Clim. Past Discuss., 9, 775–835, 2013
www.clim-past-discuss.net/9/775/2013/
doi:10.5194/cpd-9-775-2013

Climate of the Past
Discussions
Open Access

Using paleo-climate comparisons to constrain future projections in CMIP5

G. A. Schmidt[1], J. D. Annan[2], P. J. Bartlein[3], B. I. Cook[1], E. Guilyardi[4,5], J. C. Hargreaves[2], S. P. Harrison[6,7], M. Kageyama[8], A. N. LeGrande[1], B. Konecky[9], S. Lovejoy[10], M. E. Mann[11], V. Masson-Delmotte[8], C. Risi[12], D. Thompson[13], A. Timmermann[14], L.-B. Tremblay[10], and P. Yiou[8]

[1]NASA Goddard Institute for Space Studies, 2880 Broadway, New York, NY 10025, USA
[2]Research Institute for Global Change, JAMSTEC, Yokohama Institute for Earth Sciences, Yokohama, Japan
[3]University of Oregon, Eugene, OR 97403, USA
[4]NCAS-Climate, University of Reading, Whiteknights, P.O. Box 217, Reading, Berkshire, RG6 6AH, UK
[5]Laboratoire d'Océanographie et du Climat: Expérimentation et Approches Numériques/Institut Pierre Simon Laplace, UMR7617, CNRS-IRD-UPMC, 4 place Jussieu, 75252 Paris Cedex 05, France
[6]University of Reading, Whiteknights, P.O. Box 217, Reading, Berkshire, RG6 6AH, UK
[7]Macquarie University, Sydney, NSW 2109, Australia
[8]Laboratoire des Sciences du Climat et de l'Environnement (CEA-CNRS-UVSQ/Institut Pierre Simon Laplace), UMR8212, CE Saclay l'Orme des Merisiers, 91191 Gif-sur-Yvette, France
[9]Brown University, 324 Brook St., P.O. Box 1846, Providence, RI 02912, USA

[10]McGill University, 805 Sherbrooke Street West, Montreal, Quebec H3A 0B9, Canada
[11]Pennsylvania State University, University Park, PA 16802, USA
[12]Laboratoire de Météorologie Dynamique/Institut Pierre Simon Laplace, 4, place Jussieu, 75252 Paris Cedex 05, France
[13]University of Arizona, Department of Geosciences, Gould-Simpson Building #77, 1040 E 4th St. Tucson, AZ 85721, USA
[14]University of Hawaii, 2525 Correa Road, Honolulu, HI 96822, USA

Received: 11 January 2013 – Accepted: 30 January 2013 – Published: 11 February 2013

Correspondence to: G. A. Schmidt (gavin.a.schmidt@nasa.gov)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

We present a description of the theoretical framework and "best practice" for using the paleo-climate model component of the Coupled Model Intercomparison Project (Phase 5) (CMIP5) to constrain future projections of climate using the same models. The constraints arise from measures of skill in hindcasting paleo-climate changes from the present over 3 periods: the Last Glacial Maximum (LGM) (21 thousand years before present, ka), the mid-Holocene (MH) (6 ka) and the Last Millennium (LM) (850–1850 CE). The skill measures may be used to validate robust patterns of climate change across scenarios or to distinguish between models that have differing outcomes in future scenarios. We find that the multi-model ensemble of paleo-simulations is adequate for addressing at least some of these issues. For example, selected benchmarks for the LGM and MH are correlated to the rank of future projections of precipitation/temperature or sea ice extent to indicate that models that produce the best agreement with paleoclimate information give demonstrably different future results than the rest of the models. We also find that some comparisons, for instance associated with model variability, are strongly dependent on uncertain forcing timeseries, or show time dependent behaviour, making direct inferences for the future problematic. Overall, we demonstrate that there is a strong potential for the paleo-climate simulations to help inform the future projections and urge all the modeling groups to complete this subset of the CMIP5 runs.

1 Introduction

The Coupled Model Intercomparison Project (Phase 5) (CMIP5) is an ongoing coordinated project instigated by the Working Group on Coupled Modelling (WGCM) at the World Climate Research Programme (WCRP) and consisting of contributions from over 25 climate modeling groups (and over 30 climate models) from around the world (Taylor et al., 2012). Multiple experiments are being coordinated, including historical

CPD
9, 775–835, 2013

simulations for the 20th Century (starting from 1850), future simulations following multiple Representative Concentration Pathways (RCPs) and crucially, for the first time in CMIP, three sets of paleo-climate simulations for the Last Glacial Maximum (LGM) (21 K Before Present (BP)), the Mid-Holocene (MH) (6 K BP) and the Last Millennium (850–1850 CE). The paleo simulations are also part of the Paleoclimate Model Intercomparison Project (Phase 3) (PMIP3) initiative.

The CMIP5/PMIP3 paleo-simulations are true "out-of-sample" tests in that none of the models have been "tuned" to produce better paleo climates. This is not necessarily unwise (see Schneider von Deimling et al. (2006) for an example), but would complicate some of the potential analyses. Because the same models are being used for both past and future simulations, this dataset is a unique resource for research into the connections between model skill and model predictions, and has the potential to greatly improve assessments of future climate change.

There were many uncertainties in climate projections highlighted in the IPCC AR4 (Meehl et al., 2007). Many of these, such as the future of sub-tropical rainfall, El Niño/Southern Oscillation (ENSO) changes, potential declines in the North Atlantic meridional circulation, the fate of Arctic sea ice, etc. have important regional impacts. Reducing these uncertainties in the projections could have significant real world consequences for both adaptation and mitigation strategies. The three main classes of prediction uncertainty relate to: the choice of scenario, internal variability (sometimes described as initial condition uncertainty), and the imperfections in the model (or structural uncertainty) (Hawkins and Sutton, 2009). Scenario uncertainties inevitably grow in importance with time particularly after about 30 yr due to the time-scales associated with economic change, CO_2 residence time and ocean thermal inertia. Initial condition uncertainty is globally important on scales of a few years (and longer at smaller spatial scales) but predictability is fundamentally limited by the chaotic dynamics of the atmosphere and upper ocean. Thus at the multi-decadal time-horizon, reducing and/or better characterizing structural uncertainty is the only way to reduce overall uncertainty. These structural uncertainties (given a specific scenario of future emissions and other

CPD
9, 775–835, 2013

drivers) arise from a combination of model divergence – i.e. a large spread in model predictions given the same future scenario, and model inadequacy – i.e. models that are collectively either incomplete, inaccurate or are missing processes or feedbacks. The first effect is explicit (though not completely explored) in the multi-model ensemble, while the second is implicit and needs to be assessed independently.

Observations provide the means to potentially test the models and reduce these uncertainties, but unfortunately, instrumental records of useful data targets are few (essentially limited to situ networks of temperature and rainfall prior to the satellite era), and perhaps more importantly, changes in the recent past are relatively small compared to projections for the future. Furthermore, the majority of skill metrics in historical (20th Century) simulations do not constrain future projections: models that are either good or bad at simulating some aspect of modern climate – the climatology, seasonal cycle, or interannual variability – often give essentially the same spread of future projections (Santer et al., 2010; Knutti et al., 2010). Paleo-climate changes from the present offer a substantially larger signal and although paleo-climate records are often affected by substantial noise and difficulties in interpretation (Schmidt, 2010), the most robust reconstructions can provide a crucial test of model performance wider than the range of the 20th Century climates.

There has been much evaluation of paleo-climate simulations via earlier incarnations of PMIP, as well as many individual studies (see the review by Braconnot et al., 2012, and references therein). However, there has been a lack of analyses that quantitatively link future simulations or forecasts with skill or sensitivity in the paleo-climate simulations (though see Hargreaves et al., 2012b, for an example). Partly this is because (prior to CMIP5) paleo-simulations were not done with exactly the same versions of the models being used for future projections and partly through a lack of suitable skill metrics for paleo-climate change. This paper is therefore specifically not focused on understanding paleo-climate change for its own sake but rather is meant as a guide to the appropriate theoretical framework for quantitatively linking past and future that can be applied to data from the CMIP5 archive.

CPD
9, 775–835, 2013

For clarity in the rest of the text, we define the term "ensemble" to denote the full multi-model database of results across all scenarios (which here encompasses all paleoclimate, historical, idealized and future projection simulations). The future projections used here consist of the four RCP scenarios (rcp26, rcp45, rcp6, rcp85) (future possibilities that roughly produce radiative forcing at the year 2100 relative to 2000 of 2.6, 4.5, 6.0, and 8.5 $W m^{-2}$, respectively) along with idealised simulations have been included to provide clean comparisons across models (such as 1 % increasing CO_2 simulations, the response to an abrupt increase to $4 \times CO_2$, atmosphere-only simulations etc.). For ease of reference, we will use CMIP5 to refer to the entire database, including the PMIP3 simulations. Specific model simulations are referred to by their name in the CMIP5 database (i.e. rcp85, past1000, PIcontrol etc.), while the scenarios or periods when are referred to more generally using a standard abbreviation or name (e.g. the LGM, MH, RCP 4.5).

The scope of the paper is as follows: Sect. 2 discusses theoretical frameworks for dealing with the multi-model ensemble, issues arising from the use of paleo-proxy data and the use of data-synthesis products; Sects. 3 and 4 discuss specific examples of skill metrics that may have predictive power in future simulations by showing robust behaviour across paleo and future experiments, or discriminate between future projections. Sect. 5 presents some exploratory analysis of additional potentially useful metrics that either diverge over time or are too sensitive to important uncertainties; Sect. 6 concludes and discusses the potential for further work in this area. We list the models that we have used in analyses in this paper, along with the specific experiments and simulation IDs, in Table 1.

CPD
9, 775–835, 2013

2 Methodologies

2.1 Palaeoclimate reconstructions

Many of the problems in dealing with reconstructing climate from paleodata are specific to the type of record, the time period and resolution concerned – for instance, annually resolved tree rings have issues distinct from lower resolution ocean sediment or pollen records. (e.g. Kohfeld and Harrison, 2000; Ramstein et al., 2007; Jones et al., 2009; Harrison and Bartlein, 2012). There are however a number of general issues that affect the use of such data for model evaluation, including the potential for multiple climate controls on a given record, the scale over which they are representative, the need to quantify (and take into account) reconstruction uncertainties, and the sparse and uneven site coverage.

Records used for palaeoclimate reconstructions are in general influenced by several different aspects of climate as well as, potentially, non-climatic factors. For instance, oxygen or hydrogen isotopes from ice cores, carbonates or organic matter are physically meaningful variables, but do not necessarily have a one-to-one stationary relationship with temperature or precipitation (e.g. Werner et al., 2000; Schmidt et al., 2007; Masson-Delmotte et al., 2011). Vegetation, in addition to being influenced by several aspects of seasonal climate, is directly influenced by the atmospheric CO_2 concentration (Prentice and Harrison, 2009). There are several approaches that have been adopted to overcome this type of problem: the use of multi-proxy reconstruction techniques, forward modeling of the system within a climate model or using climate-model output (see an example related to coral carbonate isotopes in Sect. 5.1), and model inversion or data assimilation. Multi-proxy reconstructions rely on the idea that different types of record will be sensitive to different aspects of climate, and that pooling the information from each of these records therefore provides a more robust reconstruction of any specific climate variable. In the sense that forward modeling (and by extension model inversion techniques) are based on physical and or physiological knowledge of the given system, the use of these approaches may be a more robust

CPD
9, 775–835, 2013

Using paleo-climate comparisons to constrain future projections in CMIP5

G. A. Schmidt et al.

way of dealing with the non-stationarity issue – however, as with climate models, the results are constrained by the quality of the models and the degree to which the system is well-understood (see for example the discussion of CO_2 fertilisation in Denman et al., 2007).

The scale over which a record is representative can be a major issue in comparing paleodata and model output. All types of records are responding to local conditions, and for basic meteorological variables it is rare for a record to be representative for spatial scales of more than 50–100 km (though many records, such as tropical ice core $\delta^{18}O$, may have strong correlations to climate further afield; e.g. Schmidt et al., 2007). Comparisons at these scales often require some form of dynamical or statistical downscaling of model output, though there are many associated issues (Wilby and Wigley, 1997). Alternatively, up-scaling reconstructions (for instance, through the use of gridding) can often reveal large-scale patterns that models could be expected to resolve, although this requires a sufficiently dense network of sites. Recent developments include the use of cluster analysis to classify types of model behaviour and to determine cohesive regions for comparison with the large-scale patterns in the observations (e.g. Bonfils et al., 2004; Brewer et al., 2007; Harrison et al., 2013).

Paleoclimate reconstructions are usually accompanied by estimates of measurement or statistical uncertainty. However, in past practice these uncertainties were rarely propagated into large-scale synthetic products (except in terms of non-quantitative quality control measures, see e.g. COHMAP, 1988) and even more rarely taken into account when the reconstructions were used for model evaluation. However, quantitative measures of uncertainty have been included in more recent palaeoclimate syntheses (e.g. MARGO, 2009; Bartlein et al., 2011) and the use of fuzzy-distance measures (Guiot et al., 1999) provides an explicit way to take account data uncertainties in data-model comparisons. It is worth noting that model-data differences cannot be expected to be smaller than the data uncertainties themselves.

CPD
9, 775–835, 2013

2.2 Paleo-modelling issues

There are two particular issues that are more problematic in paleoclimate simulations than, for instance, simulations of the 20th Century: model drift and forcing uncertainty. The issue of coupled climate model drift arises because of the long (∼ thousands of years) time required to bring the deep ocean into equilibrium in coupled ocean-atmosphere models. In some cases, insufficient spin-up time may have been allowed before specific experiments are started. While drift also affects transient historical simulations, the relatively large forcings in the 20th Century mean that residual drift is usually a small component of the transient response. For simulations of the last millennium though, the forcings are much smaller, and drift in the early centuries of the simulation will be a larger fraction of the modelled change (Osborn et al., 2006; Fernández-Donado et al., 2012). One proposal to deal with this is via a correction using the drift in the control simulation (i.e. calculating a smooth trend and removing it from the perturbed simulation prior to analysis). While this works well for temperature, it is not very good for variables that exhibit threshold behaviour such as sea ice extent or precipitation. In practice, this issue needs to be assessed for each proposed comparison.

Secondly, there are important uncertainties in the forcings used for the paleoclimate experiments. This is also true for aerosols in the 20th Century simulations, for instance, but such issues are more prevalent in paleo-simulations. For example, both the magnitude of solar or volcanic forcing over the last millennium, and the size and height of ice sheets at the LGM are sources of major uncertainty. In the last millennium experiments, multiple forcing choices were proposed (Schmidt et al., 2011, 2012), but few groups have attempted (as yet) to comprehensively explore all the options, and this is also true for uncertainties associated with other time periods. If an insufficient range of different forcings is tested, it is plausible that mismatches between observations and simulations may be wrongly attributed to the model (or observations), when in fact they was related to a mis-specified forcing (e.g. Kageyama et al., 2001).

CPD
9, 775–835, 2013

It should also be noted that multi-model ensembles are not a controlled sample from a well-defined distribution of plausible simulations. Models are necessarily incomplete and there are common biases that have more to do with the state of computational technology than physics (for instance, poor or non-existent resolution of ocean eddies). Multi-model ensemble means can be informative and will generally outperform individual models (Annan and Hargreaves, 2011), but care must be taken to assess the suitability of each included model and weighting of individual models needs to be well justified (Knutti et al., 2010).

2.3 Approaches to comparing reconstructions and simulations

There has been a gradual evolution in the approaches for comparing reconstructed changes and simulations from essentially qualitative graphical comparisons of model output and reconstructions of the corresponding climatic variables (e.g. Braconnot et al., 2007) to more quantitative approaches that measure model- data mismatch via some "metric" or distance function (e.g. Sundberg et al., 2012; Izumi et al., 2013). Metrics based on correlations or RMS differences between fields of modern data and model output have been commonly used in model evaluation (e.g. Taylor, 2001; Schmidt et al., 2006; Gleckler et al., 2008). These methods provide opportunities for both inter- and intra- generational model comparisons (Reichler and Kim, 2008; Harrison et al., 2013).

