAX: Pliocene maximum sea level project PAGES News. 17(2), pp.58-59, 2009.

Raymo, M. E., Grant, B., Horowitz, M., and Rau, G. H.: Mid-Pliocene warmth: Stronger greenhouse and stronger conveyor, Marine Micropaleontology, 27, 313-326, 1996.

Roche, D. M., Crosta, X., Renssen, H. : Evaluating Southern Ocean sea-ice for the Last Glacial Maximum and pre-industrial climates : PMIP-2 models and data evidence, Quaternary Science Reviews, 56, 99-106, 2012.

Rohling, E. J., Foster, G. L., Grant, K. M., Marino, G., Roberts, A. P., Tamisiea, M. E., and Williams, F.: Sea-level and deep-sea-temperature variability over the past 5.3 million years, Nature, 508, 477-482, 2014.

Salzmann, U. et al.: A new global biome reconstruction and data-model comparison for the Middle Pliocene, Global Ecology and Biogeography, 17,432-447, 2008.

Salzmann, U., Dolan, A.M., Haywood, A M., Chan W.-L., Hill, D.J., Abe-Ouchi, A., Otto-Bliesner, B., Bragg, F., Chandler, M. A., Contoux, C., Dowsett, H.J., Jost, A., Kamae, Y., Lohmann, Lunt, D. J., Pickering, S.J., Pound M.J., Ramstein, G., Rosenbloom, N.A., Sohl, L., Stepanek, C., Ueda, H, Zhang, Z.: Challenges in reconstructing terrestrial warming of the Pliocene revealed by data-model discord. Nature Climate Change 3, 969-974, 2013.

Saint-Lu, M., Braconnot, P., Leloup, J., Lengaigne, M., & Marti, O.: Changes in the ENSO/SPCZ relationship from past to future climates. Earth and Planetary Science Letters, 412, 18–24. http://doi.org/10.1016/j.epsl.2014.12.033, 2015.

Schmidt, G.A., Jungclaus, J.H., Ammann, C.M., Bard, E., Braconnot, P., Crowley, T.J., Delaygue, G., Joos, F., Krivova, N.A., Muscheler, R., Otto-Bliesner, B.L., Pongratz, J., Shindell, D.T., Solanki, S.K., Steinhilber, F., and Vieira, L.E.A.: Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1.0). Geosci. Model Dev., 4, 33-45, doi: 10.5194/gmd-4-33-2011, 2011.

Schmidt, G.A., Jungclaus, J.H., Ammann, C.M., Bard, E., Braconnot, P., Crowley, T.J., Delaygue, G., Joos, F., Krivova, N.A., Muscheler, R., Otto-Bliesner, B.L., Pongratz, J., Shindell, D.T., Solanki, S.K., Steinhilber, F., and Vieira, L.E.A.: Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1.1). Geosci. Model Dev., 5, 185-191, doi: 10.5194/gmd-5-185-2012, 2012.

Schmidt, G. A., Annan, J. D., Bartlein, P. J., Cook, B. I., Guilyardi, E., Hargreaves, J. C., Harrison, S. P., Kageyama, M., LeGrande, A. N., Konecky, B., Lovejoy, S., Mann, M. E., Masson-Delmotte, V., Risi, C., Thompson, D., Timmermann, A., Tremblay, L.-B., and Yiou, P.: Using palaeo-climate comparisons to constrain future projections in CMIP5, Clim. Past, 10, 221-250, doi:10.5194/cp-10-221-2014, 2014.

Schurer, A.P., Tett, S.F.B., and Hegerl, G.C.: Small influence of solar variability on climate over the last millennium. Nature Geoscience, 7, 104-108, doi:10.1038/NGEO2-40, 2014.

Seki, O., Foster, G. L., Schmidt, D. N., Mackensen, A., Kawamura, K., and Pancost, R. D.: Alkenone and boron-based Pliocene pCO2 records, Earth and Planetary Science Letters, 292, 201-211, 2010.

Sherwood, S.C., Bony, S., and Dufresne, J. L.: Spread in model climate sensitivity traced to atmospheric convective mixing, *Nature*, 505, 37-42, 2014.

Sigl, M., Winstrup, M., McConnell, J.R., Welten, K.C., Plunkett, G., Ludlow, F., Büntgen, U., Caffee, M., Chellman, N., Dahl-Jensen, D., Fischer, H., Kipfstuhl, S., Kostick, C., Maselli, O.J., Mekhaldi, F., Mulvaney, R., Muscheler, R., Pasteris, D.R., Pilcher, J.R., Salzer, M., Schüpbach, S., Steffensen, J.P., Vinther, B.M., and Woodruff, T.E.: Timing and climate forcing of volcanic eruptions for the past 2,500 years", Nature, vol. 523, pp. 543-549, 2015. http://dx.doi.org/10.1038/nature14565, 2015.

