Climate model benchmarking with glacial and mid-Holocene climates: Supplementary Information

Harrison, S.P.[1,2], Bartlein, P.J.[3], Brewer, S.[4], Prentice, I.C.[1,5], Boyd, M.[6,7], Hessler, I.[8,1], Holmgren, K.[6,7], Izumi, K.[3], Willis, K.[1]

1: Department of Biological Sciences, Macquarie University, North Ryde, NSW 2109, Australia.

2: Geography & Environmental Sciences, School of Human and Environmental Sciences, Reading University, Whiteknights, Reading, UK

3: Department of Geography, University of Oregon, Eugene, Oregon, USA

4: Geography Department, University of Utah, Salt Lake City, Utah, USA

5: AXA Chair of Biosphere and Climate Impacts, Department of Life Sciences and Grantham Institute for Climate Change, Imperial College, Silwood Park, Ascot SL5 7PY, UK

6: Bert Bolin Centre for Climate Research, Stockholm University, 106 91 Stockholm, Sweden

7: Department of Physical Geography and Quaternary Geology, Stockholm University, 106 91 Stockholm, Sweden

8: MARUM, Centre for Marine Environmental Sciences, University of Bremen, Germany

Correspondence to: sandy.harrison@mq.edu.au

The SI contains a description of (1) the derivation of the benchmark data sets, (2) the modelling protocol used in PMIP2 and CMIP5 simulation, (3) the processing of model output, (4) the metric statistics, and (5) tables summarizing the metrics for each time period and model.

1. Derivation of Benchmark Data Sets

1.1. Bioclimatic variables over land: Bartlein et al. (2011)

There has been a long history of making quantitative climate or bioclimate reconstructions using pollen and plant macrofossil data (Webb and Bryson 1972; Bartlein et al. 1984; Overpeck et al. 1985; Bartlein et al. 1986; Huntley and Prentice 1988), using different statistical methods and inverse modelling. Bartlein et al. (2011) provide a global synthesis by combining and gridding existing regional data sets available for the Mid-Holocene (MH, ca 6000 yr BP) and Last Glacial Maximum (LGM, ca 21,000 yr BP). The original data were screened by applying age-selection and statistical outlier criteria. Bartlein et al. (2011) includes reconstructions of six variables: (1) mean temperature of the warmest month (MTWA), (2) the accumulated temperature sum during the growing season (growing degree days above a baseline temperature for the growth of woody plants of 5°C: GDD5), (3) mean temperature of the coldest month (MTCO), (4) mean annual temperature (MAT), (5) mean annual precipitation (MAP) and (6) the ratio of actual to equilibrium evapotranspiration (α). Analyses of these data in the original Bartlein et al. (2011) paper have shown that methodological uncertainties have little impact on the reconstructions: the reconstructed climate variables show large, spatially coherent patterns that are consistent with plausible responses to known climate forcing. Nevertheless, the dataset includes a quantified estimate of the combined uncertainties arising from data

resolution and sampling, age scale uncertainties, analytical uncertainties and calibration model uncertainty. To facilitate comparison with climate-model outputs, the original site-based reconstructions were re-gridded as anomalies on a regular latitude/longitude grid of 2° by 2° (comparable to the typical grid size of the models). The grid-cell anomaly value was obtained by simple averaging and grid-cell uncertainty is a pooled estimate of the standard error. Gridding results in the creation of climate estimates for 715 cells (out of a possible 3,687 non-ice-covered land cells) at MH and for 153 cells (out of a possible 3,618 non-ice-covered land cells) at LGM.

1.2. Surface ocean reconstructions: the MARGO data set: MARGO Project Members (2009)

Reconstructions of seasonal sea-surface temperatures (SSTs) and seasonal sea-ice cover at the LGM have been developed by the MARGO (Multiproxy Approach for the Reconstruction of the Glacial Ocean) project (Kucera et al. 2005) using biological and geochemical evidence from deep-sea sediments. MARGO initially produced data sets for different types of records and for different regions, with measurement and statistical uncertainties given for at each site. Different types of records give different information about surface ocean conditions: for example, foraminiferal assemblages can be analyzed statistically to yield seasonal variation in SSTs while alkenone data only provide a measure of mean annual sea-surface temperature. The various data sets were subsequently combined to provide gridded data sets of summer SST (SSTsum), winter SST (SSTwin), annual SST (SSTann) at 5° by 5° resolution, where summer is defined as July, August, September in the northern hemisphere and January, February, March in the southern hemisphere (and vice versa for winter). The gridded data sets were produced by averaging individual site-based reconstructions that fall into the same cell, weighted by an index of the reliability of each contributing reconstruction (MARGO Project Members 2009). Uncertainties are calculated from the calibration error of each proxy, weighted by an index based on dating quality, number of samples and the proxy reliability. This is aggregated with the standard deviation of climate estimations for all records in a grid cell, to provide an overall grid uncertainty.

There are three data sets that provide reconstructions of some aspect of sea ice cover in MARGO. One simply estimates the extent of winter and summer sea ice based on foraminifer-based SST reconstructions exceeding specified threshold values (MARGO Project Members 2009). The diatom-based reconstructions for the Southern Ocean (Gersonde et al. 2005) include estimates of average concentration per month. The dinoflagellate data set for the North Atlantic (de Vernal et al. 2005) only provides the number of months per year with sea ice. We therefore do not attempt to use seasonal measures of sea ice for the model evaluations. The de Vernal et al. (2005) data set uses a sea-ice concentration threshold of 0.50 to distinguish ice-free from ice-covered states under present conditions, while the Gersonde et al. (2005) data set uses a concentration of 0.40 to translate concentration to months of coverage. Gersonde et al, (2005) note that the ice limit is well matched by a concentration value between 0.40 and 0.50, suggesting that any bias between the different thresholds used in these two studies will be minimal. We used a value of 0.40 to encode model simulations of sea-ice concentrations into months of sea-ice coverage (SInmon), which is consistent with the common definition (or approach) for translating concentration to months of coverage (Fowler et al 2004; Maslanik et al. 2011). Thus, the gridded sea ice data are expressed as the number of months with

sea ice cover >40% (SInmon); the change in the number of months at the LGM has been calculated based on a long-term average (1971-2000) of modern observations of sea-ice cover from the HADISST data set (Rayner et al. 2003).