Focusing on the collective performance of the ensemble as a whole, Hargreaves et al. (2011) tested the ability of the PMIP2 ensemble to represent the Last Glacial Maximum in terms of its "reliability"; defined as the adequacy of the ensemble, considered in probabilistic terms, in predicting the changes documented in the paleo-climate archives during that interval. The concept of "skill" as adopted in the numerical weather prediction community is also useful as a quantitative test of model performance: that is, does a model produce a more accurate prediction (match to the paleo-climate record), than that which would be achieved by a simple null hypothesis? (Hargreaves et al., 2012b). While many studies have focused on time-slice or time-series comparisons,

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |
CPD
9, 775–835, 2013
Using paleo-climate comparisons to constrain future projections in CMIP5
G. A. Schmidt et al.

nothing precludes comparing the simulations and paleo-record in the frequency domain. Recent work, has looked at the fluctuations in forcings and data as a function of timescale, and in principle, these fingerprints could also be useful (Lovejoy and Schertzer, 2012).

2.4 Linking past and future

There are two main ways in data-model comparisons can be used as a guide to the future – either as a validation of a robust relationship across models and scenarios, or as a method to discriminate between different models. A prerequisite for the latter example is that the metric chosen actually correlates to future outcomes within the ensemble. If this is not the case, then the metric is orthogonal to the spread in the projections and cannot be used to constrain it. Even when such a relationship is found, we need to consider whether it is physically meaningful to be confident that it has not arisen either though chance due to a small sample size or as an artifact of the model or the experimental design. While connections may in principle be highly complex, it is natural as a first step to consider whether a correlation exists between past and future behaviour in the same diagnostic. The search for useful metrics (in this sense) using modern data has generally been disappointing (Knutti et al., 2010), although there have been a small number of cases where apparently meaningful relationships have been found (Boe et al., 2006; Hall and Qu, 2009; Fasullo and Trenberth, 2012). It is notable that the first two examples relate future climate changes to externally-forced changes in the modern climate (relating to decadal trend, and seasonal range, respectively), rather than using metrics based on the climatological mean state alone. This lends support to our working hypothesis that past variations seen in paleoclimate simulations might also be informative about the future as well as increasing understanding about the past.

Where a credible relationship between past and future is found, there is a range of methods that can be applied to use observations to constrain future predictions (Collins et al., 2012). One method, applied by both Boe et al. (2006) and Hall and Qu (2009),

CPD
9, 775–835, 2013

is to take the observational estimate, and use the relationship (often linear) embodied in the correlation of model output to project this value into the future. An attractive feature of this approach, beyond its simplicity, is that it readily allows extrapolation of the observed relationship in the case where the true value is suspected of lying outside the model range. An alternative approach, which has been widely applied to perturbed physics ensembles is more explicitly Bayesian, considers the ensemble as a probabilistic sample. For the prior, equal weight is typically assigned to each ensemble member. Probabilistic weights are then calculated for each member of the ensemble, according to their performance in reproducing the observations. This weighted ensemble now represents the posterior estimate of future change. This method uses the model spread as a prior constraint, which depending on one's viewpoint, and the specific case in question, may be considered either a strength or weakness of this approach. These and other methods are discussed in more detail in Collins et al. (2012).

3 Robust metrics

In this section we highlight physically-based correlations between key metrics that show similar patterns in the paleo-climate runs and in future projections (or more idealised scenarios). With evaluation via the paleo-climate record, these metrics can be considered robust, and thus provide contingent predications of one variable given a potential change in the other.

3.1 Patterns of regional climate change vs. global means

The main climate forcings for the LGM are the lower concentrations in atmospheric greenhouse gases and the presence of Laurentide and Fenno-Scandinavian ice-sheets in the northern extratropics. The ice sheets have a strong local albedo effect (e.g. Braconnot et al., 2012) but also affect the mid-latitude large-scale atmospheric circulation due to the associated change in topography (e.g. Pausata et al., 2011; Rivière et al., 2009; Laîné et al., 2009). However, away from this perturbation for the atmospheric radiative budget and for the atmospheric dynamics, we expect that the greenhouse gas forcing would be the main forcing for the LGM climate change. There could then be a relationship between LGM climate change and future climate change for a given model, which could be useful in testing the ability of climate models in reproducing regional climate change relative to the global change.

Figure 1 shows the results comparing the mean annual surface air temperature change over a region compared to the global mean change for the abrupt 4xCO2, 1pctCO2 and lgm CMIP5 simulations across a suite of models. We have considered the tropics (land + oceans) and the tropical oceans, which have been used previously in perturbed physics ensemble studies (Schneider von Deimling et al., 2006; Hargreaves et al., 2007), East Antarctica, for which the temperature change is shown to scale with global temperature change for the LGM and the CMIP3 2xCO2 and 4xCO2 changes (Masson-Delmotte et al., 2006a,b) and the well documented mid-latitude region of the North Atlantic and Europe.

For the tropics, for land and ocean points as well and ocean points only, Fig. 1 shows that the relationship between the regional and global temperature change exists for the 1pctCO2 and abrupt 4xCO2 anomalies, and is consistent across these two experiments. Such a relationship also exists for the LGM, but the slope is clearly lower than for the increased CO_2 experiments. Furthermore, the models which simulate the smallest warming for increased CO_2 are not those which simulate the smallest cooling for LGM (and similarly for the models with the largest warmings and coolings). The regional vs. global temperature change relationship appears more consistent between LGM and increased GHG forcings for East-Antarctica and, surprisingly, over the North Atlantic/Europe region. However, for the tropics, rankings of the models according to their cooling for LGM and warming for 1pctCO2 and abrupt 4xCO2 are not consistent. This shows that either the impact from the lower GHG concentrations are not symmetric compared to those for increased GHG concentrations, or that the ice-sheet remote impact extends to the tropics (Laîné et al., 2009).

CPD
9, 775–835, 2013

3.2 Land-ocean contrasts

Model results have consistently shown that for the LGM, the continents cooled more than the ocean (e.g. Braconnot et al., 2007, 2012; Laîné et al., 2009), while, in a symmetric manner, predictions for future climate show a stronger warming over land than over the oceans (e.g. Sutton et al., 2007; Drost et al., 2011). The ratio between cooling over land and cooling over the ocean for the LGM tropics was ~ 1.3 in the PMIP1 computed sea surface temperature (SST) simulations (Pinot et al., 1999), a result close to the ratio of ~ 1.5 found for the PMIP2 fully coupled LGM experiments (Braconnot et al., 2012) and conspicuously close to the 1.5 ratio found by Sutton et al. (2007) for future climate.

This relationship also holds in the most recent CMIP5 simulations (Fig. 2) not only for the tropics but also for the well-documented region of the North Atlantic and Europe, consistent with the LGM data. It is worthwhile to note that this pattern was previously used to highlight the inconsistency in an earlier compilation of tropical LGM sea surface temperatures (Rind and Peteet, 1985). We conclude that these relationships are indeed robust, although they appear imperfectly understood (Lambert et al., 2011).

3.3 Regional extremes

Extreme climate events such as heatwaves and cold spells can have long lasting impacts on society or ecosystems (IPCC SREX, 2012). The development of such events spans days to a few weeks, so that they are largely intra-seasonal by nature (Seneviratne et al., 2012). In such a context, the generally linear relationship between reconstructions and actual climate can be strongly distorted. Hence, since extreme events are by definition rare, large numbers of examples are required to get good statistics. Simulations of the past millennium offer a promising tool to investigate modeled extremes since they sample a large range of possible cases. The strongest limitation for an application of this method to paleoclimatic data has been the necessity of dealing with daily data in order to capture behavior that is non-Gaussian and the need for proxy

CPD
9, 775–835, 2013

data that record extreme variables (Jomelli et al., 2007). However, if we can demonstrate the robustness of the relationships between short and longer term statistics over long periods of time, and/or their dependence on external forcings, we can potentially predict the behavior of temperature extremes in the future.

The statistical analyses of (daily) temperature hot extremes of the 20th century have shown that temperature is generally a bounded variable, for which the upper bound can be computed from the statistical parameters of extremes (Parey et al., 2010a,b). Diagnostic studies focusing on the probability distribution of temperature and precipitation extremes are often based on the application of Extreme Value Theory (EVT), though simpler metrics have also been used (e.g. Hansen et al., 2012). EVT describes the behavior of the probability distribution near the tails, and allows one to compute return levels for return periods that are longer than the period of observation (Coles, 2001). It has been applied to meteorological observations (Parey et al., 2010a), reanalysis data (Nogaj et al., 2006) and model simulations (Kharin et al., 2005, 2007) in order to quantify trends of extremes.

It has also been shown that the extremes of hot and cold temperatures are correlated with mean temperatures over the northern extra-tropics (Yiou et al., 2009). Until now, few models had provided daily output of temperature or precipitation on multi-century timescales (Jansen et al., 2007). However, with increasing storage capacity, daily resolution data is becoming more common and was requested for simulations in the CMIP5 archive (Yiou et al., 2012).

In the extra-tropics, seasonal summer heatwaves are generally preceded by droughts in the Winter-Spring seasons (Fischer et al., 2007; Vautard et al., 2007) with a mechanism that involves a positive feedback between sensible heat fluxes, evapotranspiration and temperature (Schår et al., 1999) and this has also been found in global and regional models for future projections (Seneviratne et al., 2006, 2010; Quesada et al., 2012). A useful statistical metric to connect winter-spring precipitation and summer temperature is quantile regression. Ordinary least-squares regression focuses on the mean values of variables to be connected but by setting a threshold based on

CPD
9, 775–835, 2013

high (or low) quantiles of the variable to be predicted, one can build regression coefficients conditional to high (or low) values of this variable (Koenker, 2005). We illustrate this diagnostic in Fig. 3, by computing the quantile regression for 90th and 10th quantiles of the summer hot day frequency and winter-spring precipitation frequency in the IPSL-CM5A-MR historical simulation and the E-OBS gridded dataset (Haylock et al., 2008). The quantile regression slopes illustrate the asymmetry of the precipitation or temperature dependence for hot or cool summers in Western Europe (Hirschi et al., 2011; Mueller and Seneviratne, 2012; Quesada et al., 2012; Seneviratne and Koster, 2012).

The general picture is that a dry winter/spring tends to favor a hot summer. But while wet winter-spring conditions are generally followed by cool summers (small spread between low and high quantiles), dry winter-spring conditions can be followed by cool summers as well as heatwaves (large spread between low and high quantiles), because the genesis of heatwaves can be broken in just a few days, due to fast variations of the synoptic atmospheric circulation (Hirschi et al., 2011; Quesada et al., 2012). This feature has been tested on CMIP3 and some CMIP5 simulations for the present and A2/rcp85 scenarios. It was shown that the seasonal predictability of large European heatwaves decreases under warmer conditions, although their frequency increases (Quesada et al., 2012).

There have been many studies compiled by historians focusing on European heatwaves in recent centuries and their impacts on society (Le Roy Ladurie, 2004, 2006; Barriendos and Rodrigo, 2006; Camuffo et al., 2010). Hence, using a metric to capture heatwave dynamics is a promising approach to investigate major heatwaves that struck Europe during the last millennium, and to explore the relationship between summer temperature and winter-spring precipitation preconditioning, with different climate forcings, especially land use, though this remains a work in progress.

CPD
9, 775–835, 2013

4 Discriminating metrics

In this section we highlight metrics for which we have paleoclimate information that serve to discriminate between models that show different behaviours in future projections (or more idealised scenarios).

4.1 Rainfall change in South America

Projections of precipitation change in South America have a large spread in the CMIP3 archive (Meehl et al., 2007). In future projections, most models simulate a dipole of precipitation change in Northern South America, but the sign of this dipole depends on the model. If this feature is an intrinsic response in each model to a forcing, it might be possible to evaluate the dipole response in the paleo-climate simulations.

We define the precipitation dipole as the annual-mean precipitation averaged over 0°–8° N–60° W–50° W minus the annual-mean precipitation averaged over 5° S–15° S–45° W–35° W. Over the 16 models examined here, the models that have the 5 lowest values for rcp85-piControl precipitation dipole change are classified as group 1; those with the 5 highest values are classified as group 3, and the remainder, in group 2. The group 1 models simulate drier Guyana, Venezuela and Colombia, and wetter Nordeste and Eastern Brazil, associated with a Southward shift of the Inter-Tropical Convergence Zone (ITCZ). The group 3 models simulate a wetter Venezuela and drier Eastern Brazil, associated with a Northward shift of the ITCZ.

Figure 4 shows a strong link between precipitation changes in the future and precipitation changes in the MH. Models in group 1 show a dipole in the MH which is similar to the dipole they simulate in the future, with a strong Southward shift of the ITCZ. In contrast, models of group 3 show instead a broadening of the ITCZ in the MH. Therefore, paleo-proxies of precipitation along the South American coast could help determine which group of models is the most realistic in the MH, and, by extension, which simulation of future change has greater credibility (Silva Dias et al., 2009) The Prado et al. (2013) paleo-data synthesis for South America suggests drying everywhere except over Northeastern Brazil, a pattern that is most consistent with that simulated by the group 1 models.

To gain confidence in such a paleo-constraint, we need to understand the physical processes that explain the common behavior between past and future. This preliminary analysis will not fully answer this question, but it does illustrate how to make use of the wealth of past, future and idealized CMIP5 simulations. Table 1 shows a selection of correlations between precipitation changes and other model features. First, in the future climate, shifts in the ITCZ seem to be associated with shifts in the SST dipole in the Atlantic: models that shift the ITCZ the most southwards are those with the strongest warming south of the Equator relative to the rest of the Atlantic. ITCZ shifts in response to SST dipoles are expected (e.g. Kang et al., 2008). However, this relationship does not seem to hold for the MH to PI change. Second, the atmospheric component of the model also appears to play a key role. Some of the different model behaviors can be seen in amipFuture simulations, where all models are forced by the same pattern of SST warming. In addition, much of the different model behaviors can already be seen in sstClim4xCO2 simulations, where a quadrupling of CO_2 is imposed with SST held constant. This is consistent with the fast response to CO_2 being an important component of the total precipitation response in global warming (e.g. Bala et al., 2009). Models that decrease precipitation over Northern South America in the projections and in the MH are those that decrease precipitation over this region under $4 \times CO_2$. They also happen to be the models with the strongest land surface warming in response to both $4 \times CO_2$ and to MH forcing. Therefore, the different groups of models show different precipitation response to SST changes, orbital forcing and to $4 \times CO_2$, but the response shows similarity between all these different forcings and within each model group. This suggests that common mechanisms are involved in the precipitation response to all forcings, and that this is representative of each individual model. Finally, it is worth noting that models in group 3 often show the most significant "double ITCZ" problem in the Atlantic, an obvious, and persistent, common model bias.

CPD
9, 775–835, 2013

4.2 LGM constraints on climate sensitivity

The LGM has been a prime target for assessments of climate sensitivity since it is a quasi-stable period with significant climate differences from today, with reasonably well-known boundary conditions and sufficient data to reconstruct large-scale climate shifts (e.g. Lorius et al., 1990; Edwards et al., 2007; Köhler et al., 2010; Schmittner et al., 2011; PALAEOSENS, 2012).