Spratt, R. M. and Lisiecki, L. E.: A Late Pleistocene sea level stack, Clim. Past, 12, 1079-1092, doi:10.5194/cp-12-1079-2016, 2016.

Stap, L. B., de Boer, B., Ziegler, M., Bintanja, R., Lourens, L. J., and van de Wal, R. S. W.: CO2 over the past 5 million years: Continuous simulation and new δ11B-based proxy data, Earth and Planetary Science Letters, 439, 1-10, 2016.

Steffensen, J. P., Andersen, K. K., Bigler, M., Clausen, H. B., Dahl-Jensen, D., Fischer, H., Goto-Azuma, K., Hansson, M., Johnsen, S. J., Jouzel, J., Masson-Delmotte, V., Popp, T., Rasmussen, S. O., Rothlisberger, R., Ruth, U., Stauffer, B., Siggaard-Andersen, M. L., Sveinbjornsdottir, A. E., Svensson, A., and White, J. W. C.: High-Resolution Greenland Ice Core Data Show Abrupt Climate Change Happens in Few Years, Science, 321, 680-684, 2008.

Steig, E.J., Huybers, K., Singh, H.A., Steiger, N.J., Ding, Q.H., Frierson, D.M.W., Popp, T. and White, J.W.C.: Influence of West Antarctic Ice Sheet collapse on Antarctic surface climate. Geophysical Research Letters, 42(12): 4862-4868, 2015.

Stevenson, S., Fox-Kemper, B., Jochum, M., Rajagopalan, B., & Yeager, S. G.: ENSO Model Validation Using Wavelet Probability Analysis, J. Climate, 23, 5540–5547, http://doi.org/10.1175/2010JCLI3609.1, 2010.

Swingedouw, D., J. Mignot, P. Braconnot, E. Mosquet, M. Kageyama, R. Alkama: Impact of freshwater release in the North Atlantic under different climate conditions in an OAGCM, Journal of Climate, 22, 6377-6403, https://doi.org/10.1175/2009JCLI3028., 2009.

Tarasov, L., Dyke, A. S., Neal, R. M. and Peltier, W. R.: A data-calibrated distribution of deglacial chronologies for the North American ice complex from glaciological modeling, Earth Planet. Sci. Lett., 315–316, 30–40, doi:10.1016/j.epsl.2011.09.010, 2012.

Taylor, K.E., Stouffer, R.J., and Meehl, G. A.: An Overview of CMIP5 and the experiment design, Bull. Amer. Meteor. Soc., 93, 485-498, doi:10.1175/BAMS-D-11-00094.1, 2012.

Timm, O., Timmermann, A., Abe-Ouchi, A., Saito, F., and Segawa, T.: On the definition of seasons in paleoclimate simulations with orbital forcing, Paleoceanography, 23, doi:10.1029/2007PA001461 2008.

Toohey, M., Stevens, B., Schmidt, H., and Timmreck, C.: Easy Volcanic Aerosol (EVA v1.0): an idealized forcing generator for climate simulations, Geosci. Model Dev., 9, 4049-4070, https://doi.org/10.5194/gmd-9-4049-2016, 2016.

Toohey, M. and Sigl, M.: Volcanic stratospheric sulphur injections and aerosol optical depth from 500 BCE to 1900 CE, Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-31, in review, 2017.

Tripati, A. K., Roberts, C. D., and Eagle, R. A.: Coupling of CO2 and Ice Sheet Stability Over Major Climate Transitions of the Last 20 Million Years, Science, 326, 1394-1397, 2009.

Turney, C. S. M. and Jones, R. T.: Does the Agulhas Current amplify global temperatures during super-interglacials? J. Quat. Sci. 25, 839–843, doi:10.1002/jqs.1423, 2010.

Usoskin, I.G., Gallet, Y., Lopes, F., Kovaltsov, G.A., and Hulot, G.: Solar activity during the Holocene: the Hallstatt cycle and its consequence for grand minima and maxima. Astron. and Astroph., 587, A150, 2016.

van den Hurk, B., Kim, H., Krinner, G., Seneviratne, S. I., Derksen, C., Oki, T., Douville, H., Colin, J., Ducharne, A., Cheruy, F., Viovy, N., Puma, M. J., Wada, Y., Li, W., Jia, B., Alessandri, A., Lawrence, D. M., Weedon, G. P., Ellis, R., Hagemann, S., Mao, J., Flanner, M. G., Zampieri, M., Materia, S., Law, R. M., and Sheffield, J.: LS3MIP (v1.0) contribution to CMIP6: the Land Surface, Snow and Soil moisture Model Intercomparison Project – aims, setup and expected outcome, Geosci. Model Dev., 9, 2809-2832, doi:10.5194/gmd-9-2809-2016, 2016.