1.3. Surface ocean reconstructions: the GHOST data set

The GHOST (Global Holocene Spatial and Temporal Climate Variability) project has produced reconstructions of mean annual SST for the MH based on surface-dwelling planktonic foraminiferal Mg/Ca and alkenone measurements (Kim et al. 2004; Leduc et al. 2010). Alkenone and Mg/Ca reconstructions are derived from empirical calibrations derived under controlled environmental conditions. Although several different calibrations are used for both the alkenone and Mg/Ca estimates, following the original authors, the uncertainty resulting from this is small. The measurement error on individual reconstructions is generally estimated as about 1° C. To ensure compatibility with the MARGO dataset, we recalculated the uncertainties for the GHOST dataset, using the MARGO approach described above. The reconstructions are given for individual sites, and estimates of the uncertainties are also given. The MH time-slice reconstruction is based on averaged values of samples falling in the1000-year window around 6000 yr B.P. Anomalies are calculated relative to the average SST over the last 1000 years. Gridded values on the 5° by 5° resolution ocean grid were obtained by simple averaging and grid-cell uncertainty is a pooled estimate of the standard error.

1.4. Climate reconstructions from ice cores: Braconnot et al. (2012)

Temperature changes at the LGM (defined as the interval between 20.5 to 21.5 ka, and expressed as anomalies with respect to the mean values of the last millennium) have been obtained from isotopic records, corrected for changes in sea-water isotopic composition, from five ice cores in East Antarctica (Vostok, EDC, EDML, TALDICE and Dome F). The estimates take into account the difference in ice-sheet elevation between the LGM and present. The temperature estimates are based on the use of the present-day isotope-temperature gradient (0.8 ‰ per °C for $\delta^{18}O$), and as a result of the ±0.1 ‰ change in slope at the LGM results in an uncertainty on the LGM estimates of ± 2°C. Corrections for colder LGM moisture sources (indicated by deuterium excess data) are only available for the EDC and EDML ice cores, but could reduce the magnitude of the reconstructed cooling at other sites by approximately 1°C (Stenni et al. 2010). The most likely range of LGM Antarctic fixed-elevation temperature change, taking into account the differences between the estimates on individual cores (1.4°C), the use of the modern isotope-temperature slope (2°C), the change in elevation (1°C) and moisture source effects (1°C), is –6°C to –10°C (*57*). In this compilation, we use the original reconstructed anomaly for each site with an error bar of ± 2°C (Table S1).

1.5. Other quantitative climate reconstructions: Schmittner et al. (2011)

Schmittner et al. (2011) compiled quantitative estimates from the literature of surface temperature anomalies for the LGM. Many of these sites are included in the data sets described above. However, there are five terrestrial sites and 25 marine sites included in Schmittner et al. (2011) that are not available from other sources (Table S2). These data have been included in our data sets for completeness.

1.6. Climate reconstructions from speleothems

Calcium carbonate precipitates – speleothems – in limestone caves can provide precisely dated archives containing information about long-term climate variability (Fairchild et al. 2006; Fairchild and Treble 2009; Lachniet 2009). Significant contributions have been made by studies of stable isotope composition, trace elements, luminescent laminae and calcite fabrics of speleothems (mainly stalagmites). Interpretation requires information about site-specific conditions. Monitoring of present-day speleothem-forming processes is an important tool (Tremaine et al. 2011). Most commonly $\delta^{18}O$ in speleothems is interpreted as reflecting $\delta^{18}O$ composition of meteoric water, which in turn reflects surface temperatures (especially at high latitudes) and/or amount effects (especially in tropical oceanic sites and monsoon regions). However, other processes also influence the $\delta^{18}O$ composition in precipitation, such as changes in atmospheric circulation patterns or in climate modes which may lead to changing temperature/$\delta^{18}O$ relationship (Dayem et al. 2010; Pausata et al. 2011). Although the stable carbon isotope composition of speleothems has been regarded as more ambiguous than that of oxygen isotopes, an increasing number of studies demonstrate the potential of stable carbon-isotope variations to provide additional information on climate and environment (Dreybrodt and Scholz 2011). Measured variations can be a function of the amount of C_3 versus C_4 vegetation, vegetation cover, soil biological activity, bedrock proportion, rainfall amount and drip rate. Speleothem records are dated using uranium-thorium dating (U-Th dating) methods, which provide ages in calendar years and normally with small uncertainties (Dorale et al. 2004). Larger uncertainties are introduced if the sample is contaminated by detrital ^{230}Th. A correction factor then has to be applied (Hellstrom 2006).