We can apply the methods described in Sect. 2 to estimate of the equilibrium climate sensitivity based on the CMIP5 LGM simulations. We use an ensemble of opportunity consisting of 7 models which participated in the PMIP2 experiment, together with 4 CMIP5 models for which sufficient data are available (at time of writing). Estimates of the climate sensitivities of these models were obtained from a variety of sources and were derived using a range of methods. For the PMIP2/CMIP3 models, sensitivity was generally calculated using a slab ocean coupled to the atmospheric component (Meehl et al., 2007), whereas in CMIP5, the most readily available estimates use a regression based on a transient simulation (Andrews et al., 2012). These estimates are not perfectly commensurate, with some models reporting a 10 % difference in the two methods (Schmidt et al., 2013). Some of the PMIP2 models used for the LGM simulations may also differ from the equivalent CMIP3 versions for which the sensitivity estimates were made. Thus, the values used here may be somewhat inconsistent and imprecise, although we expect the uncertainty arising from these sources to be modest in comparison to the range of values represented across the ensemble. The boundary conditions for the LGM simulations are essentially unchanged between PMIP2 and CMIP5 (save for changes in the shape of the imposed ice sheets), allowing us to consider these experiments as broadly equivalent (though there are some systematic biases, Kageyama et al., 2012). Limitations in the boundary conditions (such as the exclusion of dust and vegetation effects) may, however, introduce additional bias and uncertainty into our result, which we do not attempt to account for here. For these and other reasons discussed below, these results should be considered as a proof of concept rather than conclusive.

The LGM was associated with a large negative radiative forcing anomaly with respect to the pre-industrial including substantially lower concentrations of greenhouse gases (e.g. Köhler et al., 2010). However, the ensemble does not show an expected negative correlation between climate sensitivities and their globally averaged LGM temperature anomalies (over the full 100 yr of simulation output) (Fig. 5a, see also Crucifix, 2006). There is a strong negative correlation in the tropics, most strongly in the latitude band 10° S–30° N (Fig. 5b) (Hargreaves et al., 2012a). The correlation is weaker at higher latitudes where the feedbacks in response to large cryospheric changes may be very different to those exhibited in a future warmer climate. There is also a strong positive correlation in the southern ocean (i.e., colder LGM anomalies are linked with *lower* sensitivity), possibly due to a large range of biases in the control climate (Fig. 5c). The correlation of piControl temperatures to sensitivity points to the Arctic and the southern oceans as regions where base climatology impacts sensitivity, probably via cloud effects (see Trenberth and Fasullo, 2010, for a discussion). The strong negative correlation ($r = -0.8$) between the LGM temperature anomalies in the latitude band 10° S–30° N, and the climate sensitivities of the models (Fig. 6), is physically plausible, since this region is far from the cryospheric and sea ice changes of the LGM, and the forcing here is dominated by the reduction in greenhouse gas concentrations.

If we assume that the correlation with tropical temperatures provides a valid constraint on the real climate system, we can use this correlation to project an observational estimate of the past change onto the future, as in Boe et al. (2006). Recently, Annan and Hargreaves (2012) generated a new estimate of LGM temperature changes, based on a combination of several multiproxy data sets, and the ensemble of PMIP2 models. The method does not depend on the magnitude of changes estimated by the models, but only their spatial patterns. Using the resulting estimate of LGM temperature change in this latitude band of -2.2 ± 0.7 °C (at 90 % confidence), the predicted value for climate sensitivity arising from the correlation is 2.7 °C, with a 90 % interval of

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

CPD
9, 775–835, 2013

Using paleo-climate comparisons to constrain future projections in CMIP5

G. A. Schmidt et al.

1.2–4.1 °C calculated by Monte Carlo sampling, but this range is somewhat sensitive to the reconstruction uncertainties.

For a more explicitly Bayesian approach, we initially assign equal probability to each model in the ensemble. This choice can be questioned, given both the range of model complexities, and also the possible inter or intra-generational similarities between models of related origins (Masson and Knutti, 2010). However, quantifying these issues is far from straightforward, so we make our choice for reasons of practicality and in order to demonstrate the utility of the overall method. A standard kernel density estimation based on the ensemble leads to the prior distribution presented as the green curve in Fig. 7, which has a 90 % range of 1.7–4.9 °C and a mean of 3.4 °C. The observationally-derived estimate of tropical temperature gives rise to the natural likelihood function $L(M|O) = P(O|M)$ from which the weights are calculated (where O represents the observations, M the model simulation, $L(M|O)$ the likelihood of the model result given the observed data, and $P(O|M)$ the probability of the observations assuming that the models are correct). The posterior distribution is shown in red, the bulk of which has been shifted to lower values with the mean reducing to 2.8 °C. Its 90 % probability range has only moved slightly, however, to 1.6–4.7 °C. The reason for the upper limit here remaining high is that the highest sensitivity model in the ensemble has been assigned a fairly large weight since it matches the reconstructions well. The small size of the ensemble means that this approach is rather sensitive to the presence or absence of particular models in the ensemble.

The two approaches differ considerably in their use of the model ensemble. In the latter case, the ensemble is directly used as a prior estimate, which therefore imposes quite a strong constraint on climate sensitivity even before these observational constraints are used. The former method may be considered as roughly equivalent to using a prior that is uniform in the observed variable (here tropical temperature), although this approach is rarely presented in explicitly Bayesian terms. Despite the different assumptions and approaches, these methods both generate rather similar estimates for the climate sensitivity – both assigning highest probability towards the lower end of the

CPD
9, 775–835, 2013

model range. The ranges are comparable with other paleo-climate derived estimates of 2.3–4.8 °C (68 % confidence interval, PALAEOSENS, 2012) but, given the small ensemble size and possible naïvety of the assumptions made here, these estimates may not be robust and need to be tested using a larger ensemble.

4.3 Arctic Sea ice sensitivity constraints from the mid-Holocene

The rate and pattern of Arctic sea ice change in the future is of key scientific interest due both to the surprisingly rapid changes currently occurring and the large spread in model estimates in, for instance, the onset of summertime "ice-free" conditions (Stroeve et al., 2012; Massonet et al., 2012).

Recent studies (Mahlstein and Knutti, 2012; Abe et al., 2011) have demonstrated that biases in sea ice volume have a strong impact on the simulated responses to radiative perturbations, and that there maybe a possibility to discriminate among models based on interannual modes of variability. The mid-Holocene simulations (driven mainly by changes in orbital forcing) may provide a orthogonal test of Arctic sea ice sensitivity. MH insolation changes imply that NH summers were warmer than summers today (see Kutzbach, 1981, and many subsequent papers). Paleo-data from the circum-Arctic region indicates that this warmth was accompanied by reductions in sea ice extent at least during some months of the year (Dyke and Savelle, 2001; de Vernal et al., 2005; McKay et al., 2008; Funder et al., 2011; Polyak et al., 2010; Moros et al., 2006).

The CMIP5 MH simulations (Fig. 8) consistently show decreases in sea ice extent from August through to November. Changes in winter months are not coherent across the models, though these changes are not well characterised in the paleo-data either. There is a relationship (Fig. 9) between the size of the anomaly at the MH and in future projections, presumably reflecting the underlying sensitivity of the sea ice model and Arctic climate in general (see also O'ishi and Abe-Ouchi, 2011). This correlation exists despite the variations in the cause of the ice loss (summer insolation versus greenhouse-gas-related forcing). Although, the small size of the ensemble raises questions of robustness of the relationships, it should be possible to use the MH ice extent

CPD
9, 775–835, 2013

anomaly to estimate the likely loss in future projections. However, it may also be possible to use more specific or local diagnostics to compare to a wider proxy network for a similar constraint (Tremblay et al., 2013).

5 Exploratory metrics and limitations

In this section we provide examples of where the paleo-climate information is ambiguous, or where connections seen in paleo-climate changes do not translate into the future for some reason. This may be related to forcing ambiguities, climate-change related divergence, or potentially, a misunderstanding of the dominant processes. While these examples are not directly informative about the future, they illustrate how the limitations of our outlined approaches can be explored in ways that illuminate key uncertainties.

5.1 20th-century changes in tropical Pacific climate

The response of the tropical Pacific Ocean to anthropogenic climate change is uncertain, partly because we do not fully understand how the region has responded to anthropogenic influences during the 20th century. Instrumentally based estimates of SST do not depict an internally consistent view (Deser et al., 2010), and model simulations similarly disagree regarding the 20th-century trend (Thompson et al., 2011). Understanding trends in the tropical Pacific is particularly challenging because the instrumental record is sparse even for the early 20th Century and long-term in situ measurements of SST are uncommon. High-resolution paleoclimate records, particularly the large network of tropical Pacific coral $\delta^{18}O_{calcite}$ records, can be used in conjunction with the observational record and help interpret tropical climate trends. These proxy records respond to the combined effects of SST and the isotopic composition of seawater ($\delta^{18}O_{sw}$) (which is strongly correlated to sea surface salinity, SSS) and can reveal changes on longer time scales. To address the limitations of each individual

CPD
9, 775–835, 2013

archive, Thompson et al. (2011) proposed using a forward-modeling approach to generate synthetic coral records (i.e., pseudocorals) from observational and climate model output and test whether these pseudocorals are in agreement with the network of coral $\delta^{18}O_c$ observations. If they agree with the $\delta^{18}O_c$ records, then the causes of change in the region may be inferred, while disagreement may reveal key uncertainties in the data.

The coral $\delta^{18}O$ model from Thompson et al. (2011) calculates isotopic variations as a function of SST and SSS, with an SST-$\delta^{18}O_c$ slope of -0.22 ‰ °C^{-1} and the SSS-$\delta^{18}O_{sw}$ slope varying by region (LeGrande and Schmidt, 2006). When driven with historical SST and SSS data, this simple model of $\delta^{18}O_c$ was able to capture the spatial and temporal pattern of ENSO and the linear trend observed in 23 Indo-Pacific coral records between 1958 and 1990 (Thompson et al., 2011). The observed trends were driven primarily by warming at the coral sites, though SSS trends were responsible for approximately 40 % of the shared $\delta^{18}O_c$ trend. These results not only indicate a significant SSS trend in the tropical Pacific, but also the importance of $\delta^{18}O_{sw}$ in simulating the observed $\delta^{18}O_c$.

However, pseudocoral records calculated from CMIP3 historical simulations did not reproduce the magnitude of the secular trend, the change in mean state, or the change in ENSO-related variance observed in the coral network from 1890 to 1990. Similarly large discrepancies are present between CMIP5 simulations and the observations, with none of the individual CMIP5 pseudocoral networks producing trends as strong as in the observed 20th Century coral records. While the observational coral network suggests a reduction in ENSO-related variance and an El Niño-like trend over the 20th century, CMIP3 and CMIP5 simulations vary greatly on both points.

The differences between observed and GCM-derived $\delta^{18}O_c$ trends may stem from the simplicity of the forward model for $\delta^{18}O_c$, bias in the coral records, and/or errors in the GCM SST and SSS fields. In particular, the potential role of non-climatic trends in $\delta^{18}O_c$ and the magnitude and spatiotemporal pattern of the $\delta^{18}O_{sw}$-SSS relationship needs to be further investigated. Preliminary tests with data from isotope-enabled

CPD
9, 775–835, 2013

coupled control simulations (LeGrande and Schmidt, 2011) support a minor role for short term isotope variability, though it remains to be seen if longer-term trends in the 20th Century simulations are important. Previous work has highlighted potential biases in simulated salinity fields as a potential source of the observed-simulated discrepancy (Thompson et al., 2011, 2012). For example, CMIP3 and CMIP5 simulations display weak and spatially heterogeneous SSS trends, such that the magnitude of the $\delta^{18}O_c$ trend in CMIP3 and CMIP5 simulated pseudocorals is indistinguishable from the trends observed in individual centuries of an unforced control run (Fig. 10, upper panel). We also find that the trends in mean state and change in ENSO-related variance within the basin are highly variable among the CMIP5 models, and even between ensemble members of the same model. On the other hand, while pseudocorals, modeled from the new SODA 20th-century reanalysis of SST and SSS, display greater agreement with the observed coral trends, two recent versions of this product disagree regarding the relative contribution of SST and SSS. These results suggest that more work is needed to constrain the magnitude of the observed 20th-century salinity trend throughout the tropical Pacific Ocean.

Despite the disagreement among models and runs regarding the change over the 20th Century, the CMIP5 projections converge upon a more El Niño-like (e.g. warmer eastern equatorial Pacific) mean state change by 2100 under RCP 4.5 (with only one model suggesting the opposite), consistent with the CMIP3 projections (Meehl et al., 2007). However, the models still disagree about the change in ENSO-related variance. Further, there is no clear relationship between the magnitude of the simulated 20th-century $\delta^{18}O_c$ trend and the projected future $\delta^{18}O_c$ trend in the CMIP5 ensemble (Fig. 10, lower panel). This suggests that an agreement of the simulated 20th-century change in the tropical Pacific with that of the observational coral network would not be a reliable indicator of future trends.

CPD
9, 775–835, 2013

5.2 Spectra and fluctuation analyses

As mentioned above, there is no restriction on what kind of variables, means, variances or higher-order statistics can be used in these analyses. In this section we highlight two analyses in the frequency domain that demonstrate the important role of relatively uncertain forcings in assessing skill.

In Fig. 11, we show the maximum-entropy method (MEM) spectra (using 30 poles) for the NH mean land surface temperature over 8 last-millennium simulations with the GISS-E2-R model that were run with different combinations of plausible solar, volcanic and land use forcings (Schmidt et al., 2011, 2012). The spectra are similar for models that have the same volcanic forcing, and significantly different when the volcanic forcing is derived from a different dataset or where there is no volcanic forcing at all. Specifically, interannual to multi-decadal variability is much larger when volcanoes are imposed, and the larger the volcanic forcing, the greater the variability, with the largest response in simulations using the Gao et al. (2008) reconstruction, compared to the Crowley et al. (2008) reconstruction. In contrast, the difference between two different solar forcings (Vieira et al., 2011; Steinhilber et al., 2009) is not detectable in this metric. (Note that the implementation of the Gao et al., 2008, volcanic forcing in these simulations was mis-specified and gave roughly twice the expected radiative forcing. Although part of the increase in variance seen here was unanticipated, given the uncertainties in specifying the forcing, the exercise is useful in highlighting the role of the forcings in determining variance.)

Another analysis in the spectral domain is one focused on power law scaling (Lovejoy and Schertzer, 1986). Several scaling studies of GCMs demonstrate that they generally simulate the statistics (including spectral scaling exponents) reasonably well up to ≈ 10 yr scales (e.g. Fraedrich and Blender, 2003; Zhu et al., 2006; Rybski et al., 2008; Lovejoy and Schertzer, 2012; Vyushin et al., 2012). This already gives us confidence in the decadal scale responses of GCM. However, tests at lower frequencies will depend on solar and volcanic forcings as well as the possible impacts of slow processes such

CPD
9, 775–835, 2013

as deep ocean or land-ice dynamics which are currently missing or perhaps poorly represented in the models.

Following Lovejoy et al. (2012), we calculate the Root Mean Square (RMS) fluctuation as a function of time-scale, from months to centuries, for the NH land temperatures using the same eight runs of the GISS-E2-R model used above for the period 1500–1900 CE. Since simulations are strongly clustered according to changes in the volcanic forcing used (Fig. 11), for simplicity we averaged over the three GRE and three CEA volcanic and the two no-volcanic runs.

For comparison, we show the mean of the same metric from three multiproxy reconstructions (Huang, 2004; Moberg et al., 2005; Ljundqvist, 2010). The multiproxy average is processed with and without the 20th Century to indicate the importance of that period for the scaling behaviour – in all cases the variance in the multi-decadal to century scale is greatly enhanced by the recent anthropogenic trend. These curves show fluctuations decreasing with scale over the low frequency weather regime (months to decades) but increasing in the climate regime (decades to centuries).

The comparison with the GISS-E2-R simulations is illuminating. First, we note that at the decadal scale, the sign of the all the slopes changes. However, the simulations vary in the opposite direction from the data: first growing and then decreasing with scale. Only the volcano-free runs (bottom) qualitatively follow the reconstructions by first decreasing and then increasing with scale. When compared to the surface data and multiproxy reconstructions we see that at ~ 10 yr, the simulations have variance that is too large while at longer scales (> 100 yr) the variance is too small.