Vieira, L.E.A., Solanki, S.K., Krivova, N.A., and Usoskin, I.G.: Evolution of the solar irradiance during the Holocene. Astron. Astroph., 531, A6., 2011.

Wardlaw, B.R. and Quinn, T.M.: The record of Pliocene sea-level change at Enewetak Atoll. Quaternary Science Reviews. 10(2-3), pp.247-258, 1991.

Webb, M. J., Andrews, T., Bodas-Salcedo, A., Bony, S., Bretherton, C. S., Chadwick, R., Chepfer, H., Douville, H., Good, P., Kay, J. E., Klein, S. A., Marchand, R., Medeiros, B., Siebesma, A. P., Skinner, C. B., Stevens, B., Tselioudis, G., Tsushima, Y., and Watanabe, M.: The Cloud Feedback Model Intercomparison Project (CFMIP) contribution to CMIP6, Geosci. Model Dev., 10, 359-384, https://doi.org/10.5194/gmd-10-359-2017, 2017.

Winsor, K., Carlson, A.E., Klinkhammer, G.P., Stoner, J.S. and Hatfield, R.G.: Evolution of the northeast Labrador Sea during the last interglaciation. Geochemistry Geophysics Geosystems, 13, 2012.

Wittenberg, A. T.: Are historical records sufficient to constrain ENSO simulations? Geophys. Res. Lett., 36, L12702, doi:10.1029/2009GL038710, 2009.

Yoshimori, M., Hargreaves, J. C., Annan, J. D., Yokohata, T. and Abe-Ouchi, A.: Dependency of feedbacks on forcing and climate state in physics parameter ensembles, Journal of Climate, 24, 6440-6455, 2011.

Zanchettin, D., Khodri, M., Timmreck, C., Toohey, M., Schmidt, A., Gerber, E. P., Hegerl, G., Robock, A., Pausata, F. S. R., Ball, W. T., Bauer, S. E., Bekki, S., Dhomse, S. S., LeGrande, A. N., Mann, G. W., Marshall, L., Mills, M., Marchand, M., Niemeier, U., Poulain, V., Rozanov, E., Rubino, A., Stenke, A., Tsigaridis, K., and Tummon, F.: The Model Intercomparison Project on the climatic response to Volcanic forcing (VolMIP): experimental design and forcing input data for CMIP6, Geosci. Model Dev., 9, 2701-2719, doi:10.5194/gmd-9-2701-2016, 2016.

Zhang, R., Yan, Q., Zhang, Z. S., Jiang, D., Otto-Bliesner, B. L., Haywood, A. M., Hill, D. J., Dolan, A. M., Stepanek, C., Lohmann, G., Contoux, C., Bragg, F., Chan, W.-L., Chandler, M. A., Jost, A., Kamae, Y., Abe-Ouchi, A., Ramstein, G., Rosenbloom, N. A., Sohl, L., and Ueda, H.: Mid-Pliocene East Asian monsoon climate simulated in the PlioMIP, Clim. Past, 9, 2085-2099, doi:10.5194/cp-9-2085-2013, 2013.

Zheng, W., Braconnot, P.: Characterization of model spread in PMIP2 Mid-Holocene simulations of the African Monsoon, Journal of Climate, 26, 1192-1210, 2013.

Zheng, W., P. Braconnot, E. Guilyardi, U. Merkel, and Y. Yu: ENSO at 6ka and 21ka from ocean-atmosphere coupled model simulations, Climate Dynamics, 30, 745-762, doi:DOI 10.1007/s00382-007-0320-3, 2008

TABLES

Period	Purpose	CMIP6 Priority
Last millennium (*past1000*) 850-1849 CE	a) Evaluate the ability of models to capture reconstructed 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) Provide a longer-term perspective for detection and attribution studies.	Tier 1*
Mid-Holocene (*midHolocene*) 6 kyr ago	a) Compare the model response to known orbital forcing changes and changes in greenhouse gas concentrations to paleodata, describing major temperature and hydrological changes. b) Relationships between changes in mean state and variability	Tier 1* PMIP4-CMIP6 entry card
Last Glacial Maximum (*lgm*) 21 kyr ago	a) Compare the model response to ice-age boundary conditions with paleodata. b) Attempt to provide empirical constraints on global climate sensitivity.	Tier 1* PMIP4-CMIP6 entry card
Last Interglacial (*lig127k*) 127 kyr ago	a) Evaluate climate model for warm period in northern hemisphere and high sea-level stand. b) Impacts of this climate on sea ice and ice sheets.	Tier 1*
Mid-Pliocene Warm Period (midPlioceneEoi400) 3.2 Ma ago	a) Earth System response to a long term to CO_2 forcing analogous to that of the modern. b) Significance of CO_2-induced polar amplification for the stability of the ice sheets, sea-ice and sea-level.	Tier 1*