We reviewed published speleothem records covering one or both of the MH and LGM time periods. We evaluated the reliability of each record, using the following criteria:

a) The record has a reliable age model, i.e. it is based on U-Th dating, using Thermal Ionization Mass Spectrometry (TIMS) or Inductively Coupled Plasma Mass Spectrometry (ICP-MS) techniques, it yields most ages in stratigraphic order and the uncertainty intervals of individual dates are less than the age intervals between dates. Records with large uncertainty intervals are included if a change in the age model within the uncertainty interval does not affect the climate interpretation.

b) The record has bracketing dates within 1000 yr of 6000 yr BP and within 2000 yr of 21000 yr BP.

c) The record lacks bracketing dates within 1000 yr of 6000 yr BP or within 2000 yr of 21000 yr BP, but has a climate signal that is consistent throughout the period bracketed within the closest dates.

d) The interpretation of the record in terms of temperature and/or precipitation changes is robust.

e) The record is related to present-day conditions, shown either by (1) continuous speleothem growth until present day; (2) long-term monitoring of present day processes clearly distinguishing the present conditions from the past;

(3) a halt in growth until the present indicating wetter conditions during past growth (applied only in one case: Yudamutana Cave). Records that do not relate to present-day conditions are listed in Table S3 but were not used in further evaluation.

f) The record yields quantitative information. Methods for retrieving quantitative data include the analysis of fluid inclusions, clumped isotope thermometry and calibration using instrumental record and or other quantitative climate records.

g) The record yields qualitative information. Reliable interpretation of retrieved data in terms of relative changes in past temperature and/or precipitation relies on the understanding of the relationship between the data measured ($\delta^{18}O$, $d^{13}C$, trace elements etc.) and atmospheric processes, and requires a present-day benchmark to indicate the response to known conditions.

Speleothem records from a total number of 65 caves met the above criteria. Of these 28 did not relate to the present-day conditions and were not used in further analyses (Table S3). Of the remaining 37 records (Table S4; Figure S1), most yield information on relative precipitation changes while fewer records provide information on relative temperature changes. Quantitative estimates are available from only 6 caves.

1.7. Combined climate reconstructions data set

The combined data sets provide 10 climate benchmarks: MAT, MTCO, MTWA, GDD5, MAP, α over land and SSTsum, SSTwin, SSTann, SInmon for the oceans. We have combined the data from all sources to produce new estimates on a 2° by 2° resolution land or ocean grid, again using simple averaging to derive a grid-cell value and the pooled estimate of the standard error to derive the uncertainty. The addition of a comparatively small number of extra data points to the existing gridded data sets does not change the patterns or magnitudes of the reconstructions, but ensures that the current data set is as comprehensive as possible. The final gridded data sets for each benchmark for each period are shown in Figures S2 through S5.

2. Climate modelling protocol

The MH and LGM simulations are equilibrium experiments, presenting a “snapshot” of climate at a specific time. Table S5 shows the boundary conditions used for MH and LGM experiments in the second phase of the Palaeoclimate Modelling Intercomparison Project (PMIP2) compared to the boundary conditions used for these experiments in the Coupled Modelling Intercomparison Project (CMIP5). The ultimate external forcing of climate is change in incoming solar radiation (insolation) as determined by changes in the Earth’s orbit. Changes in insolation can be specified precisely. Due to the slow variations of Earth’s orbital parameters, the seasonal and latitudinal distribution of MH insolation was different from present, enhancing the magnitude of the seasonal contrast in the Northern Hemisphere by about 60 Wm^{-2}. Insolation forcing at the LGM was very similar to present. Atmospheric

composition (long-lived greenhouse gas concentrations) during the MH and at the LGM are prescribed to conform to ice-core measurements. Changes in ice sheet extent can be specified from geomorphic evidence, but there are several possible configurations of LGM ice-sheet orography that are consistent with relative sea-level and geoidal constraints. In the PMIP2 experiments, the LGM ice sheets were specified from Peltier (2004). The CMIP5 ice sheets are a blended product made from three recent ice-sheet reconstructions (Braconnot et al. 2012). Although there were some small ice-sheet relicts in eastern Canada at 6 ka, they have been ignored in both the PMIP2 and CMIP5 experiments. Some of the models in PMIP2 simulated vegetation dynamics explicitly (i.e. were fully coupled ocean-atmosphere-vegetation general circulation models, OAVs) but vegetation cover and albedo had to be specified for the coupled ocean-atmosphere general circulation models (OAs). Processes associated with the terrestrial and marine carbon cycle were ignored in the PMIP2 experiments, but are included as interactive components of some of the models (here designated as OACs) used in CMIP5.

3. Processing climate model output

We use outputs from all of the models archived at either PMIP2 or CMIP5 on August 15th 2012 (Table S6). One model (CSIRO3.1.2) has performed the CMIP5 MH experiment, but is not an official CMIP5 simulation because the mandatory future experiments have not been made with this model version. Nevertheless, for simplicity, we refer to this simulation as a CMIP5 paleo-experiment. Climate simulations of the MH are available for 13 OAs and 6 OAVs from the PMIP2 archive and 10 OA simulations and 5 coupled carbon-cycle (OAC) simulations from the CMIP5 archive. Climate simulations of the LGM are available for 7 OA simulations and 2 OAV simulation from the PMIP2 archive, and 3 OA and 3 OAC simulations from the CMIP5 archive. No MH SSTann data were archived by August 15th 2012 for four models: ECHAM, EARTH, FGOALSG2, and FGOALS2. For these models, comparisons were therefore restricted to land variables. There were no LGM sea-ice data archived for the PMIP2 IPSL4 model, and no SST or sea-ice data for the COSMOS model by August 15th 2012. For these two models, comparisons were restricted to the subset of variables available.

We analyzed five climate variables from these archives: near-surface air temperature (*tas*), precipitation flux (*pr*), cloud-area fraction (*clt*), sea-surface temperature (*tos*), and sea-ice fraction (*sic*). The first three variables were used to create bioclimatic variables, while the last two were used to characterize SSTs and sea ice. The general strategy involves the calculation of long-term means from the archived time-series data, interpolation of the means on to a 0.5° grid, the calculation of variables directly comparable to paleoclimatic reconstructions, and the aggregation of those data to coarser-resolution grids. For the variables used to calculate bioclimatic variables, calculation on the 0.5° grid is important for propagating the observed spatial variability of climate (sensed by the paleodata) into the coarser-resolution simulations, while for SST and (which are sometimes calculated on rotated-pole or other non-latitude-by-longitude grids), interpolation on to a common grid facilitates the creation of ensembles, and aggregation to resolutions appropriate for comparisons with the paleo-reconstructions. The first analysis steps were identical for each variable, while subsequent steps treated near-surface temperature, precipitation and clouds as a group, and SSTs and sea ice as a group.