These results demonstrate clear mismatches in behaviour between the models' simulated variance at different scales and the inferred variability from multi-proxy reconstructions. However, there are strong sensitivities to the (uncertain) external forcing functions, precluding a straightforward attribution of the mismatch to potentially misspecified forcings, missing mechanisms, insufficient "slow" variability or data problems in the reconstructions.

CPD
9, 775–835, 2013

5.3 Hydroclimate divergence

Distinct from temperature, hydroclimate variability can be quantified using a range of variables, including precipitation, soil moisture, lake levels, or other synthetic indices (e.g. Nigam and Ruiz-Barradas, 2006). Most models provide output for these diagnostics, but often these variables are not available directly from paleo-climate archives, creating a challenge when conducting model-data comparisons. However, calibrations of networks of precipitation sensitive tree ring widths have been used to reconstruct the Palmer Drought Severity Index (PDSI) in North America and Asia over the Common Era (Cook et al., 2004, 2010). PDSI is calculated using temperaturederived estimates of the evapo-transpiration and precipitation, and nominally represents a normalized index of soil moisture, with negative values indicating drought and positive values indicate wetter than normal conditions. There are many outstanding issues with using variations of the index globally to assess drought, in definition and availability and quality of inputs and sensitivity (e.g. contrast Sheffield et al., 2012; Dai, 2012). However, we focus here on the question of how well does this index, if derived from GCM output, reflect actual model soil moisture and whether this relationship changes over time.

From two GCMs (GISS-E2-R and MIROC-ESM), we calculated PDSI using model temperature and precipitation (the Thornthwaite method) and compared this index against the standardized (zero mean, unit standard deviation, over the 1850–1950 period, 10 yr smoothing) total column soil moisture model output for the Central Plains of North America (105° W–90° W; 32° W–48° W) (Fig. 13). Prior to the start of the industrial period in 1850, PDSI and soil moisture track each other closely in both models (GISS-E2-R: $r = 0.82$; MIROC-ESM: $r = 0.50$). Beginning near the middle of the twentieth century, however, the two indices begin to diverge dramatically. In one model (GISS-E2-R) the correlation weakens considerably ($r = 0.33$), while in the other model (MIROC-ESM) the sign of the correlation actually reverses ($r = -0.29$).

PDSI changes over the twenty–first century indicate severe and unprecedented drought, in contrast to the model soil moisture trends, which indicate a modest shift

CPD
9, 775–835, 2013

towards drying (GISS-E2-R) or even wetter conditions over the coming decades (MIROC-ESM). The reason for this divergence is in the treatment of evapotranspiration (ET) in the model soil moisture versus in the PDSI (Thornthwaite) calculation. In this PDSI calculation, temperature is used as a proxy for the energy available while in the GCMs the soil energy and moisture budgets are calculated directly using explicit physical models. In reality, Thornthwaite ET becomes increasingly decoupled from temperature as the temperature increases, a factor reflected in the model soil moisture but not in the PDSI index. For time periods with strong transient forcing in temperature (e.g., the late twentieth century and into the future), our analysis suggests that the usefulness of PDSI for diagnosing drought and hydroclimate trends is limited. This suggests caution should be used when trying to convert projected variables to those defined from the paleoclimate record.

6 Conclusions and recommendations

In this paper, we have focused the opportunities provided by inclusion of "out-of-sample" paleo-climate experiments within the CMIP5 framework, and specifically how measures of skill in modelling paleo-climate change might inform future projections of climate change.

We have shown that some relationships are robust across the ensemble of models, simulations and paleo-data (Sect. 3) and furthermore that there are skill measures that are well correlated to the simulated magnitude of future change, thus allowing the likely magnitude of future changes to be constrained (Sect. 4). However, there is a need for caution because of the limitations with models, the experimental setup used in CMIP5, or with the paleo-climate data itself (Sect. 5).

Our examples suggest that there are some general requirements for attempts to use the paleo-climate simulations to quantitatively constrain future projections. Each example makes use of a specific target (or targets) from a paleo-climate reconstruction

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |
CPD
9, 775–835, 2013
Using paleo-climate comparisons to constrain future projections in CMIP5
G. A. Schmidt et al.

of change, defines a metric of skill that quantifies the accuracy of the modeled changes and assesses the connection to a future prediction. We recommend that ideally:

– paleo-data targets be spatially representative synthesis products with well-characterised uncertainties,

– the chosen metrics shouqld be robust to uncertainties in external forcing,

– the chosen metrics should not be overly sensitive to the model representation of key phenomena, and are within the scope of the modelled system,

– any relationship between the targets in the past and the future predictions should be examined, and not simply assumed.

Under these conditions, the likelihood of a significant constraint is much greater.

We underline the need for paleo-simulations to be performed with models that are also being used for future projections and that model diagnostics are commensurate (see also Schmidt, 2012). Although the robustness of some of our analyses is limited by the small number of paleo-simulations currently available in the CMIP5 database, we hope that the demonstration of their potential to address questions relevant to the future should encourage other modeling groups to complete and archive these simulations.

There are also important lessons here for the paleo-data community. Our analyses rely heavily on the use of synthesis data products, for instance the MARGO dataset for the LGM (MARGO, 2009), pollen-based reconstructions for the Mid-Holocene (Bartlein et al., 2011), multi-proxy reconstructions of hemispheric temperature (e.g. Moberg et al., 2005), or gridded tree-ring based reconstructions of PDSI for the last millennium (Cook et al., 2010). Such products are invaluable, but there is a need for increased transparency of included uncertainties and continued expansion e.g. see Müller et al. (2011) for sea ice extent. Increasing model complexity, for instance by including a carbon cycle, fire models or online tracers such as water isotopes, necessitates the creation of new syntheses (e.g. charcoal records: Daniau et al., 2012; or sea surface carbonate isotopes: Oppo et al., 2007).

CPD
9, 775–835, 2013

The periods and experiments chosen in paleo-climate experiments are far more limited than the number of interesting features in the paleo-climate record. The three periods selected for CMIP5 were chosen on the basis of their relative maturity (the existence of prior sets of experiments, already tested issues, existing data syntheses), but additional periods are also potentially useful – the mid-Pliocene (2.5 million yr ago), the 8.2 kyr event, the last interglacial, the peak Eocene etc. (see Schmidt, 2012 for justifications). Some of these periods are already being examined in a coordinated fashion (e.g. Haywood et al., 2012, and Dolan et al., 2012, for the Pliocene), and it is to be hoped that more will be started. Further expansion of the model experiments will increasingly produce higher frequency diagnostics (daily and sub-daily variations), and perturbed physics ensembles, to better characterise the model structural uncertainty. Both of these expansions will create possibilities for more, and better, tests of model performance. In the meantime, there is already a huge scope for more informative comparisons that can be made using the existing databases.

Acknowledgements. This paper arose from a workshop at the Bishop Museum, Honolulu in March 2012. Funding from NASA, NOAA, NSF, AORI, PAGES and CLIVAR is gratefully acknowledged. Organisation of the workshop was by the PAGES-CLIVAR Intersection Panel. We acknowledge the WCRP WGCM, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table 1 of this paper) for producing and making available their model output. For CMIP the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We thank Rumi Ohgaito and Tetsuo Sueyoshi from JAMSTEC for help in some of the analysis.

CPD
9, 775–835, 2013

References

Abe, M., Shiogama, H., Nozawa, T., and Emori, S.: Estimation of future surface temperature changes constrained using the future-present correlated modes in inter-model variability of CMIP3 multimodel simulations, J. Geophys. Res., 116, D18104, doi:10.1029/2010JD015111, 2011.

Annan, J. D. and Hargreaves, J. C.: Understanding the CMIP3 multimodel ensemble, J. Climate, 24, 4529–4538, doi:10.1175/2011JCLI3873.1, 2011.

Annan, J. D. and Hargreaves, J. C.: A new global reconstruction of temperature changes at the Last Glacial Maximum, Clim. Past Discuss., 8, 5029–5051, doi:10.5194/cpd-8-5029-2012, 2012.

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. E., Williams, J., and Wu, H.: Pollen-based continental climate reconstructions at 6 and 21 ka: a global synthesis, Clim. Dynam., 37, 775–802, 2011.

Barriendos, M. and Rodrigo, F.: Study of historical flood events on Spanish rivers using documentary data, Hydrolog. Sci. J., 51, 765–783, 2006.

Boe, J. L., Hall, A., and Qu, X.: September sea-ice cover in the Arctic Ocean projected to vanish by 2100, Nat. Geosci., 2, 1–3, 2009.

Bonfils, C., de Noblet-Ducoudré, N., Guiot, J., and Bartlein, P.: Some mechanisms of mid-Holocene climate change in Europe, inferred from comparing PMIP models to data, Clim. Dynam., 23, 79–98, 2004.

Braconnot, P., Otto-Bliesner, B., Harrison, S., Joussaume, S., Peterchmitt, J.-Y., Abe-Ouchi, A., Crucifix, M., Driesschaert, E., Fichefet, Th., Hewitt, C. D., Kageyama, M., Kitoh, A., Laîné, A., Loutre, M.-F., Marti, O., Merkel, U., Ramstein, G., Valdes, P., Weber, S. L., Yu, Y., and Zhao, Y.: Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum – Part 1: experiments and large-scale features, Clim. Past, 3, 261–277, doi:10.5194/cp-3-261-2007, 2007.

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, Natural Climate Change, 2, 417–424, doi:10.1038/nclimate1456, 2012.

CPD
9, 775–835, 2013

Brewer, S., Guiot, J., and Torre, F.: Mid-Holocene climate change in Europe: a data-model comparison, Clim. Past, 3, 499–512, doi:10.5194/cp-3-499-2007, 2007.
Camuffo, D., Bertolin, C., Diodato, N., Barriendos, M., Dominguez-Castro, F., Cocheo, C., della Valle, A., Garnier, E., and Alcoforado, M. J.: The western Mediterranean climate: how will it respond to global warming?, Clim. Change, 100, 137–142, 2010.
Carton, J. A. and Giese, B. S.: A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA), Mon. Weather Rev., 136, 2999–3017, 2008.
COHMAP Members: Climatic changes of the last 18 000 yr: observations and model simulationsm, Science, 241, 1043–1052, 1988.
Coles, S.: An introduction to statistical modeling of extreme values, Springer, London, New York, 2001.
Collins, M., Chandler, R. E., Cox, P. M., Huthnance, J. M., Rougier, J., and Stephenson, D. B.: Quantifying future climate change, Natural Climate Change, 2, 403–409, 2012.
Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The twentieth century reanalysis project, Q. J. Roy. Meteor. Soc., 137, 1–28, 2011.
Cook, E. R., Woodhouse, C. A., Eakin, C. M., Meko, D. M., and Stahle, D. W.: Long-term aridity changes in the Western United States, Science, 306, 1015–1018, 2004.
Cook, E. R., Seager, R., Heim Jr., R. R., Vose, R. S., Herweijer, C., and Woodhouse, C., Megadroughts in North America: placing IPCC projections of hydroclimatic change in a long-term palaeoclimate context. J. Quaternary Sci., 25, 48–61, 2010.
Crowley, T. J., Zielinski, G., Vinther, B., Udisti, R., Kreutzs, K., Cole-Dai, J., and Castellano, E.: Volcanism and the Little Ice Age, PAGES Newslett., 16, 22–23, 2008.
Crucifix, M.: Does the Last Glacial Maximum constrain climate sensitivity? Geophys. Res. Lett., 33, L18701, doi:10.1029/2006GL027137, 2006.
Daniau, A.-L., Bartlein, P. J., Harrison, S. P., Prentice, I. C., Brewer, S., Friedlingstein, P., Harrison-Prentice, T. I., Inoue, J., Izumi, K., Marlon, J. R., Mooney, S., Power, M. J., Stevenson, J., Tinner, W., Andrič, M., Atanassova, J., Behling, H., Black, M., Blarquez, O., Brown, K. J., Carcaillet, C., Colhoun, E. A., Colombaroli, D., Davis, B. A. S., D'Costa, D., Dodson, J., Dupont, L., Eshetu, Z., Gavin, D. G., Genries, A., Haberle, S., Hallett, D. J.,

Hope, G., Horn, S. P., Kassa, T. G., Katamura, F., Kennedy, L. M., Kershaw, P., Krivonogov, S., Long, C., Magri, D., Marinova, E., McKenzie, G. M., Moreno, P. I., Moss, P., Neumann, F. H., Norström, E., Paitre, C., Rius, D., Roberts, N., Robinson, G. S., Sasaki, N., Scott, L., Takahara, H., Terwilliger, V., Thevenon, F., Turner, R., Valsecchi, V. G., Vannière, B., Walsh, M., Williams, N., and Zhang, Y.: Predictability of biomass burning in response to climate changes, Global Biogeochem. Cy., 26, GB4007, doi:10.1029/2011GB004249, 2012.

de Vernal, A., Hillaire-Marcel, C., Solignac, S., Radi, T., and Rochon, A.: Reconstructing seaice conditions in the Arctic and Subarctic prior to human observations, in: Arctic Sea Ice Decline: Observations, Projections, Mechanisms and Implications, edited by: de Weaver, E., AGU Monograph Series, 180, 2745, 2008.

Denman, K. L., Brasseur, G., Chidthaisong, A., Ciais, P., Cox, P. M., Dickinson, R. E., Hauglustaine, D., Heinze, C., Holland, E., Jacob, D., Lohmann, U., Ramachandran, S., da Silva Dias, P. L., Wofsy, S. C., and Zhang, X.: Couplings between changes in the climate system and biogeochemistry, in: IPCC, Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., and Miller, H., Cambridge University Press, Cambridge, United Kingdom and United States, 499–587, 2007.

Deser, C., Phillips, A. S., and Alexander, M. A.: Twentieth century tropical sea surface temperature trends revisited, Geophys. Res. Lett., 37, L10701, doi:10.1029/2010GL043321, 2010.

Dolan, A. M., Koenig, S. J., Hill, D. J., Haywood, A. M., and DeConto, R. M.: Pliocene Ice Sheet Modelling Intercomparison Project (PLISMIP) – experimental design, Geosci. Model Dev., 5, 963–974, doi:10.5194/gmd-5-963-2012, 2012.

Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F. F., Cozic, A., Cugnet, D., de Noblet, N., Duvel, J.-P., Ethé, C., Fairhead, L., Fichefet, T., Flavoni, S., Friedlingstein, P., Grandpeix, J.-Y., Guez, L., Guilyardi, E., Hauglustaine, D., Hourdin, F., Idelkadi, A., Ghattas, J., Joussaume, S., Kageyama, M., Krinner, G., Labetoulle, S., Lahellec, A., Lefèbvre, M.-P., Lefèvre, F., Lévy, C., Li, Z. X., Lloyd, J., Lott, F., Madec, G., Mancip, M., Marchand, M., Masson, S., Meurdesoif, Y., Mignot, J., Musat, I., Parouty, S., Polcher, J., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Talandier, C., Terray, P., and Viovy, N.: Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5, Clim. Dynam., in review, 2013.

Dyke, A. S. and Savelle, J. M.: Holocene history of the Bering Sea bowhead whale (Balaena mysticetus) in its Beaufort Sea summer grounds off southwestern Victoria Island, Western Canadian Arctic, Quaternary Res., 55, 371–379, 2001.

Edwards, T. L., Crucifix, M., and Harrison, S. P.: Using the past to constrain the future: how the palaeorecord can improve estimates of global warming, Prog. Phys. Geogr., 31, 481–500, doi:10.1177/0309133307083295, 2007.

Fasullo, J. T. and Trenberth, K. E.: A Less Cloudy Future: The Role of Subtropical Subsidence in Climate Sensitivity, Science, 338, 792–794, 2012.