Table 1: Characteristics, purpose and CMIP6 priority of the five PMIP4-CMIP6 experiments. * All experiments can be run independently. It is not mandatory to perform all Tier 1 experiments to take part in PMIP4-CMIP6, but it is mandatory to run at least one of the PMIP4-CMIP6 entry cards.

Period	GHG	Astronomical parameters	Ice-sheets	Tropospheric aerosols *	Land surface**	Volcanoes	Solar activity	Reference to be cited
PMIP4-CMIP6 entry cards								
Mid-Holocene (*midHolocene*) 6 ky ago	CO_2: 264.4 ppm CH_4: 597 ppb N_2O: 262 ppb CFC : 0 O_3: pre-industrial	6 kyr BP	as in PI	modified (if possible)	Interactive vegetation OR Interactive carbon cycle OR fixed to present day (depending on model complexity)	as in PI	as in PI	Otto-Bliesner et al, 2017
Last Glacial Maximum (*lgm*) 21 ky ago	CO_2: 190 ppm CH_4: 375 ppm N_2O: 200 ppb CFC 0 O_3: pre-industrial	21 kyr BP	modified (larger)	modified (if possible)	Interactive vegetation OR Interactive carbon cycle OR fixed to present day (depending on model complexity)	as in PI	as in PI	Kageyama et al, 2017
Tier 1 PMIP4-CMIP6 experiments								
Last millennium (*past1000*) 850-1849 CE	Time varying (Meinshausen et al., 2017)	time varying (Berger 1978, Schmidt et al., 2011)	as in PI	as in PI	time varying (land use)	time varying radiative forcing due to stratospheric aerosols	time varying	Jungclaus et al, 2017
Last Interglacial (*lig127k*) 127 ky ago	CO_2: 275 ppm CH_4: 685 ppb N_2O: 255 ppb CFC 0 O_3: pre-industrial	127 ky BP	as in PI	modified (if possible)	Interactive vegetation OR Interactive carbon cycle OR fixed to present day (depending on model complexity)	as in PI	as in PI	Otto-Bliesner et al, 2017
Mid-Pliocene Warm Period (midPlioceneEoi400) 3.2 My ago	CO_2: 400 ppm	as in PI	modified (smaller)	as in PI	Interactive vegetation OR Modified to mid-Pliocene OR fixed to present day (depending on model complexity)	as in PI	as in PI	Haywood et al, 2016

5 **Table 2: summary of change in boundary conditions with respect to *piControl* (abbreviated as "PI") for each PMIP4-CMIP6 experiment * Only for models without fully interactive dust (see section 3.3). ** interactive carbon cycle, with computation of some characteristics of the vegetation such as the leaf area index (LAI), but without full vegetation dynamics.**

Atmospheric variables	top of atmosphere energy budget (global and annual average)
	surface energy budget (global and annual average)
	northern surface air temperature (annual average over northern hemisphere)
	southern surface air temperature (annual average over southern hemisphere)
Oceanic variables	Sea surface temperatures (global and annual average)
	deep ocean temperatures (global and annual average over depths below 2500m)
	deep ocean salinity (global and annual average over depths below 2500m)
	Atlantic Meridional Overturning Circulation (maximum overturning between 0 and 80°N and below 500 m depth)
Sea ice variables	northern sea-ice (annual average over northern hemisphere)
	southern sea-ice (annual average over southern hemisphere)
Carbon cycle variables	Global carbon budget

5   **Table 3: Variables to be saved for the documentation of the spin-up phase of the models.**