Step 1 (all variables): Calculation of long-term means and standard deviations on the native grids of the models. Where monthly time series data were available (most PMIP2 models and all CMIP5 models), we calculated long-term means and standard deviations from the last 100 years of the available simulations (taking care to synchronize the observations across variables, so that the means were based on identical time spans). In the case of data sets shorter than 100 years (PMIP2 CCSM and GISSmodelE, CMIP5 CCSM 4), the long-term means were based on the whole length of the time series archived. Only pre-computed long-term means ("SE" files in the PMIP2 data base) were available for some PMIP2 simulations (ECBILT.LV, ECHAM (OAV), HadCM3), and those were adopted without modification. Near-surface air temperature and SST were transformed to degrees Celsius (°C), and precipitation flux (in kg m^2 sec^{-1}) was converted to total monthly precipitation (mm $month^{-1}$), taking into account the variable number of days per month according to the specific calendar employed by each model (e.g. 360-day vs. 365-day), and the changes over time related to orbital variations.

Step 2 (all variables): Interpolation to a 0.5° grid. The long-term means and standard deviations on the model grids were interpolated to a 0.5° grid that followed the Climate Research Unit (CRU) CL 2.0 data set used here to provide "modern" high-resolution data. We used bilinear interpolation for temperature, precipitation and clouds, as implemented in the CDO package (https://code.zmaw.de/projects/cdo) (remapbil). For ocean grids that were based on rotated-pole or non-latitude-longitude grids, bilinear interpolation resulted in expansion of continental areas (consequently masking the near-shore location of many of the marine paleodata points), and so we adopted nearest-neighbor interpolation (remapnn) for SSTs and sea ice.

Step 3 (temperature, precipitation, clouds): Calculation of anomalies. Anomalies or long-term mean differences (experiment minus control) of temperature, precipitation and clouds were calculated in the standard way.

Step 4 (temperature, precipitation, clouds): Application of anomalies to modern CRU data. The anomalies calculated in Step 3 were applied to the CRU CL 2.0 1961-1990 long-term means to create "paleo" values that retain the spatial variation in the modern. Such variation is important for the proper calculation of bioclimatic variables and for the simulation of vegetation (Harrison et al. 1998). The CRU CL 2.0 data were first extrapolated on to the continental shelves, using the union of 0.5° grid points that are above sea level at present or would have been at 6 ka or at 21 ka, as implied by the ICE-5G paleotopography data set (Peltier 2004). Topography and bathymetry were derived from the ETOPO-1 data set. Generic lapse rates for each variable were used to adjust the extrapolated values to reflect elevational trends in the data.

Step 5: (temperature, precipitation, clouds): Calculation of bioclimatic variables. Bioclimatic variables were calculated using the anomalies from Step 4 on the 0.5° grid using the approach of Prentice et al. (1992). The original routines of Cramer and Prentice (1988) and Prentice et al. (1993) were modified to include snow-moisture accounting (where snowmelt is parameterized after Tarboton and Luce 1996) and adapted to use the IGBP-DIS multi-layer soil-properties data set (Global Soil Data Task 2000). The specific variables calculated were GDD5, MTWA, MTCO, MAT, MAP and α. The calculation of equilibrium evapotranspiration requires data on the percent of possible sunshine. These data were estimated from the cloudiness data using the Doorenbos and Pruitt (1984) approach. A separate regression equation was fit for each month of the year and for each 0.5° grid cell using the

cloudiness data in CRU CL 1.0 and the percent possible sun data in CLU CL 2.0 in a moving window centered on each grid point.

Step 6: (bioclimatic variables): Calculation of bioclimatic variable anomalies. The values of bioclimatic variables calculated in Step 5 were subtracted from the model-based values calculated in Step 4 to create a set of anomalies.

Step 7: (temperature, precipitation): Gridding of temperature and precipitation. The anomalies of tas and pr were re-gridded to a 2° latitude x longitude grid by simple averaging of the 0.5° data within each 2° cell.

Step 8: (bioclimatic variables): Gridding of bioclimatic variable anomalies. In order to facilitate comparisons with the terrestrial reconstructions, the anomalies of the bioclimatic variables calculated in the previous step were re-gridded to a 2° latitude x longitude grid by simple averaging of the 0.5° data within each 2° cell.

Step 9: (sea-surface temperature and sea-ice concentration): Calculation of sea-surface variables. The monthly values of SST and sea-ice concentration were used to calculate SSTann, SSTwin, and SSTsum, and SInmon (all defined in the same way as the seasonal reconstructions, see S1.2).

Step 10: (sea-surface summary variables): Calculation of sea-surface variable anomalies. Experiment minus control anomalies of the sea-surface summary variables were calculated in the usual way.

Step 11: *(sea-surface summary variables): Re-gridding of the sea-surface variables*. To facilitate comparisons with the ocean reconstructions, the anomalies calculated in Step 9 were re-gridded to a 2° latitude x 2° longitude grid by simple averaging of the 0.5° data within each 2° cell. (Note: the data were also re-gridded to 5° by 5° cells, but although available these data are not used in the current analyses.)