Fernández-Donado, L., González-Rouco, J. F., Raible, C. C., Ammann, C. M., Barriopedro, D., García-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.: Temperature response to external forcing in simulations and reconstructions of the last millennium, Clim. Past Discuss., 8, 4003–4073, doi:10.5194/cpd-8-4003-2012, 2012.

Fischer, E. M., Seneviratne, S. I., Lüthi, D., and Schär, C.: Contribution of land-atmosphere coupling to recent European summer heat waves, Geophys. Res. Lett., 34, L06707, doi:10.1029/2006GL029068, 2007.

Fraedrich, K. and Blender, K.: Scaling of atmosphere and ocean temperature correlations in observations and climate models, Phys. Rev. Lett., 90, 108501, doi:10.1103/PhysRevLett.90.108501, 2003.

Funder, S., Goosse, H., Jepsen, H., Kaas, E., Kjær, K. H., Korsgaard, N. J., Larsen, N. K., Linderson, H., Lyså, A., Möller, P., Olsen, J., and Willerslev, E.: A 10 000-year record of Arctic Ocean sea-ice variability: view from the beach, Science, 333, 747–750, doi:10.1126/science.1202760, 2011.

Gao, C., Robock, A., and Ammann, C.: Volcanic forcing of climate over the last 1500 years: an improved ice-core based index for climate models, J. Geophys. Res., 113, D2311, doi:10.1029/2008JD010239, 2008.

Giese, B. S. and Ray, S.: El Niño variability in simple ocean data assimilation (SODA), 1871–2008, J. Geophys. Res., 116, C02024, doi:10.1029/2010JC006695, 2011.

Gleckler, P. J., Taylor, K. E., and Doutriaux, C.: Performance metrics for climate models, J. Geophys. Res., 113, D06104, doi:10.1029/2007JD008972, 2008.

Guiot, J., Boreux, J. J., Braconnot, P., and Torre, F.: Data-model comparison using fuzzy logic in paleoclimatology, Clim. Dynam., 15, 569–581, 1999.

CPD

9, 775–835, 2013

Hall, A. and Qu, X.: Using the current seasonal cycle to constrain snow albedo feedback in future climate change, Geophys. Res. Lett., 33, L03502, doi:10.1029/2005GL025127, 2006.
Hansen, J., Mki. Sato, M., Ruedy, R., Kharecha, P., Lacis, A., Miller, R. L., Nazarenko, L., Lo, K., Schmidt, G. A., Russell, G., Aleinov, I., Bauer, S., Baum, E., Cairns, B., Canuto, V., Chandler, M., Cheng, Y., Cohen, A., Del Genio, A., Faluvegi, G., Fleming, E., Friend, A., Hall, T., Jackman, C., Jonas, J., Kelley, M., Kiang, N. Y., Koch, D., Labow, G., Lerner, J., Menon, S., Novakov, T., Oinas, V., Perlwitz, Ja., Perlwitz, Ju., Rind, D., Romanou, A., Schmunk, R., Shindell, D., Stone, P., Sun, S., Streets, D., Tausnev, N., Thresher, D., Unger, N., Yao, M., and Zhang, S.: Climate simulations for 1880–2003 with GISS ModelE, Clim. Dynam., 29, 661–696, doi:10.1007/s00382-007-0255-8, 2007.
Hargreaves, J. C., Paul, A., Ohgaito, R., Abe-Ouchi, A., and Annan, J. D.: Are paleoclimate model ensembles consistent with the MARGO data synthesis?, Clim. Past, 7, 917–933, doi:10.5194/cp-7-917-2011, 2011.
Hargreaves, J. C., Annan, J. D., Yoshimori, M., and Abe-Ouchi, A.: Can the last glacial maximum constrain climate sensitivity?, Geophys. Res. Lett., 39, L24702, doi:10.1029/2012GL053872, 2012a.
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 Discuss., 8, 3481–3511, doi:10.5194/cpd-8-3481-2012, 2012b.
Harrison, S. P. and Bartlein, P. J.: Records from the past, lessons for the future: what the palaeo-record implies about mechanisms of global change, in: The Future of the World's Climates, edited by: Henderson-Sellers, A. and McGuffie, K., Elsevier, 403–436, 2012.
Harrison, S. P., Bartlein, P. J., Brewer, S., Prentice, I. C., Boyd, M., Hessler, I., Holmgren, K., Izumi, K., and Willis, K.: Model benchmarking with glacial and mid-Holocene climates, Clim. Dynam., submitted, 2013.
Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in regional climate predictions, B. Am. Meteorol. Soc., 90, 1095–1107, doi:10.1175/2009BAMS2607.1, 2009.
Haylock, M. R., Hofstra, N., Tank, A. M. G. K., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, J. Geophys. Res., 113, D20119, doi:10.1029/2008JD010201, 2008.
Hirschi, M., Seneviratne, S. I., Alexandrov, V., Boberg, F., Boroneant, C., Christensen, O. B., Formayer, H., Orlowsky, B., and Stepanek, P.: Observational evidence for soil-moisture impact on hot extremes in southeastern Europe, Nat. Geosci., 4, 17–21, 2011.

CPD
9, 775–835, 2013

Using paleo-climate comparisons to constrain future projections in CMIP5

G. A. Schmidt et al.

Huang, S.: Merging information from different resources for new insights into climate change in the past and future, Geophys. Res, Lett., 31, L13205, doi:10.1029/2004GL019781, 2004.
IPCC SREX: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change, edited by: Field, C. B., Barros, V., Stocker, T. F., Qin, D., Dokken, D. J., Ebi, K. L., Mastrandrea, M. D., Mach, K. J., Plattner, G.-K., Allen, S. K., Tignor, M., and Midgley, P. M., Cambridge University Press, Cambridge, UK, and New York, NY, USA, 582 pp., 2012.
Izumi, K., Bartlein, P. J., and Harrison, S. P.: Consistent behaviour of the climate system in response to past and future forcing, Geophys. Res. Lett., submitted, 2013.
Jansen, E., Overpeck, J., Briffa, K., Duplessy, J.-C., Joos, F., Masson-Delmotte, V., Olago, D., Otto-Bliesner, B., Peltier, W., Rahmstorf, S., Ramesh, R., Raynaud, D., Rind, D., Solomina, O., Villalba, R., and Zhang, D.: Palaeoclimate, in: Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., and Miller, H., Cambridge University Press, Cambridge, 2007.
Jomelli, V., Grancher, D., Naveau, P., Cooley, D., and Brunstein, D.: Assessment study of lichenometric methods for dating surfaces, Geomorphology, 86, 131–143, 2007.
Jones, P. D., Briffa, K. R., Osborn, T. J., Lough, J. M., van Ommen, T. D., Vinther, B. M., Luterbacher, J., Wahl, E. R., Zwiers, F. W., Mann, M. E., Schmidt, G. A., Ammann, C. M., Buckley, B. M., Cobb, K. M., Esper, J., Goosse, H., Graham, N., Jansen, E., Kiefer, T., Kull, C., Küttel, M., Mosley-Thompson, E., Overpeck, J. T., Riedwyl, N., Schulz, M., Tudhope, A. W., Villalba, R., Wanner, H., Wolff, E., and Xoplaki, E.: High-resolution palaeoclimatology of the last millennium: A review of current status and future prospects, Holocene, 19, 3–49, doi:10.1177/0959683608098952, 2009.
Kageyama, M., Peyron, O., Pinot, S., Tarasov, P., Guiot, J., Joussaume, S., G. Ramstein: The Last Glacial Maximum climate over Europe and Western Siberia: a PMIP comparison between models and data, Clim. Dynam., 17, 23–43, 2001.
Kageyama, M., Braconnot, P., Bopp, L., Caubel, A., Foujols, M.-A., Guilyardi, E., Khodri, M., Lloyd, J., Lombard, F., Mariotti, V., Marti, O., Roy, T., and Woillez, M.-N.: Mid-Holocene and Last Glacial Maximum climate simulations with the IPSL model. Part I: compar-

ing IPSL_CM5A to IPSL_CM4, Clim. Dynam., online first, doi:10.1007/s00382-012-1488-8, 2012.
Kaplan, J. O., Krumhardt, K. M., Ellis, E. C., Ruddiman, W. F., Lemmen, C., and Klein Goldewijk, K.: Holocene carbon emissions as a result of anthropogenic land cover change, Holocene, 21, 775–791, doi:10.1177/0959683610386983, 2011.
Kharin, V. V., Zwiers, F. W., and Zhang, X. B.: Intercomparison of near-surface temperature and precipitation extremes in AMIP-2 simulations, reanalyses, and observations, J. Climate, 18, 5201–5223, 2005.
Kharin, V. V., Zwiers, F. W., Zhang, X. B., and Hegerl, G. C.: Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations, J. Climate, 20, 1419–1444, 2007.
Kohfeld, K. E. and Harrison, S. P.: How well can we simulate past climates? Evaluating earth system models using global palaeoenvironmental datasets, Quaternary Sci. Rev., 19, 321–346, 2000.
Kohfeld, K. E. and Harrison, S. P.: DIRTMAP: the geological record of dust, Earth Sci. Rev., 54, 81–114, 2001.
Köhler, P., Bintanja, R., Fischer, H., Joos, F., Knutti, R., Lohmann, G., and Masson-Delmotte, V.: What caused Earth's temperature variations during the last 800 000 years? Data-based evidence on radiative forcing and constraints on climate sensitivity, Quaternary Sci. Rev., 29, 129–145, doi:10.1016/j.quascirev.2009.09.026, 2010.
Knutti, R., Abramowitz, G., Collins, M., Eyring, V., Gleckler, P. J., Hewitson, B., and Mearns, L.: Good practice guidance paper on assessing and combining multi model climate projections, in: Meeting Report of the Intergovernmental Panel on Climate Change Expert Meeting on Assessing and Combining Multi Model Climate Projections, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., and Midgley, P. M., IPCC Working Group I Technical Support Unit, University of Bern, Bern, Switzerland, 2010a.
Knutti R., Furrer, R., Tebaldi, C., Cermak, J., and Meehl, G. A.: Challenges in combining projections from multiple climate models, J. Climate, 23, 2739–2758, 2010b.
Koenker, R.: Quantile Regression, Cambridge University Press, Cambridge, 2005.
Krivova, N. A., Balmaceda, L., and Solanski, S. K.: Reconstruction of solar total irradiance since 1700 from the surface magnetic field flux, Astron. Astrophys., 467, 335–346, 2007.
Kutzbach, J. E.: Monsoon climate of the early Holocene: climate experiment with the Earth's orbital parameters for 9000 years ago, Science, 214, 59–61, 1981.

CPD
9, 775–835, 2013

Using paleo-climate comparisons to constrain future projections in CMIP5

G. A. Schmidt et al.

Laîné, A., Kageyama, M., Braconnot, P., and Alkama, R.: Impact of greenhouse gas concentration changes on the surface energetics in the IPSL-CM4 model: regional warming patterns, land/sea warming ratio, glacial/interglacial differences, J. Climate, 22, 4621–4635, 2009.

Lambert, F. H., Webb, M. J., and Joshi, M. M.: The relationship between land–ocean surface temperature contrast and radiative forcing, J. Climate, 24, 3239–3256, doi:10.1175/2011JCLI3893.1, 2011.

Le Roy Ladurie, E.: Histoire humaine et comparée du climat, Canicules et glacier, XIIIè-XVIIIè siècle, Fayard, Paris, 2004.

Le Roy Ladurie, E.: Histoire humaine et comparée du climat, Disettes et révolutions 1740–1860, Fayard, Paris, 2006.

LeGrande, A. N. and Schmidt, G. A.: Global gridded data set of the oxygen isotopic composition in seawater, Geophys. Res. Lett., 33, L12604, doi:10.1029/2006GL026011, 2006.

LeGrande, A. N. and Schmidt, G. A.: Water isotopologues as a quantitative paleosalinity proxy, Paleoceanography, 26, PA3225, doi:10.1029/2010PA002043, 2011.

Ljundqvist, F. C.: A new reconstruction of temperature variability in the extra – tropical Northern Hemisphere during the last two millennia, Geogr. Ann. A, 92, 339–351, 2010.

Lovejoy, S. and Schertzer, D.: Scale invariance in climatological temperatures and the spectral plateau, Ann. Geophys. B, 4, 401–410, 1986.

Lovejoy, S. and Schertzer, D.: The Weather and Climate: Emergent Laws and Multifractal Cascades, Cambridge University Press, Cambridge, 660 pp., 2012.

Lovejoy, S., Schertzer, D., and Varon, D.: Stochastic and scaling climate sensitivities: solar, volcanic and orbital forcings, Geophys. Res. Lett., 39, L11702, doi:10.1029/2012GL051871, 2012.

McKay, J. L., de Vernal, A., Hillaire-Marcel, C., Not, C., Polyak, L., and Darby, D.: Holocene fluctuations in Arctic sea-ice cover: dinocyst-based reconstructions for the Eastern Chukchi Sea, Can. J. Earth Sci., 45, 1377–1397, 2008.

Mahlstein, I. and Knutti, R.: September Arctic sea ice predicted to disappear near 2 °C global warming above present, J. Geophys. Res., 117, D06104, doi:10.1029/2011JD016709, 2012.

MARGO Project Members: Constraints on the magnitude and patterns of ocean cooling at the Last Glacial Maximum, Nat. Geosci., 2, 127–132, doi:10.1038/NGEO411, 2009.

Mann, M. E., Zhang, Z., Hughes, M. K., Bradley, R. S., Miller, S. K., Rutherford, S., and Ni, F.: Proxy-based reconstructions of hemispheric and global surface temperature variations over

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

the past two millennia, P. Natl. Acad. Sci., 105, 13252–13257 doi:10.1073/pnas.0805721105, 2008.
Masson-Delmotte, V., Braconnot, P., Hoffmann, G., Jouzel, J., Kageyama, M., Landais, A., Lejeune, Q., Risi, C., Sime, L., Sjolte, J., Swingedouw, D., and Vinther, B.: Sensitivity of interglacial Greenland temperature and $\delta^{18}O$: ice core data, orbital and increased CO_2 climate simulations, Clim. Past, 7, 1041–1059, doi:10.5194/cp-7-1041-2011, 2011.
Massonnet, F., Fichefet, T., Goosse, H., Bitz, C. M., Philippon-Berthier, G., Holland, M. M., and Barriat, P.-Y.: Constraining projections of summer Arctic sea ice, The Cryosphere Discuss., 6, 2931–2959, doi:10.5194/tcd-6-2931-2012, 2012.
Meehl, G. A., Stocker, T. F., Collins, W. D., Friedlingstein, P., Gaye, A. T., Gregory, J. M., Kitoh, A., Knutti, R., Murphy, J. M., Noda, A., Raper, S. C., Watterson, I. G., Weaver, A. J., and Zhao, Z.-C.: Global climate projections. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., Cambridge University Press, Cambridge, UK and New York, NY, USA, 2007.
Moberg, A., Sonnechkin, D. M., Holmgren, K., and Datsenko, N. M.: Highly variable Northern Hemisphere temperatures reconstructed from low- and high-resolution proxy data, Nature, 433, 613–617, 2005.
Moros, M., Andrews, J. T., Eberl, D. D., and Jansen, E.: The Holocene history of drift ice in the northern North Atlantic – evidence for different spatial and temporal modes, Palaeoceanography, 21, PA2017, doi:10.1029/2005PA001214, 2006.
Mueller, B. and Seneviratne, S. I.: Hot days induced by precipitation deficits at the global scale, P. Natl. Acad. Sci., 109, 12398–12403, doi:10.1073/pnas.1204330109, 2012.
Müller, J., Wagner, A., Fahl, K., Stein, R., Prange, M., and Lohmann, G.: Towards quantitative sea ice reconstructions in the northern North Atlantic: a combined biomarker and numerical modelling approach, Earth Planet. Sc. Lett., 306, 137–148, 2011.
Nigam, S. and Ruiz-Barradas, A.: Seasonal hydroclimate variability over North America in global and regional reanalyses and AMIP Simulations: varied representation, J. Climate, 19, 815–837, doi:10.1175/JCLI3635.1, 2006.
Nogaj, M., Yiou, P., Parey, S., Malek, F., and Naveau, P.: Amplitude and frequency of temperature extremes over the North Atlantic region, Geophys. Res. Lett., 33, L10801, doi:10.1029/2005GL024251, 2006.