Reference	Variables	Time period	Comments	Data available from
Mann et al. (2009)	MAT	500-2006 CE	Gridded data set (5°)	http://science.sciencemag.org/content/suppl/2009/11/25/326.5957.1256.DC1
PAGES 2k Consortium (2013)	MAT	past 2000 years	Individual sites; Arctic data updated 2014	https://www.ncdc.noaa.gov/cdo/f?p=519:1:::::P1_STUDY_ID:12621
Bartlein et al. (2011)	MAT, MAP, α, MTCO, MTWA	6000±500 yr BP; 21000±1000 yr BP	Gridded data set (2°)	https://www.ncdc.noaa.gov/paleo/study/9897
MARGO Project Members (2009)	Mean annual, winter, summer SST	21000±2000 yr BP	Gridded data set (5°)	http://www.ncdc.noaa.gov/paleo/study/12034 http://doi.pangaea.de/10.1594/PANGAEA.733406
Turney and Jones (2010)	MAT, SST	Maximum warmth during LIG	Individual sites (100 terrestrial; 162 marine)	http://onlinelibrary.wiley.com/store/10.1002/jqs.1423/asset/supinfo/JQS_1423_sm_suppInfo.pdf?v=1&s=1726938c44b8762e15aaf17514fc076c855b8ed1
Capron et al. (2014); Capron et al. (subm.)	MAT, summer SST	114-116ka, 119-121ka, 124-126ka, 126-128 ka, 129-131ka	47 high latitude sites	doi.pangaea.de/10.1594/PANGAEA.841672
Dowsett et al, (2012)	SST	3.264-3.025 Ma	Further information available in Dowsett et al. (2016)	http://www.nature.com/nclimate/journal/v2/n5/full/nclimate1455.html#supplementary-information
Salzmann et al., 2013	MAT	3.3-3.0 Ma		http://www.nature.com/nclimate/journal/v3/n11/extref/nclimate2008-s1.pdf

Table 4: Examples of data syntheses for the PMIP4-CMIP6 periods. MAT: Mean Annual Temperature, MAP: Mean Annual Precipitation, α: ratio of the actual evaporation over potential evaporation, MTCO: Mean Temperature of the Coldest Month, MTWA: Mean Temperature of the Warmest Month, SST: Sea Surface Temperature.

MIP abbreviaton	MIP full name	Themes of interactions
CF-MIP	Cloud Feedback Model Intercomparison Project	dedicated common idealized sensitivity experiment to be run in aquaplanet set up, *AMIPminus4K*, to be co-analysed in CF-MIP and PMIP.
ISMIP6	Ice Sheet Model Intercomparison Project for CMIP6	Assessment of the climate and cryosphere interactions and the sea level changes associated with large ice sheets. In particular, the *lig127k* simulation 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.
OMIP	Ocean Model Intercomparison Project	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.
SIMIP	Sea Ice Model Intercomparison Project	Assessment of role of sea-ice in climate changes
AerChemMIP	Aerosols and Chemistry Model Intercomparison Project	Assessment of role of aerosols in climate changes, very helpful since this is a new aspect in PMIP experiments for the midHolocene, last interglacial and LGM
LS3MIP	Land Surface, Snow and Soil Moisture Model Intercomparison Project	Assessment of role of land surface processes in climate changes.
C4MIP	Coupled Climate Carbon Cycle Model Intercomparison Project	Assessment of carbon-cycle evolution and feedbacks between sub-components of the Earth System. Evaluation of paleo reconstructions of carbon storage.
LUMIP	Land-Use Model Intercomparison Project	Analysis of climate changes associated with Land Use changes (*past1000* experiment)
VolMIP	Volcanic Forcings Model Intercomparison Project	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 *past1000* 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.
DAMIP	Detection and Attribution Model Intercomparison Project	*past1000* simulations provide long-term reference background including natural climate variability for detection and attribution.
RFMIP	Radiative Forcing Model Intercomparison Project	Compare radiative forcing from LGM GHG as computed by climate models and by off-line fine-scale radiative transfer codes.