4. Metrics for comparison of reconstructed and simulated climate variables

4.1 Means, medians and interquartile ranges

Mean, median and the interquartile range (IQR) of each variable were calculated separately (a) for all the appropriate (land/ocean) 2˚ x 2˚ grid cells in each set of model results, (b) for just those grid cells corresponding for which there are observations. Calculation (b) was used as the basis for comparisons of models and data. The (signed) difference between the median values of a variable based on model results and observations is the 'median bias'. A perfect model would be expected to show a bias close to zero, within the limits of precision of the data, so models can be ranked for a given variable based on the absolute magnitude of their median bias. The *ratio* between the IQR's in the model results and observations provides a robust measure of the tendency of the model to under- or overestimate the degree of spatial variability in the observations. A perfect model in this respect would be expected to show an IQR ratio close to unity. Models can be ranked based on the absolute magnitude of the logarithm of their IQR ratio.

4.2 Kendall's Tau

Kendall's rank correlation coefficient tau (τ) was used as the basis for a dissimilarity metric expressing the extent of agreement in the rank order of anomalies among grid cells (here 2˚ x 2˚ cells). This measure does not assess bias in the modeled values (measured by the median bias) nor disagreement in their variability among grid cells (measured

by the ratio of interquartile ranges); it measures the similarity or difference of their spatial patterns without regard for magnitudes. τ is computed by considering all possible pairs of grid cells and asking whether they are concordant (i.e. one member of the pair is greater than the other in the modeled and observed sets of values) or discordant. It is defined (in the absence of ties) by:

$$\tau = (n_c - n_d)/[n(n-1)/2]$$

where the number of concordant pairs is n_c, the number of discordant pairs is n_d, and the denominator is the total number of pairs. A standard procedure exists for the handling of tied values. τ takes the value 1 for perfect agreement and –1 for perfect disagreement in the rankings. A τ of zero would mean no rank correlation. As a dissimilarity metric we therefore defined one minus tau, which takes the value zero for perfect agreement, one for 'no correlation', and values greater than one for anti-correlation. Here we consider values of <0.5 to indicate good agreement, values between 0.5 and 0.8 to indicate moderate agreement, values between 0.8 and 0.95 to indicate poor agreement, and values >0.95 to indicate no agreement.

4.3 Fuzzy distance

One challenge in comparing between data and models is how to include measures of uncertainty in either source of information. Guiot et al. (1999) proposed a distance measure based on fuzzy logic that explicitly allows for uncertainties in calculating similarity. The method requires that the variable under investigation is defined as a triangular fuzzy number for both model and data. Triangular fuzzy numbers are defined as $A(a_1,a_2,a_3)$, where a_1 defines the lower limit of the number, a_2 the point of maximum membership and a_3 the upper limit. We use this to compare each variable at each grid-cell for which we have both model and observed values.

Fuzzy limits for the data were defined using the reconstruction uncertainties. For each set of observed variables, we generated a probability density function of potential reconstruction values at each site, using a conditional Gaussian simulation. This uses the semi-variogram associated with each dataset to constrain the possible values obtained during any single iteration, reflecting the observed spatial dependency. Fuzzy number limits were then defined as the 2.5^{th} and 97.5^{th} percentile of this PDF. As no uncertainty range was available for the models (see above), the model value is defined using a crisp number, where $a_1=a_2=a_3$. We use a newer fuzzy distance measure, proposed by Tran and Duckstein (2002), which is more robust to large differences in the magnitude of the two fuzzy numbers (*A*, *B*) to be compared:

$$D^2(A,B) = (a_2 - b_2)^2 + \frac{1}{3}(a_2 - b_2)[(a_3 + a_1) - (b_3 + b_1)]$$
$$+ \frac{1}{18}[(a_3 - a_2)^2 + (a_2 - a_1)^2 + (b_3 - b_2)^2 + (b_2 - b_1)^2]$$
$$- \frac{1}{18}[(a_2 - a_1)(a_3 - a_2) + (b_2 - b_1)(b_3 - b_2)]$$
$$+ \frac{1}{12}(2a_2 - a_1 - a_3)(2b_2 - b_1 - b_3)$$

Once the distance measure D^2 has been calculated for all grid cells, we calculate the median, and use this as the metric for comparison. This distance metric synthesizes both the spatial pattern and magnitude of difference, and makes for a useful supplement to the measures of bias. For example, two models may each have a median bias of close to zero, but only the first has a spatial pattern similar to that of the data. In this case, the median distance for the second model would be larger than the first. We have chosen to calculate the fuzzy distance assuming that the relevant uncertainties are those associated with the reconstructions. In the absence of ensemble runs for each model (which would at least address the uncertainty due to internal variability) or perturbed-physics ensembles for a given model (which would address structural uncertainty), we cannot estimate individual-model uncertainty. We are therefore forced to assume that model uncertainty is represented by the differences among models.

5. Summaries of Metrics

Tables S7 and S8 summarize the numeric values of global benchmark metrics for the LGM and MH respectively for different model ensembles and Tables S9 and S10 summarize the global benchmark metrics for the LGM and MH respectively for individual models. These tables are the basis for the construction of Figure 7 and Figure 8 in the main text. We present ensemble results for all of the models (ALL), all of the PMIP2 models (PMIP2), all of the CMIP5 models (CMIP5), all of the PMIP2 ocean-atmosphere models (PMIP2 OA), all of the PMIP2 ocean-atmosphere-vegetation models (PMIP2 OAV), all of the CMIP5 ocean-atmosphere models (CMIP5 OA) and all of the CMIP5 ocean-atmosphere-carbon cycle models (CMIP5 OAC).

Figures S6 and S7 show comparisons of simulated and reconstructed zonal mean annual sea-surface temperature (SSTann) at the LGM and MH respectively.