CPD
9, 775–835, 2013

O'ishi, R. and Abe-Ouchi, A.: Polar amplification in the mid-Holocene derived from dynamical vegetation change with a GCM, Geophys. Res. Lett., 38, L14702, doi:10.1029/2011GL048001, 2011.

Oppo, D. W., Schmidt, G. A., and LeGrande, A. N.: Seawater isotope constraints on tropical hydrology during the Holocene, Geophys. Res. Lett., 34, L13701, doi:10.1029/2007GL030017, 2007.

Osborn, T., Raper, S., and Briffa, K.: Simulated climate change during the last 1000 years: comparing the ECHO-G general circulation model with the MAGICC simple climate model, Clim. Dynam., 27, 185–197, 2006.

PALAEOSENS Project Members: Making sense of palaeoclimate sensitivity, Nature, 491, 683–691, doi:10.1038/nature11574, 2012.

Parey, S., Dacunha-Castelle, D., and Hoang, T. T. H.: Mean and variance evolutions of the hot and cold temperatures in Europe, Clim. Dynam., 34, 345–359, 2010a.

Parey, S., Hoang, T. T. H., and Dacunha-Castelle, D.: Different ways to compute temperature return levels in the climate change context, Environmetrics, 21, 698–718, 2010b.

Polyak, L., Alley, R. B., Andrews, J. T., Brigham-Grette, J., Cronin, T. M., Darby, D. A., Dyke, A. S., Fitzpatrick, J. J., Funder, S., Holland, M., Jennings, A. E., Miller, G. H., O'Regan, M., Savelle, J., Serreze, M., St. John, K., White, J. W. C., and Wolff, E.: History of sea ice in the Arctic, Quat. Sci. Rev., 29, 1757–1778, doi:10.1016/j.quascirev.2010.02.010, 2010.

Pongratz, J., Reick, C. H., Raddatz, T., and Claussen, M.: A reconstruction of global agricultural areas and land cover for the last millennium, Global Biogeochem. Cy., 22, GB3018, doi:10.1029/2007GB003153, 2008.

Prentice, I. C. and Harrison, S. P.: Ecosystem effects of CO_2 concentration: evidence from past climates, Clim. Past, 5, 297–307, doi:10.5194/cp-5-297-2009, 2009.

Prado, L. F., Wainer, I., Chiessi, C. M., Ledru, M.-P., and Turcq, B.: Mid-Holocene climate reconstruction for eastern South America, Clim. Past Discuss., 8, 5925–5961, doi:10.5194/cpd-8-5925-2012, 2012.

Quesada, B., Vautard, R., Yiou, P., Hirschi, M., and Seneviratne, S. I.: Asymmetric European summer heat predictabilityfrom wet and dry southern winters and springs, Natural Climate Change, 2, 736–741, doi:10.1038/nclimate1536, 2012.

Ramstein, G., Kageyama, M., Guiot, J., Wu, H., Hély, C., Krinner, G., and Brewer, S.: How cold was Europe at the Last Glacial Maximum? A synthesis of the progress achieved since

the first PMIP model-data comparison, Clim. Past, 3, 331–339, doi:10.5194/cp-3-331-2007, 2007.
Reichler, T. and Kim, J.: How well do coupled models simulate today's climate?, B. Am. Meteorol. Soc., 89, 303–311, 2008.
Rind, D. and Peteet, D.: Terrestrial conditions at the last glacial maximum and CLIMAP sea-surface temperature estimates: are they consistent?, J. Geophys. Res., 94, 12851–12871, 1985.
Rybski, D., Bunde, A., and von Storch, H.: Long-term memory in 1000-year simulated temperature records, J. Geophys. Res., 113, D02106, doi:10.1029/2007JD008568, 2008.
Santer, B. D., Taylor, K. E., Gleckler, P. J., Bonfils, C., Barnett, T. P., Pierce, D. W., Wigley, T. M. L., Mears, C., Wentz, F. J., Brüggemann, W., Gillett, N. P., Klein, S. A., Solomon, S., Stott, P. A., and Wehner, M. F.: Incorporating model quality information in climate change detection and attribution studies, P. Natl. Acad. Sci., 106, 14778–14783, 2009.
Schär, C., Lüthi, D., Beyerle, U., and Heise, E.: The soil-precipitation feedback: a process study with a regional climate model, J. Climate, 12, 722–741, 1999.
Schmidt, G. A: Enhancing the relevance of palaeoclimate model/data comparisons for assessments of future climate change, J. Quaternary Sci., 25, 79–87, doi:10.1002/jqs.1314, 2010.
Schmidt, G. A., Ruedy, R., Hansen, J. E., Aleinov, I., Bell, N., Bauer, M., Bauer, S., Cairns, B., Canuto, V., Cheng, Y., Del Genio, A., Faluvegi, G., Friend, A. D., Hall, T. M., Hu, Y., Kelley, M., Kiang, N. Y., Koch, D., Lacis, A. A., Lerner, J., Lo, K. K., Miller, R. L., Nazarenko, L., Oinas, V., Perlwitz, Ja. Perlwitz, Ju. Rind, D., Romanou, A., Russell, G. L., Sato, M., Shindell, D. T., Stone, P. H., Sun, S., Tausnev, N., Thresher, D., and Yao, M.-S.: Present day atmospheric simulations using GISS Model E: comparison to in-situ, satellite and reanalysis data, J. Climate, 19, 153–192, doi:10.1175/JCLI3612.1, 2006.
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

CPD
9, 775–835, 2013

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.

Seneviratne, S. I. and Koster, R. D.: A revised framework for analyzing soil moisture memory in climate data: Derivation and interpretation, J. Hydrometeorol., 13, 404–412, 2012.

Seneviratne, S. I., Lüthi, D., Litschi, M., and Schär, C.: Land-atmosphere coupling and climate change in Europe, Nature, 443, 205–209, 2006.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J.: Investigating soil moisture-climate interactions in a changing climate: a review, Earth Sci. Rev., 99, 125–161, 2010.

Seneviratne, S. I., Nicholls, N., Easterling, D., Goodess, C. M., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., Reichstein, M., Sorteberg, A., Vera, C., and Zhang, X.: Changes in climate extremes and their impacts on the natural physical environment, in: A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC SREX Report), edited by: Field, C. B., Barros, V., Stocker, T. F., Qin, D., Dokken, D. J., Ebi, K. L., Mastrandrea, M. D., Mach, K. J., Plattner, G.-K., Allen, S. K., Tignor, M., and Midgley, P. M., Cambridge University Press, Cambridge, 2012.

Shindell, D. T., Faluvegi, G., Unger, N., Aguilar, E., Schmidt, G. A., Koch, D. M., Bauer, S. E., and Miller, R. L.: Simulations of preindustrial, present-day, and 2100 conditions in the NASA GISS composition and climate model G-PUCCINI, Atmos. Chem. Phys., 6, 4427–4459, doi:10.5194/acp-6-4427-2006, 2006.

Silva Dias, P. L., Turcq, B., Silva Dias, M. A. F., Braconnot, P., and Jorgetti, T.: Mid-Holocene Climate of Tropical South America: A Model-Data Approach, in: Past Climate Variability in South America and Surrounding Regions, Dev. Paleoenviron. Res., 14, 259–281, Springer Netherlands, doi:10.1007/978-90-481-2672-9_11, 2009.

Steinhilber, F., Beer, J., and Fröhlich, C.: Total solar irradiance during the Holocene, Geophys. Res. Lett., 36, L19704, doi:10.1029/2009GL040142, 2009.

Stroeve, J. C., Kattsov, V., Barrett, A., Serreze, M., Pavlova, T., Holland, M., and Meier, W. N.: Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations, Geophys. Res. Lett., 39, L16502, doi:10.1029/2012GL052676, 2012.

Sundberg, R., Moberg, A., and Hind, A.: Statistical framework for evaluation of climate model simulations by use of climate proxy data from the last millennium – Part 1: Theory, Clim. Past, 8, 1339–1353, doi:10.5194/cp-8-1339-2012, 2012.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res., 106, 7183–7192, 2001.
Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An Overview of CMIP5 and the experiment design, B. Am. Meteorol. Soc., 93, 485–498, doi:10.1175/BAMS-D-11-00094.1, 2011.
Tremblay, L.-B, Huard, D., Schmidt, G. A., and de Vernal, A.: Mid Holocene constraints on future Arctic climate change, in preparation, 2013.
Trenberth, K. E. and Fasullo, J. T.: Simulation of present-day and twenty-first-Century energy budgets of the Southern Oceans, J. Climate, 23, 440–454, doi:10.1175/2009JCLI3152.1, 2010.
Vautard R., Yiou, P., D'Andrea, F., de Noblet, N., Viovy, N., Cassou, C., Polcher, J., Ciais, P., Kageyama, M., and Fan, Y.: Summertime European heat and drought waves induced by wintertime Mediterranean rainfall deficit, Geophys. Res. Lett., 34, L07711, doi:10.1029/2006GL028001, 2007.
Vieira, L. E. A., Solanki, S. K., Krivova, N. A., and Usoskin, I.: Evolution of the solar irradiance during the Holocene, Astron. Astrophys., 531, A6, doi:10.1051/0004-6361/201015843, 2011.
Vyushin, D. I., Kushner, P. J., and Zwiers, F.: Modeling and understanding persistence of climate variability, J. Geophys. Res., 117, D21106, doi:10.1029/2012JD018240, 2012.
Wilby, R. L. and Wigley, T. M. L.: Downscaling general circulation model output: a review of methods and limitations, Prog. Phys. Geogr., 21, 530–548, 1997.
Wittenberg, A. T.: Are historical records sufficient to constrain ENSO simulations?, Geophys. Res. Lett., 36, L12702, doi:10.1029/2009GL038710, 2009.
Yiou, P., Dacunha-Castelle, D., Parey, S., and Hoang, T. T. H.: Statistical representation of temperature mean and variability in Europe, Geophys. Res. Lett., 36, L04710, doi:10.1029/2008GL036836, 2009.
Yiou, P., Servonnat, J., Yoshimori, M., Swingedouw, D., Khodri, M., and Abe-Ouchi, A.: Stability of weather regimes during the last millennium from climate simulations, Geophys. Res. Lett., 39, L08703, doi:10.1029/2012GL051310, 2012.
Zhu, X., Fraederich, L., and Blender, R.: Variability regimes of simulated Atlantic MOC, Geophys. Res. Lett., 33, L21603, doi:10.1029/2006GL027291, 2006.

Table 1. List of models, institutions and experiments used in the analyses in this paper. Experiment names use the CMIP5 database shorthand, and run numbers are the "rip" coding for each experiment.

Model Name	Model Institution	Experiments	Run numbers
ACCESS-1.0	CSIRO (Commonwealth Scientific and Industrial Research Organisation, Australia), and BOM (Bureau of Meteorology, Australia)	Historical	r1i1p1
BCC-CSM1	Beijing Climate Center, China Meteorological Administration, China	piControl	r1i1p1
		midHolocene	r1i1p1
		rcp85	r1i1p1
CanESM2	Canadian Centre for Climate Modelling and Analysis, Canada	historical	r[1-5]i1p1
		rcp45	r[1-5]i1p1
CNRM-CM5	Centre National de Recherches Météorologiques / Centre Européen de Recherche et Formation Avancée en Calcul Scientifique, France	piControl	r1i1p1
		historical	r[1-10]i1p1
		midHolocene	r1i1p1
		lgm	r1i1p1
		1pctCO2	r1i1p1
		abrupt4xCO2	r1i1p1
		rcp45	r1i1p1
		rcp85	r1i1p1
CSIRO-Mk3-6-0	Commonwealth Scientific and Industrial Research Organisation in collaboration with the Queensland Climate Change Centre of Excellence, Australia	piControl	r1i1p1
		historical	r[1-10]i1p1
		midHolocene	r1i1p1
		rcp45	r[1-10]i1p1
		rcp85	r1i1p1
EC-EARTH	EC-Earth consortium	piControl	r1i1p1
		historical	r7i1p1
		midHolocene	r1i1p1
		rcp45	r[1,2,6-9,11,12, 14]i1p1
		rcp85	r1i1p1
FGOALS-g2	LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences; and CESS, Tsinghua University, China	piControl	r1i1p1
		midHolocene	r1i1p1
		rcp85	r1i1p1
GFDL-CM2.1	NOAA Geophysical Fluid Dynamics Laboratory, US	historical	r1i1p1
GFDL-ESM2G	NOAA Geophysical Fluid Dynamics Laboratory, US	piControl	r1i1p1
		historical	r1i1p1
		midHolocene	r1i1p1
		rcp85	r1i1p1
GFDL-ESM2M	NOAA Geophysical Fluid Dynamics Laboratory, US	piControl	r1i1p1
		historical	r1i1p1
		midHolocene	r1i1p1
		rcp85	r1i1p1
GISS-E2-H	NASA Goddard Institute for Space Studies, US	piControl	r1i1p1
		historical	r[1-5]i1p[12]

CPD
9, 775–835, 2013

Table 1. Continued.

Model Name	Model Institution	Experiments	Run numbers
GISS-E2-R	NASA Goddard Institute for Space Studies, US	piControl	r1i1p1, r1i1p141
		historical	r1i1p[12], r[45]i1p3
		past1000	r1i1p12[1-8]
		midHolocene	r1i1p1
		lgm	r1i1p15[01]
		1pctCO2	r1i1p1
		abrupt4xCO2	r1i1p1
		rcp45	r[1-5]i1p1
		rcp85	r1i1p1
HadCM3	Hadley Center, UK Met. Office, UK	historical	r[1-10]i1p1
HadGEM2-CC	Hadley Center, UK Met. Office, UK	piControl	r1i1p1
		historical	r1i1p1
		midHolocene	r1i1p1
		rcp45	r1i1p1
		rcp85	r1i1p1
HadGEM2-ES	Hadley Center, UK Met. Office, UK	piControl	r1i1p1
		historical	r1i1p1
		midHolocene	r1i1p1
		rcp45	r[1-3]i1p1
		rcp85	r1i1p1
INM-CM4	Institute for Numerical Mathematics, Russia	piControl	r1i1p1
		historical	r1i1p1
		midHolocene	r1i1p1
		rcp45	r1i1p1
		rcp85	r1i1p1
IPSL-CM5A-LR	Institut Pierre-Simon Laplace, France	piControl	r1i1p1
		historical	r[1-4]i1p1
		midHolocene	r1i1p1
		lgm	r1i1p1
		1pctCO2	r1i1p1
		abrupt4xCO2	r1i1p1
		rcp45	r[1-4]i1p1
		rcp85	r1i1p1
IPSL-CM5A-MR	Institut Pierre-Simon Laplace, France	piControl	r1i1p1
		historical	r1i1p1
		midHolocene	r1i1p1
		rcp45	r1i1p1
		rcp85	r1i1p1
MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies, Japan	piControl	r1i1p1
		midHolocene	r1i1p1
		lgm	r1i1p1
		past1000	r1i1p1
		1pctCO2	r1i1p1
		abrupt4xCO2	r1i1p1
		rcp85	r1i1p1

CPD
9, 775–835, 2013

Using paleo-climate comparisons to constrain future projections in CMIP5

G. A. Schmidt et al.

Table 1. Continued.