5   **Table 5:interactions of PMIP with other CMIP6 MIPs**

 FIGURE CAPTIONS

Figure 1: Context of the PMIP4 experiments (from left to right: mPWP, Mid-Pliocene Warm Period; LIG, last interglacial; LGM, last glacial maximum; MH, mid-Holocene; LM, last millennium; H, CMIP6 historical simulation): (a)-(d) insolation anomalies (differences from 1950 CE), for July at 65°N, calculated using the programs of Laskar et al. (2004, panel (a)) and Berger (1978, panels (b)-(d)); (e) $\delta^{18}O$ (magenta, Lisiecki and Raymo, 2005, scale at left), and sea level (blue line, Rohling et al., 2014; blue shading, a density plot of eleven Mid-Pliocene sea level estimates (Dowsett and Cronin 1990; Wardlaw and Quinn, 1991; Krantz, 1991; Raymo et al., 2009; Dwyer and Chandler, 2009; Naish and Wilson, 2009; Masson-Delmotte et al., 2013; Rohling et al., 2014; Dowsett et al., 2016) scale at right); (f) and (g) $\delta^{18}O$ (magenta, Lisiecki and Raymo, 2005, $\delta^{18}O$ scale at left), and sea level (blue dots, with light-blue 2.5, 25, 75 and 97.5 percentile bootstrap confidence intervals, Spratt and Lisiecki, 2015; blue rectangle, LIG high-stand range, Dutton et al., 2015; dark blue lines, Lambeck et al., 2014, sea-level scale at right on panel (g)), (h) sea level (Kopp, et al., 2016, scale at right); (i) CO_2 for the interval 3.0-3.3 Ma shown as a density plot of eight Mid-Pliocene estimates (Raymo et al., 1996; Stap et al., 2016; Pagani et al., 2010; Seki et al., 2010; Tripati et al., 2009; Bartoli et al., 2011; Seki et al., 2010; Kurschner et al., 1996); (j) and (k) CO_2 measurements (Bereiter et al., 2015, scale at left); (l) CO_2 measurements (Schmidt et al, 2011, scale at right); (m) and (n) CH_4 measurements (Loulergue et al., 2008, scale at left); (o) CH_4 measurements (Schmidt et al, 2011, scale at right); (p) volcanic radiative forcing (Schmidt et al., 2012, scale at right); (q) total solar irradiance (Schmidt et al., 2012, scale at right).

Figure 2: Changes in boundary conditions related to changes in ice sheets for the midPliocene-eoi400 (top) and lgm (middle: ICE-6G_C and bottom: GLAC-1D) experiments. Coastlines for palaeo-period shown as brown contours. Ice sheet boundaries for each period shown as red contour. Bright shading: changes in altitude over regions covered by ice sheets during the considered palaeo-period. Faded shading: changes in altitude over ice-free regions.

Figure 3: Maps of dust deposition (g m-2 a-1) simulated with the Community Earth System Model for the a. PI (pre-industrial) (Albani et al., 2016), b. MH (Mid-Holocene) (Albani et al., 2015), and c. LGM (Albani et al., 2014). Maps of dust deposition (g m-2 a-1) for the LGM d. simulated with the Hadley Centre Global Environment Model 2-Atmosphere (Hopcroft et al, 2015), and reconstructed from a global interpolation of paleodust data (Lambert et al., 2015).