Supplementary References

Affek HP, Bar-Matthews M, Ayalon A, Matthews A, Eiler J.M. (2008) Glacial/interglacial temperature variations in Soreq cave speleothems as recordes by “clumped isotope” thermometry. Geochim Cosmochim Ac 72**:**5351-5360

Arz HW, Pätzold J, Müller PJ, Moammar MO (2003) Influence of Northern Hemisphere climate and global sea level rise on the restricted Red Sea marine environment during termination I. Paleoceanography 18:053 doi:10.1029/2002PA000864

Asmeron Y, Polyak VJ, Burns SJ, Rassmussen J (2007) Solar forcing of Holocene climate: New insights from a speleothem record, southwestern United States. Geology 35:1-4

Asmeron Y, Polyak VJ, Burns SJ (2010) Variable winter moisture in the southwestern United States linked to rapid glacial climate shifts. Nature Geoscience 3:114-117.

Barker S, Diz P, Vautravers MJ, Pike J, Knorr G, Hall IR, Broecker WS (2009) Interhemispheric Atlantic seesaw response during the last deglaciation. Nature 457:1097-1050

Bar-Matthews M, Ayalon A, Kaufman A, Wasserburg GJ (1999) The Eastern Mediterranean paleoclimate as a reflection of regional events: Soreq cave, Israel. Earth Planet Sci Lett 166:85-95

Bar-Matthews M, Ayalon A, Gilmour M, Matthews A, Hawkesworth CJ (2003) Sea-land oxygen isotopic relationships from planktonic foraminifera and speleothems in the Eastern Mediterranean region and their implication for paleorainfall during interglacial intervals. Geochim Cosmochim Ac 67:3181-3199

Bar-Matthews M, Ayalon A (2010) Mid-Holocene climate variations revealed by high-resolution speleothem records from Soreq Cave, Israel and their correlation with cultural changes. Holocene 21:163-171

Barrows TT, Lehman SJ, Fifield LK, De Deckker P (2007) Absence of cooling in New Zealand and the adjacent ocean during the younger dryas chronozone. Science 318:86-89

Bartlein PJ, Webb III T, Fleri EC (1984) Holocene climatic change in the northern Midwest: Pollen-derived estimates. Quat Res 22:361-374

Bartlein PJ, Prentice IC, Webb III T (1986) Climatic response surfaces from pollen data for some eastern North American taxa. J Biogeog 13:35-57

Bartlein PJ, Harrison SP, Brewer S, Connor S, Davis BAS, Gajewski K, Guiot J, Harrison-Prentice TI, Henderson A, Peyron O, Prentice IC, Scholze M, Seppä H, Shuman B, Sugita S, Thompson RS, Viau AE, Williams J, Wu H (2011) Pollen-based continental climate reconstructions at 6 and 21 ka: a global synthesis. Clim Dyn 37:775-802

Benway HM, Mix AC, Haley BA, Klinkhammer GP (2006) Eastern Pacific Warm Pool paleosalinity and climate variability: 0-30 kyr. Paleoceanography doi:10.1029/2005PA001208

Bernal JP, Lachniet M, McCulloch M, Mortimer G, Morales P, Cienfuegos E (2011) A speleothem record of Holocene climate variability from southwestern Mexico. Quat Res 75:104-113

Braconnot P, Harrison SP, Kageyama M, Bartlein PJ, Masson-Delmotte V, Abe-Ouchi A, Otto-Bliesner B, Zhao Y (2012) Evaluation of climate models using palaeoclimatic data. Nature Climate Change. doi: 10.1038/NCLIMATE1456

Budich R, Giorgetta M, Jungclaus J, Redler R, Reick C (2010) The MPI-M Millennium Earth System Model: An Assembling Guide for the COSMOS Configuration. Available online. http://cosmos.enes.org/fileadmin/user_upload/cosmos/documents/Documentation/M-ESM-Assembling-Guide.pdf

Cai Y, Tan L, Cheng H, An Z, Edwards RL, Kelly MJ, Kong X, Wang X **(**2010) The variation of summer monsoon precipitation in central China since the last deglaciation. Earth Planet Sci Lett 291**:**21-31

Calvo E, Pelejero C, De Deckker P, Logan GA (2007) Antarctic deglacial pattern in a 30 kyr record of sea surface temperature offshore South Australia. Geophys Res Lett. doi: 10.1029/2007GL029937

Carlson AE, Oppo DW, Came RE, LeGrande AN, Keigwin LD, Curry WB (2008) Subtropical Atlantic salinity variability and Atlantic meridional circulation during the last deglaciation. Geology 36:991-994

Castañeda IS, Schefuß E, Pätzold J, Sinninghe Damsté JS, Weldeab S, Schouten S (2010) Millennial-scale sea surface temperature changes in the eastern Mediterranean (Nile River Delta region) over the last 27,000 years. Paleoceanography doi: 10.1029/2009PA001740

Cheng H, Zhang PZ, Spötl C, Edwards RL, Cai YJ, Zhang DZ, Sang WC, Tan W, An ZS (2012) The climatic cyclicity in semiarid-arid central Asia over the past 500,000 years. Geophys Res Lett 39**:**L01705. doi:10.1029/2011GL050202

Collins WJ, Bellouin N, Doutriaux-Boucher M, Gedney N, Halloran P, Hinton T, Hughes J, Jones CD, Joshi M, Liddicoat S, Martin G, O'Connor F, Rae J, Senior C, Sitch S, Totterdell I, Wiltshire A, Woodward S (2011) Development and evaluation of an Earth-system model HadGEM2. Geosci Model Dev 4:1051-1075

Constantin S, Bojar A-V, Lauritzen S-E, Lundberg J (2007) Holocene and Late Pleistocene climate in the sub-Mediterranean continental environment: A speleothem record from Poleva Cave (Southern Carpathians, Romania). Palaeogeogr, Palaeoclimatol, Palaeoecol 243:322-338

Cosford J, Qing H, Yuan D, Zhang M, Holmden C, Patterson W, Hai C (2008a) Millennial-scale variability in the Asian monsoon: Evidence from oxygen isotope records from stalagmites in southeastern China. Palaeogeogr, Palaeoclimatol, Palaeoecol 266:3-12