Model Name	Model Institution	Experiments	Run numbers
MIROC5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology, Japan	piControl	r1i1p1
		historical	r1i1p1
		midHolocene	r1i1p1
		rcp45	r[1-3]i1p1
		rcp85	r1i1p1
MPI-ESM-P	Max Planck Institute for Meteorology, Hamburg, Germany	piControl	r1i1p1
		historical	r1i1p1
		past1000	r1i1p1
		lgm	r1i1p1
		midHolocene	r1i1p1
		1pctCO2	r1i1p1
		abrupt4xCO2	r1i1p1
		rcp85	r1i1p1
MPI-ESM-LR	Max Planck Institute for Meteorology, Hamburg, Germany	piControl	r1i1p1
		historical	r[1-3]i1p1
		rcp85	r1i1p1
MRI-CGCM3	Meteorological Research Institute, Tsukuba, Japan	piControl	r1i1p1
		midHolocene	r1i1p1
		lgm	r1i1p1
		1pctCO2	r1i1p1
		abrupt4xCO2	r1i1p1
		rcp85	r1i1p1
NCAR-CCSM4	National Center for Atmospheric Research, US	piControl	r1i1p1
		midHolocene	r1i1p1
		lgm	r1i1p1
		1pctCO2	r1i1p1
		abrupt4xCO2	r1i1p1
		rcp45	r1i1p1
		rcp85	r1i1p1
NCAR-CESM1	National Center for Atmospheric Research, US	historical	r[5-7]i1p1
NorESM1-M	Norwegian Climate Centre, Norway	piControl	r1i1p1
		historical	r[1-3]i1p1
		midHolocene	r1i1p1
		rcp45	r1i1p1
		rcp85	r1i1p1

CPD
9, 775–835, 2013

Table 2. Correlation of different variables with future precipitation change in the RCP 8.5 scenario. Precipitation changes are defined as in Fig. 4: annual-mean precipitation averaged over 0°–8° N–60° W–50° W minus annual-mean precipitation averaged over 5° S–15° S–45° W–35° W; SST dipole changes are defined as the annual-mean change in SST averaged over 3° N–15° N–50° W–20° W minus the annual-mean change in SST averaged over 3° S–15° S–30° W–20° W (see boxes on Fig. 4b). Land surface warming is the annual-mean warming averaged over 15° S–0°–70° W–50° W. The double ITCZ index is defined as the annual-mean precipitation averaged over the Southern branch (7° S–3° S–35° W–20° W) minus the annual-mean precipitation averaged over the Northern branch (7° N–3° N–35° W–20° W).

Variable	correlation (r)	No. of models	p-value
midHolocene-piControl: Δ precip	0.93	9	0.0001
rcp85-piControl: Δ SST dipole	0.67	16	0.002
midHolocene-piControl: Δ SST dipole	−0.08	9	Not significant
amipFuture-amip: Δ precip	0.78	5	0.06
sstClim4xCO2-sstClim: Δ precip	0.92	7	0.002
sstClim4xCO2-sstClim: Δ SAT (land)	0.71	8	0.024
midHolocene-piControl: Δ SAT (land)	0.73	9	0.013
double ITCZ index in piControl	−0.66	16	0.003

CPD
9, 775–835, 2013

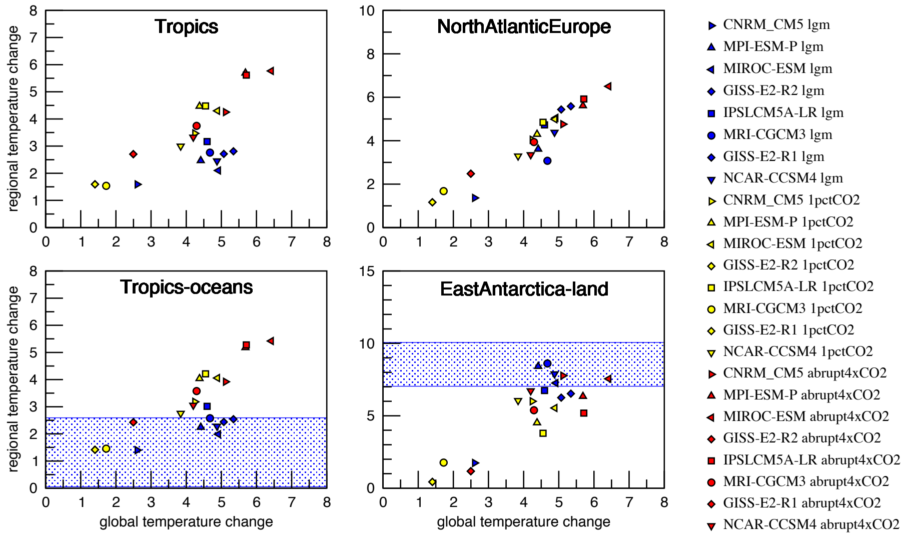

Fig. 1. Average regional temperature changes vs. global temperature changes for the glacial (in blue, the pre-industrial – LGM difference is shown), years 120 to 140 of the 1pctCO2 simulations (in yellow, in comparison to piControl) and the years 100 to 150 of the abrupt 4xCO2 simulations (in red, in comparison to piControl). For the bottom plots, the model averages are taken only from grid-boxes that correspond to proxy data sites within the defined region (reconstruction range shown in blue shading). Definition of the regions: Tropics: 23° S–23° N, North Atlantic Europe: 45° W–45° E, 30–50° N, East Antarctica: 5° W–165° E, 70–80° S. The results have been computed for all models in the database on 23 July 2012 for which there were results for the lgm, piControl, 1pctCO2 and abrupt4xCO2 simulations. Reconstructions are from the MARGO (2009) and Bartlein et al. (2011).

CPD
9, 775–835, 2013

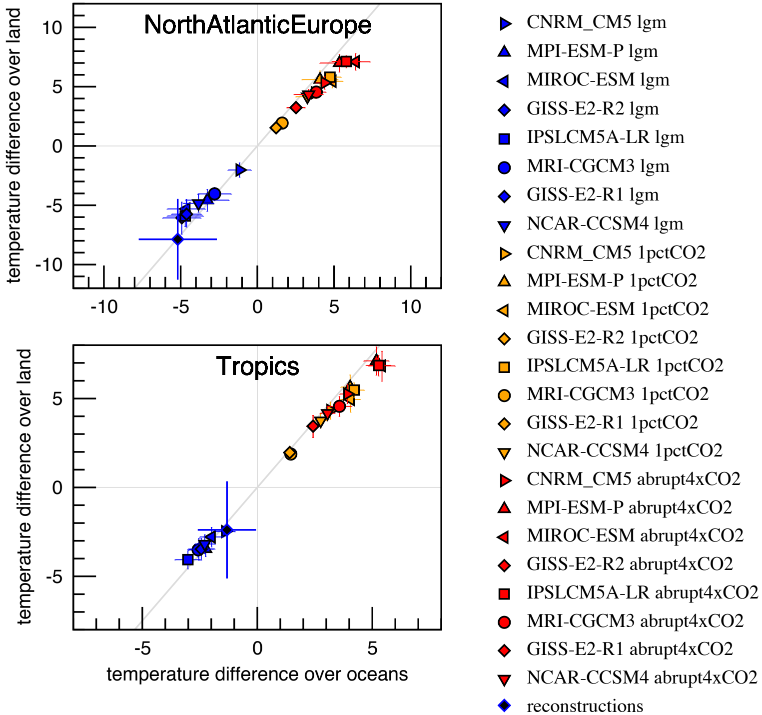

Fig. 2. Average surface air temperature change, compared to piControl, over land compared to over the oceans for the North Atlantic and Europe region (45° W–45° E, 30°–50° N) and the tropics (23° S–23° N). LGM – piControl in blue, 1pctCO2 – piControl in orange, abrupt4xCO2 – piControl in red. For the latter 2 periods, the averages have been computed over the same years as Fig. 1 above. The results have been computed for all models in the database on 23 July 2012 for which there were results for the lgm, piControl, 1pctCO2 and abrupt4xCO2. The grey lines indicate the 1 : 1.5 ratio in both plots. Reconstructions are as in Fig. 1.

CPD
9, 775–835, 2013

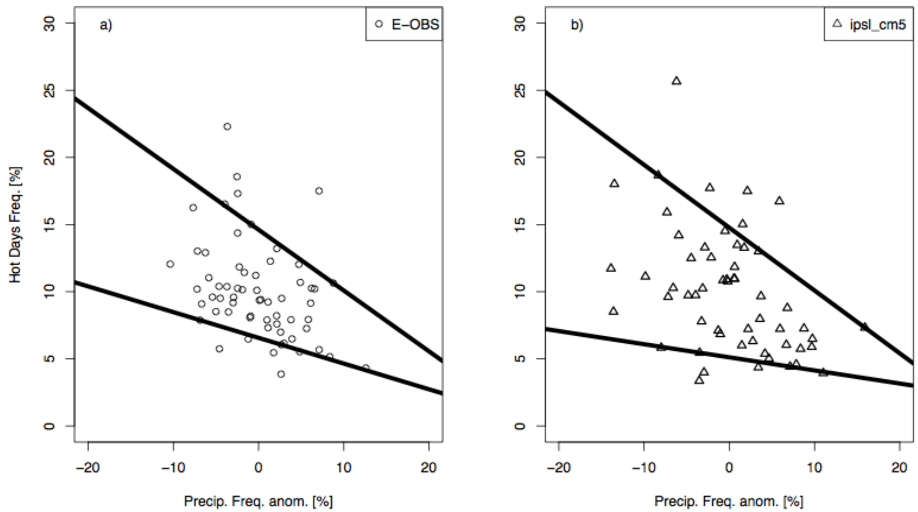

Fig. 3. Illustration of quantile regressions between the percentage of summer hot days in Europe (i.e. exceeding the 90th quantile of daily mean temperature in June-July-August) and the precipitation frequency anomaly with respect to the mean in winter-spring (January to May). The precipitation frequency is computed over southern Europe (36° N to 46° N) and is defined as the percentage of days with precipitation exceeding 0.5 mm. The quantile regressions are computed for the 10th and 90th quantiles of the hot day frequency, following Quesada et al. (2012). **(a)** shows the quantile regression for Western Europe from the EOBS dataset (Haylock et al., 2008) between 1950 and 2011 where each point represents a year. **(b)** is for the "historical" simulation (1960–2008) of the IPSL-CM5A-MR model (Dufresne et al., 2013). Both panels show a widening of the quantile regression for low values of precipitation frequency, indicating a consistency of the model simulation with observations.

CPD
9, 775–835, 2013

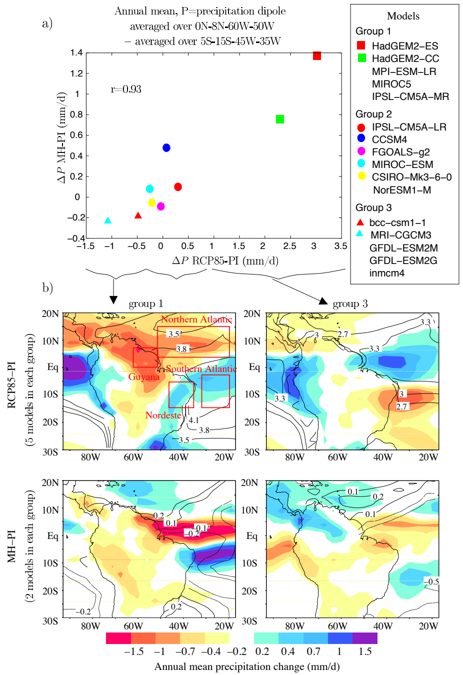

Fig. 4. (a) Relationship between the precipitation dipole change from pre-industrial to future climate under RCP 8.5 for the 2080–2100 and the precipitation dipole change from pre-industrial to mid-Holocene. Only those models within each group that had both rcp85 and midHolocene data available at the time are plotted, other models that provided only rcp85 data are listed for completeness. **(b)** Maps of precipitation changes from piControl to rcp85 (top) and from piControl to midHolocene (bottom) in average over all available models in group 1 (left) and in group 3 (right). Contours show corresponding SST changes. The boxes over land and ocean show the areas used in the dipole definitions.

CPD
9, 775–835, 2013

Using paleo-climate comparisons to constrain future projections in CMIP5

G. A. Schmidt et al.

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀| |▶
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

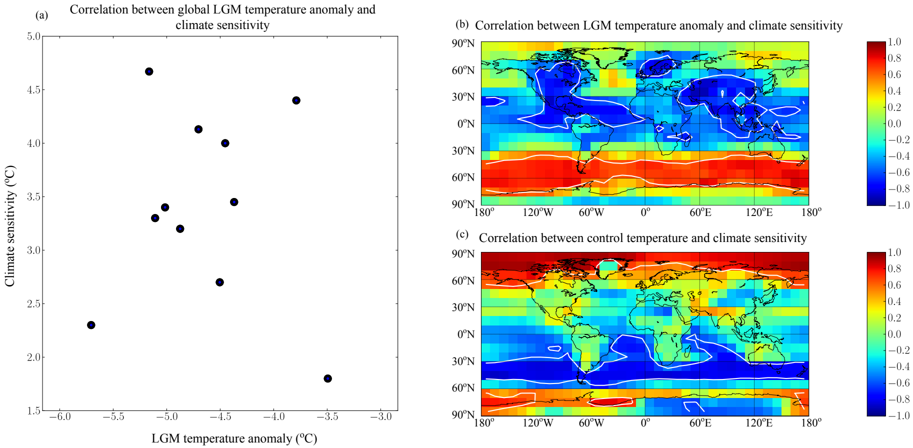

Fig. 5. (a) Global mean LGM temperature change versus overall climate sensitivity to 2xCO2. **(b)** Correlation between local air temperature anomaly and climate sensitivity across the model ensemble. **(c)** Correlation across the model ensemble between control run temperatures and climate sensitivity.

CPD
9, 775–835, 2013

Using paleo-climate comparisons to constrain future projections in CMIP5

G. A. Schmidt et al.

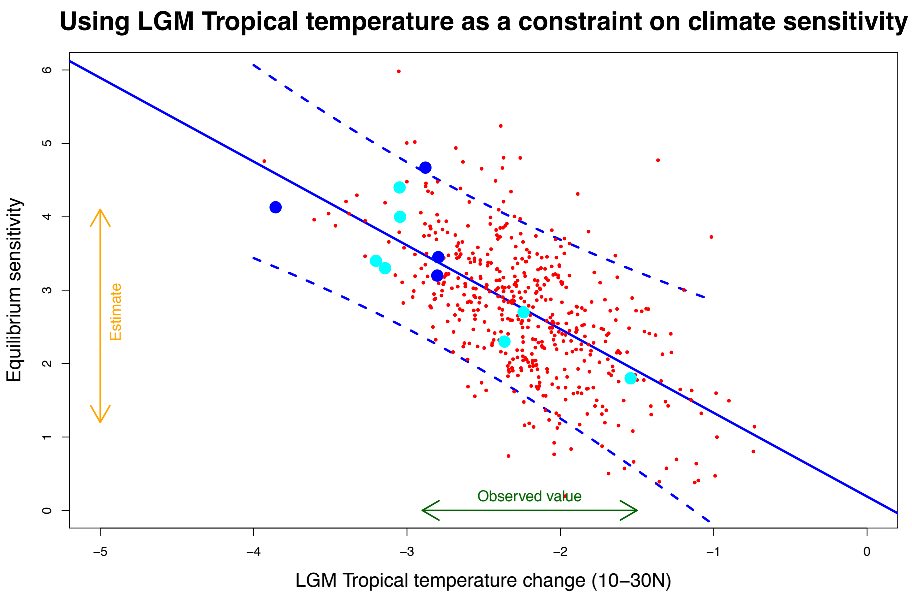

Fig. 6. Using LGM Tropical temperature as a constraint on climate sensitivity. Cyan and blue dots represent PMIP2 and CMIP5 simulations. Linear correlation and predictive uncertainty range are plotted as solid and dashed blue lines respectively. Small red dots represent a Monte Carlo sample from the estimated proxy-derived observational value, mapped onto the climate sensitivity.