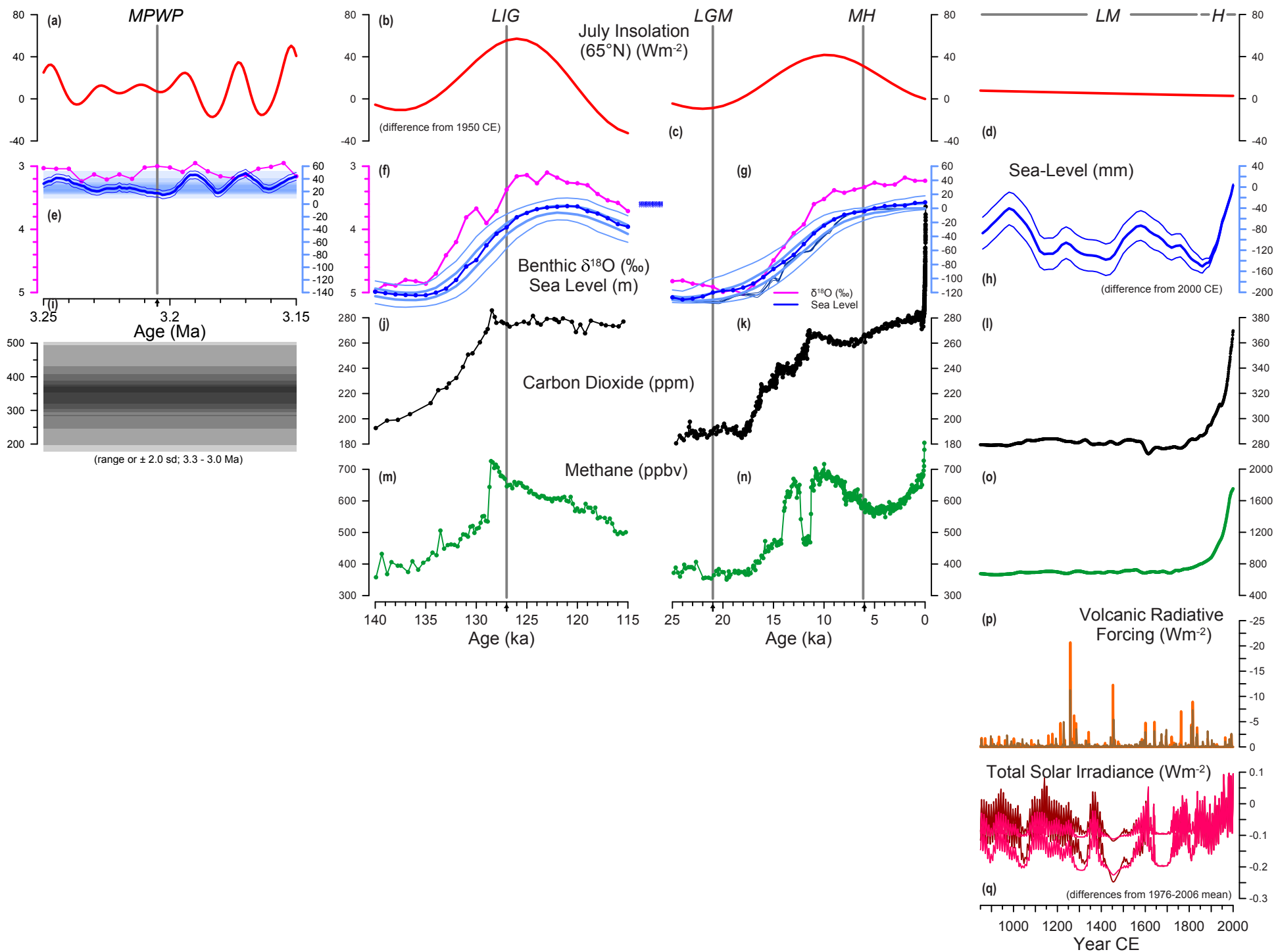
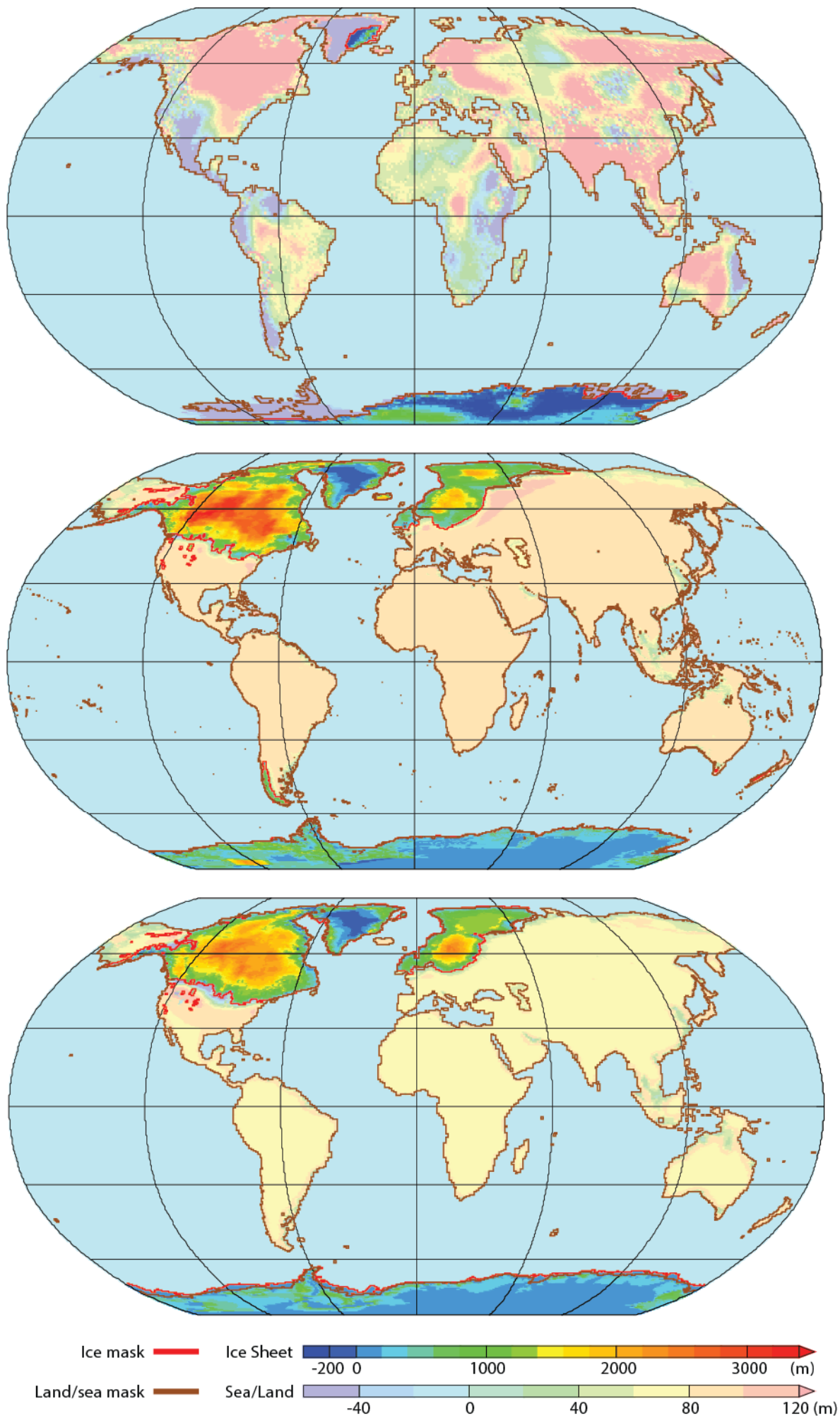
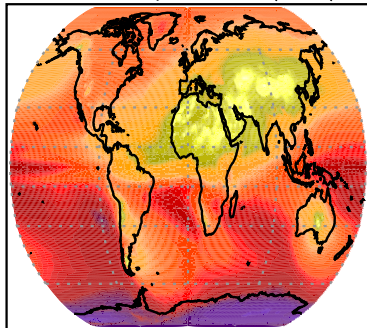
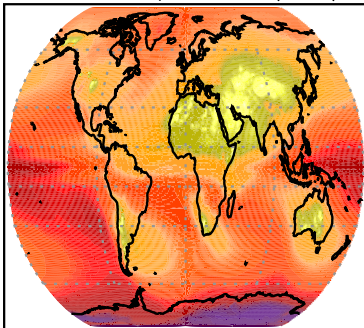
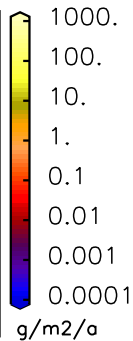
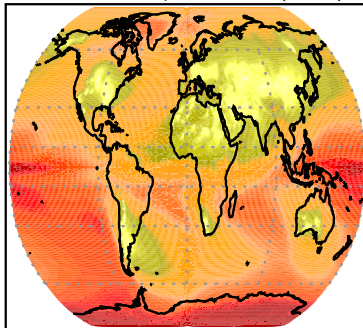
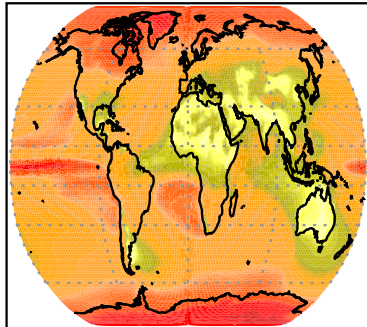
Figure 4: The calendar effect: (a) month-length anomalies, 140 ka to present, with the PMIP4 experiment times indicated by vertical lines. The month-length anomalies were calculated using the formulation in Kutzbach and Gallimore (1988). (b and c) The calendar effect on October temperature at 6 and 127 ka, calculated using Climate Forecast System Reanalysis near-surface air temperature (https://www.earthsystemcog.org/ projects/obs4mips/), 1981-2010 long-term means, and assuming the long-term mean differences in temperature are zero everywhere. (e and f) The calendar effect on October precipitation at 6 and 127 ka, calculated using the CPC Merged Analysis of Precipitation (CMAP) enhanced precipitation (http://www.esrl.noaa.gov/ psd/data/gridded/ data.cmap.html), 1981-2010 long-term means, and again assuming that the long-term mean differences in temperature are zero everywhere. Calendar effects were calculated by interpolating present-day monthly temperature or precipitation to a daily time step as in Pollard and Reusch, 2002 (but using a mean-preserving algorithm for pseudo-daily interpolation for monthly values; Epstein, 1991), and then recalculating the monthly means using the appropriate paleo calendar (Bartlein and Shafer, 2016). Note that the 6 and 127 ka map patterns for both variables, while broadly similar, are not simply rescaled versions of one another.

5   Figure 5: the PMIP4-CMIP6 experiments in the framework of CMIP6, with associated MIPs, and in the framework of PMIP4, with its working groups.

Figure 1

Ice mask ———
Land/sea mask ———

Ice Sheet
-200 0   1000   2000   3000   (m)

Sea/Land
-40   0   40   80   120 (m)

a. PI Dust deposition flux (Albani)   b. MH Dust deposition flux (Albani)   c. LGM Dust deposition flux (Albani)

d. LGM Dust deposition flux (Hopcroft)   e. LGM Dust deposition flux (Lambert)

1000.
100.
10.
1.
0.1
0.01
0.001
0.0001

g/m2/a

Figure 3