Cosford J, Qing H, Eglington B, Mattey D, Yuan D, Zhang M, Cheng H (2008b) East Asian monsoon variability since the Mid-Holocene recorded in a high-resolution, absolute-dated aragonite speleothem from eastern China. Earth Planet Sci Lett 275:296-307

Cosford J, Qing H, Mattey D, Eglington B, Zhang M (2009) Climatic and local effects on stalagmite $\delta^{13}C$ values at Lianhua Cave, China. Palaeogeogr, Palaeoclimatol, Palaeoecol 280:235-244

Cosford J, Qing H, Eglington B, Mattey D, Chen YG, Zhang M, Cheng H (2010) The East Asian monsoon during MIS 2 expressed in a speleothem $\delta^{18}O$ record from Jintanwan Cave, Huan, China. Quat Res 73:541-549

Cramer W, Prentice IC (1988) Simulation of regional soil moisture deficits on a European scale. Norsk Geograf Tids 42:149-151

Cruz FW, Burns SJ, Karmann I, Sharp WD, Vuille M, Cardoso AO, Ferrari JA, Dias PL, Viana O (2005) Insolation-driven changes in atmospheric circulation over the past 116,000 years in subtropical Brazil. Nature 434:63-66

Cruz FW, Burns SJ, Jercinovic M, Karmann I, Sharp WD, Vuille M (2007) Evidence of rainfall variations in Southern Brazil from trace element ratios (Mg/Ca and Sr/Ca) in a Late Pleistocene stalagmite. Geochim Cosmochim Ac 71:2250-2263

Dayem KE, Molnar P, Battisti DS, Roe GH (2010) Lessons learned from oxygen isotopes in modern precipitation applied to interpretation of speleothem records of paleoclimate from eastern Asia. Earth Planet Sci Lett 295:219-230

deMenocal P, Ortiz J, Guilderson T, Sarnthein M (2000) Coherent high- and low-latitude climate variability during the Holocene warm period. Science 288:2198-2202

Denniston RF, González, Asmerom Y, Baker RG, Reagan MK, Bettis III EA (1999) Evidence for increased cool season moisture during the middle Holocene. Geology 27:815-818

de Vernal A, Eynaud F, Henry M, Hillaire-Marcel C, Londeix L, Mangin S, Matthiessen J, Marret F, Radi T, Rochon A, Solignac S, Turon J-L (2005) Reconstruction of sea-surface conditions at middle to high latitudes of the Northern Hemisphere during the Last Glacial Maximum (LGM) based on dinoflagellate cyst assemblages. Quat Sci Rev 24:897-924

de Vries P, Weber SL (2005) The Atlantic freshwater budget as a diagnostic for the existence of a stable shut down of the meridional overturning circulation Geophys Res Lett 32:L09606. doi: 10.1029/2004GL021450

Dong J, Wang Y, Cheng H, Hardt B, Edwards RL, Kong X, Wu J, Chen S, Liu D, Jiang X, Zhao K (2010) A high-resolution stalagmite record of the Holocene East Asian monsoon from Mt Shennongjia, central China. Holocene 20:257-264

Doorenbos J, Pruitt WO (1984) Crop water requirements. FAO Irrigation and Drainage Paper No. 24, Food and Agriculture Organization of the United Nations, Rome, pp 144

Dorale JA, Edwards RL, Alexander CA, Shen C-C, Richards DA, Cheng H (2004) Uranium-series dating of speleothems: current techniques, limits and applications. In: Studies of Cave Sediments. Physical and Chemical Records of Palaeoclimate (eds Sasowsky ID, Mylroie J) Kluwer Academic, New York, pp 177-197

Dreybrodt W, Scholz D (2011) Climatic dependence of stable carbon and oxygen isotope signals recorded in speleothems: From soil water to speleothem calcite. Geochim Cosmochim Ac 75:734-752

Drysdale R, Zanchetta G, Hellstrom J, Maas R, Fallick A, Pickett M, Cartwright I, Piccini L (2006) Late Holocene drought responsible for the collapse of Old World civilizations is recorded in an Italian cave flowstone. Geology 34:101-104

Dykoski CA, Edwards RL, Cheng H, Yuan D, Cai Y, Zhang M, Lin Y, Qing J, An Z, Revenaugh J (2005) A high-resolution, absolute-dated Holocene and deglacial Asian monsoon record from Dongge Cave, China. Earth Planet Sci Lett 233:71-86

Essery RLH, Best MJ, Betts RA, Cox PM, Taylor CM (2003) Explicit representation of subgrid heterogeneity in a GCM land surface scheme. J Hydrometeor 4:530–543

Fairchild IJ, Smith CL, Baker A, Fuller L, Spötl C, Mattey D, McDermott F, EIMF (2006) Modification and preservation of environmental signals in speleothems. Earth Sci Rev 75:105-153

Fairchild IJ, Treble PC (2009) Trace elements in speleothems as recorders of environmental change. Quat Sci Rev 28:449-468

Farmer EC, deMenocal PB, Marchitto TM (2005) Holocene and deglacial ocean temperature variability in the Benguela upwelling region: Implications for low-latitude atmospheric circulation. Paleoceanography 20

Fleitmann D, Burns SJ, Neff U, Mangini A, Matter A (2003) Changing moisture sources over the last 330,000 years in Northern Oman from fluid-inclusion evidence in speleothems. Quat Res 60:223-232

Fleitmann D, Burns SJ, Mangini A, Mudelsee M, Kramers J, Villa I, Neff U, Al-Subbary AA, Buettner A, Hippler D, Matter A (2007) Holocene ITCZ and Indian monsoon dynamics recorded in stalagmites from Oman and Yemen (Socotra). Quat Sci Rev 26:170-188