CPD
9, 775–835, 2013

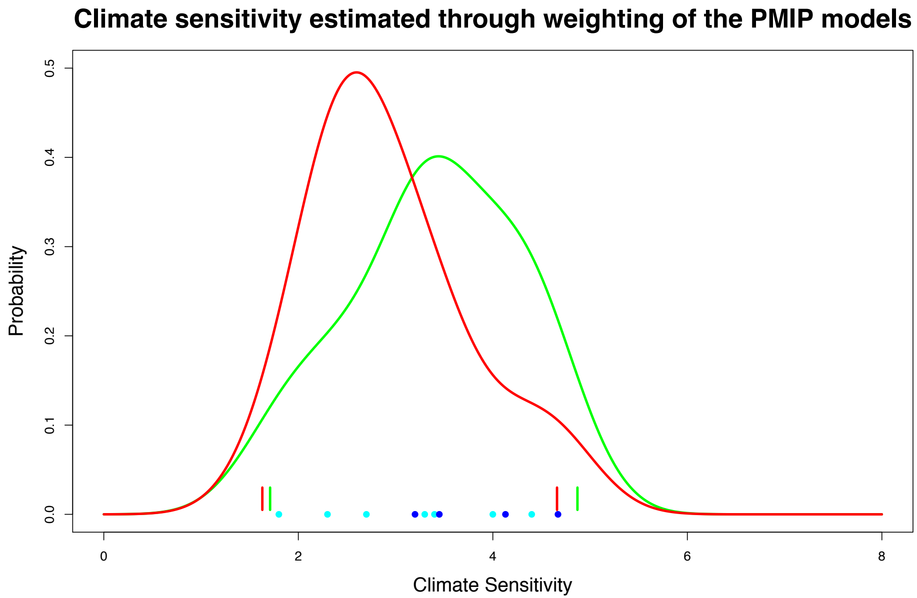

Fig. 7. Climate sensitivity estimated through weighting of the PMIP models. Cyan and blue dots represent PMIP2 and CMIP5 simulations. Green curve shows prior distribution of climate sensitivity (based on equal weighting of the models). Red curve shows posterior distribution, after weighting according to match to the LGM tropical temperature. Vertical bars indicate 5–95 % ranges.

CPD
9, 775–835, 2013

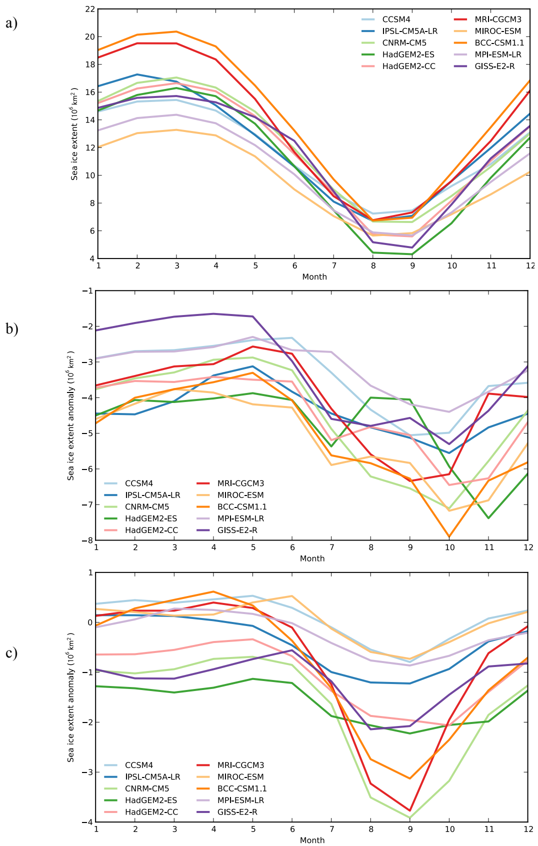

Fig. 8. Sea-ice extent in CMIP5 models in $10^6\,km^2$. **(a)** 30-yr mean seasonal cycle for the historical period (1870–1900), **(b)** the anomaly in sea ice extent for the period 2036–2065 in RCP 8.5, and **(c)** the anomaly at the mid-Holocene.

CPD
9, 775–835, 2013

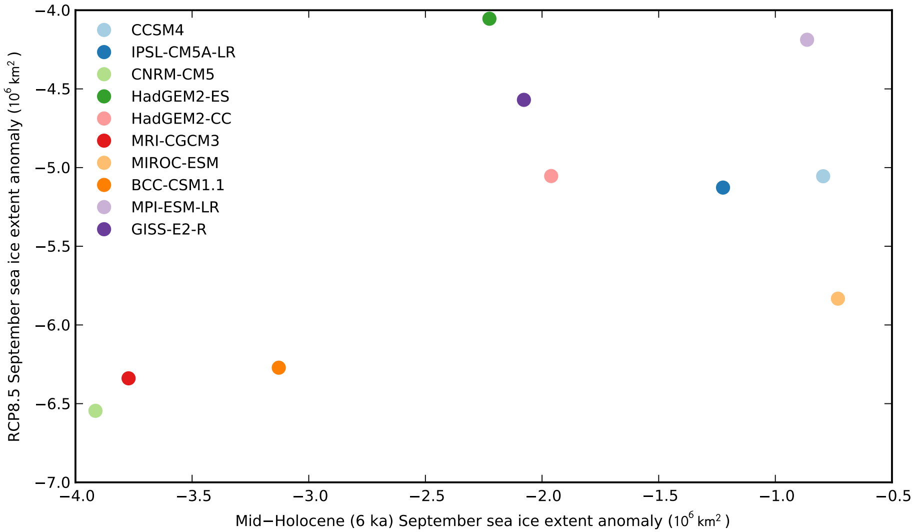

Fig. 9. Relationship between the September MH anomaly and the September RCP 8.5 anomaly across the CMIP5 models.

CPD
9, 775–835, 2013

Using paleo-climate comparisons to constrain future projections in CMIP5

G. A. Schmidt et al.

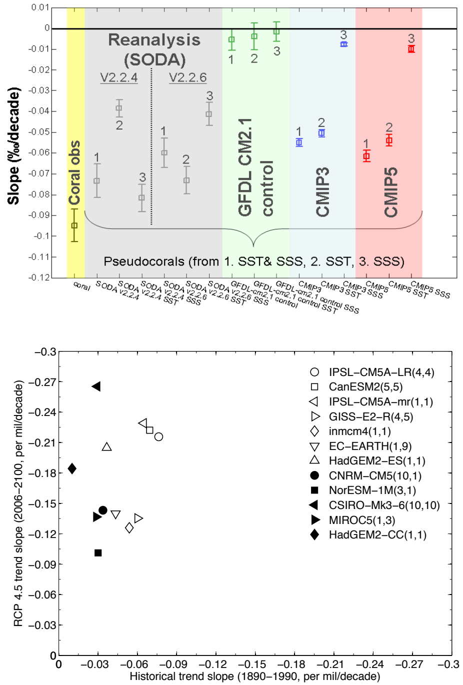

Fig. 10. Upper panel: Magnitude of the trend in $\delta^{18}O_c$ (‰/decade, computed from a simple linear regression through the trend PC) in corals (far left), Simple Ocean Data Assimilation (SODA) 20th-century reanalysis (Carson and Giese, 2008; Giese and Ray, 2011; Compo et al., 2011), a 500-yr control run from GFDL CM2.1 (Wittenberg, 2009), and the CMIP3 and CMIP5 multi-model ensembles. In each case, $\delta^{18}O_c$ was modeled from SST and SSS (1), SST only (2), and SSS (3). Lower Panel: Magnitude of the $\delta^{18}O_c$ trend (‰/decade, computed from a simple linear regression through the trend PC) over 1890–1990 in pseudocorals modeled from CMIP5 historical simulations and over 2006–2100 in the RCP 4.5 projections where numbers in parenthesis indicate the number of runs in the historical and RCP4.5 ensemble, respectively.

CPD
9, 775–835, 2013

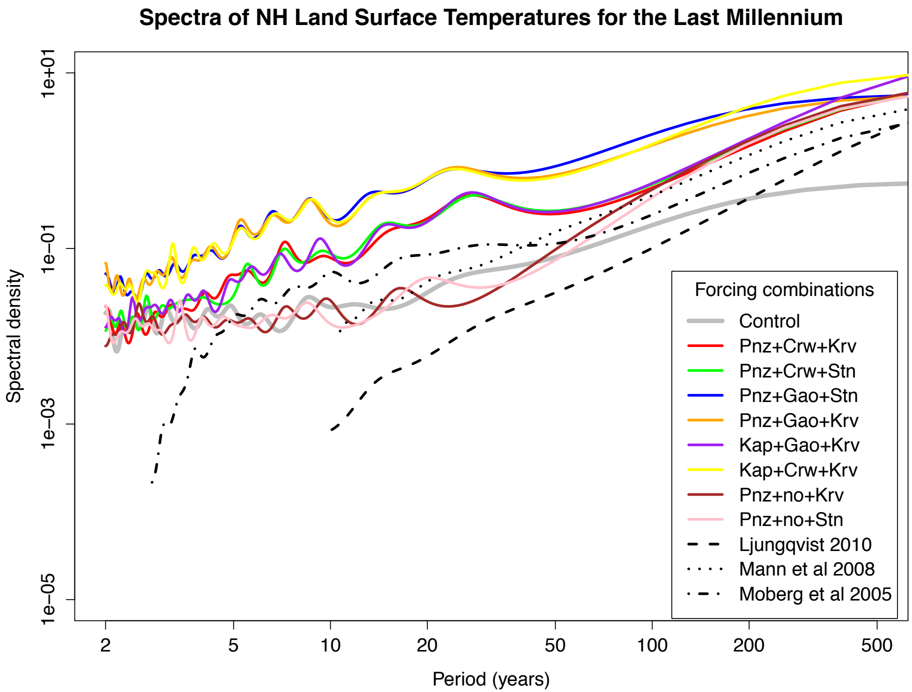

Fig. 11. Spectra from an ensemble of LM simulations using the same model but driven with different sets of forcings compared with Ljundqvist (2010), Mann et al. (2008) and Moberg et al. (2006) reconstructions. The clustering of simulations is driven entirely by changes in the volcanic forcing dataset used, with the simulations with the most decadal and multi-decadal variability using the Gao et al. (2008) reconstruction. Only in the examples where no volcanic forcing is used at all is the impact of different solar forcing reconstructions detectable. Spectra derived using MEM with 30 poles, from 850 to 2005, after correction for control run drift using a loess low-frequency estimate derived from the control run. Key abbreviations: Land use: Pnz (Pongratz et al., 2008), Kap (Kaplan et al., 2011); Solar: Krv (Vieira et al., 2011), Stn (Steinhilber et al., 2009); Volcanic: Gao (Gao et al., 2008), Crw (Crowley et al., 2008).

CPD
9, 775–835, 2013

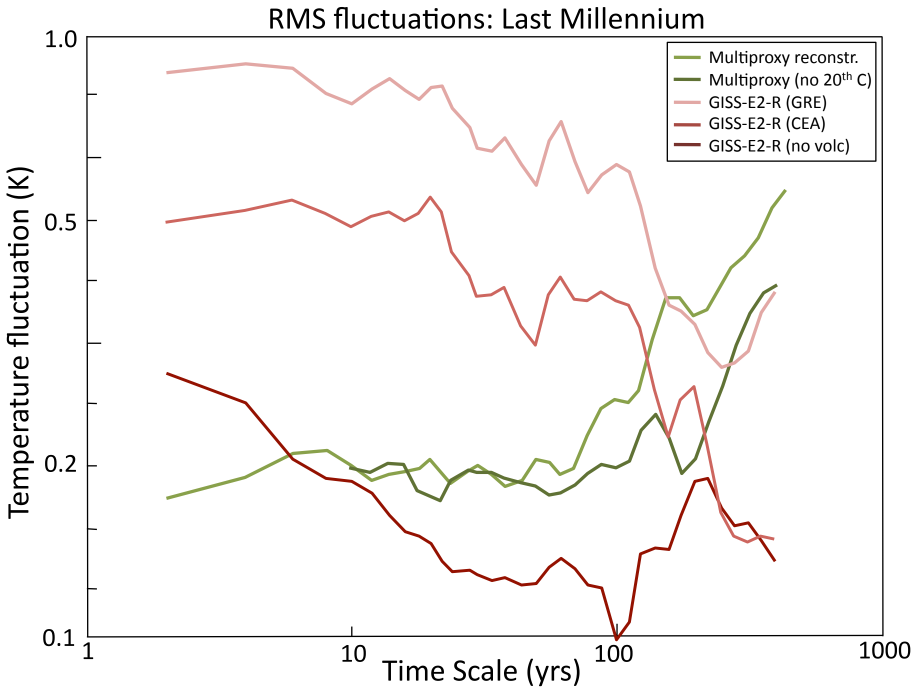

Fig. 12. RMS fluctuations of instrumental and paleoclimate reconstructions compared to simulations of the Northern Hemisphere land temperature for the period 1500–1900. GRE, CEA refer to GISS-E2-R simulations using the Gao et al., (2008), and Crowley et al. (2008) reconstructions of volcanic forcing. The multiproxy reconstruction used is an average of three NH estimates, and the RMS fluctuations are separately shown for the periods 1000–1900 and 1000–1980.

CPD
9, 775–835, 2013

Using paleo-climate comparisons to constrain future projections in CMIP5

G. A. Schmidt et al.

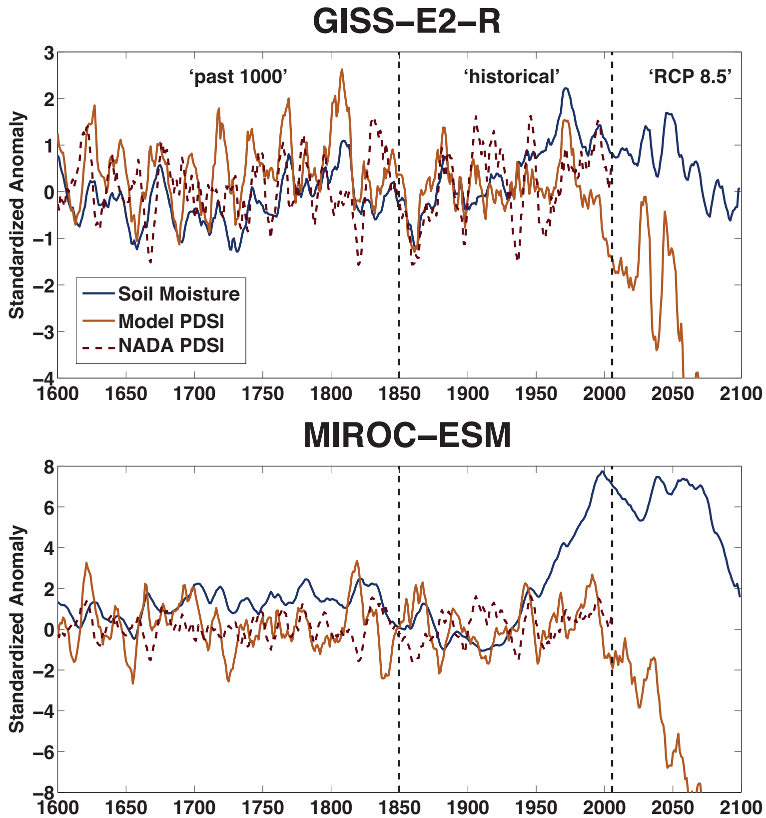

Fig. 13. Standardized anomalies for PDSI and soil moisture in two models (GISS-E2-R and MIROC-ESM) using a past1000 simulation, and a historical+rcp85 continuation. For reference, the tree-ring based reconstruction is plotted (dashed-line) (Cook et al., 2010), though this would not be expected to line up exactly with the model simulations. All data smoothed with a 10-yr running mean.