Fleitmann D, Cheng H, Badertscher S, Edwards RL, Mudelsee M, Göktürk OM, Fankhauser A, Pickering R, Raible CC, Matter A, Kramers J, Tüysüz O (2009a) Timing and climatic impact of Greenland interstadials recorded in stalagmites from northern Turkey. Geophys Res Lett 36:L19797

Fleitmann D, Matter A (2009b) The speleothem record of climate variability in Southern Arabia. Geoscience 341:633-642

Flower BP, Hastings DW, Hill HW, Quinn TM (2004) Phasing of deglacial warming and Laurentide ice sheet meltwater in the Gulf of Mexico. Geology 32:597-600

Fowler C, Emery WJ, Maslanik J (2004) Satellite-derived evolution of Arctic sea ice age: October 1978 to March 2003. IEEE Geosci Remote Sens Letters 1:71-74

Frisia S, Borsato A, Spötl C, Villa IM, Cucchi F (2005) Climate variability in the SE Alps of Italy over the past 17 000 years reconstructed from a stalagmite record. Boreas 34:445-455

Frumkin A, Ford DC, Schwarcz HP (1999) Continental oxygen isotopic record of the last 170,000 years in Jerusalem. Quat Res 51:317-327

Gent PR, Danabasoglu G, Donner LJ, Holland MM, Hunke EC, Jayne SR, Lawrence DM, Neale RB, Rasch PJ, Vertenstein M, Worley PH, Yang Z-L, Zhang M (2011) The community climate system model version 4. J Clim 24:4973-4991

Genty D, Blamart D, Ghaleb B, Plagnes V, Causse Ch, Bakalowicz M, Zouari K, Chkir N, Hellstrom J, Wainer K, Bourges F (2006) Timing and dynamics of the last deglaciation from European and North African $\delta^{13}C$ stalagmite profiles - comparison with Chinese and South Hemisphere stalagmites. Quat Sci Rev 25:2118–2142

Gersonde R, Crosta X, Abelmann A, Armand L (2005) Sea-surface temperature and sea ice distribution of the Southern Ocean at the EPILOG Last Glacial Maximum: A circum-Antarctic view based on siliceous microfossil records. Quat Sci Rev 24:869-896

Global Soil Data Task (2000) Global Soil Data Products, CD-ROM. International Geosphere–Biosphere Programme Data and Information System, Potsdam, Germany. Available from Oak Ridge National Laboratory Distributed Active Archive Center (http://www.daac.ornl.gov)

Göktürk OM, Fleitmann D, Badertscher S, Cheng H, Edwards RL, Leuenberger M, Fankhauser A, Tüysüz O, Kramers J (2011) Climate on the southern Black Sea coast during the Holocene: implications from the Sofular Cave record. Quat Sci Rev 30:2433-2445

Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mitchell JFB, Wood RA (2000) The simulation of SST, sea ice extents and ocean heat transport in a version of the Hadley Centre coupled model without adjustments. Clim Dyn 16:147-168

Griffiths ML, Drysdale RN, Gagan MK, Zhao J-X, Ayliffe LK, Hellstrom JC, Hantoro WS, Frisia S, Feng Y-X, Cartwright I, Pierre ESt, Fischer MJ, Suwargadi BW (2009) Increasing Australian-Indonesian monsoon rainfall linked to early Holocene sea-level rise. Nature Geoscience 2:636-639

Griffiths ML, Drysdale RN, Vonhof HB, Gagan MK, Zhao J-X, Ayliffe LK, Hantoro WS, Hellstrom JC, Cartwright I, Frisia S, Suwargadi BW (2010a) Younger Dryas-Holocene temperature and rainfall history of southern Indonesia from $\delta^{18}O$ in speleothem calcite and fluid inclusions. Earth Planet Sci Lett 295:30-36

Griffiths ML, Drysdale RN, Gagan MK, Frisia S, Zhao JX, Ayliffe LK, Hantoro WS, Hellstrom JC, Fischer MJ, Feng YX, Suwargadi BW (2010b) Evidence for Holocene changes in Australian-Indonesian monsoon rainfall from stalagmite trace element and stable isotope ratios. Earth Planet Sci Lett 292:27-38

Guiot J, Boreux JJ, Braconnot P, Torre F (1999) Data-model comparisons using fuzzy logic in palaeoclimatology. Clim Dyn 15:569–581

Hardt B, Rowe HD, Springer GS, Cheng H, Edwards RL (2010) The seasonality of east central North American precipitation based on three coeval Holocene speleothems from southern West Virginia. Earth Planet Sci Lett 295:342-348

Harrison SP, Jolly D, Laarif F, Abe-Ouchi A, Dong B, Herterich K, Hewitt C, Joussaume S, Kutzbach JE, Mitchell J, de Noblet N, Valdes P (1998) Intercomparison of simulated global vegetation distribution in response to 6 kyr BP orbital forcing. J Clim 11:2721-2742

Hazeleger W, Wang X, Severijns C, Stefanescu S, Bintanja R, Sterl A, Wyser K, Semmler T, Yang S, van den Hurk B, van Noije T, van der Linden E, van der Wiel K (2012) EC-Earth V2.2: description and validation of a new seamless Earth system prediction model. Clim Dyn 39:2611-2629, doi:10.1007/s00382-011-1228-5.

Hellstrom J, McCulloch M, Stone J (1998) A detailed 31,000-year record of climate and vegetation change, from the isotope geochemistry of two New Zealand speleothems. Quat Res 50:167-178

Hellstrom J (2006) U-Th dating of speleothems with high initial 230Th using stratigraphical constraint. Quat Geochr 1:289-295

Holmgren K, Lee-Thorp JA, Cooper GJ, Lundblad K, Partridge TC, Scott L, Sithaldeen R, Talma AS, Tyson PD (2003) Persistent millennial-scale climatic variability over the past 25 thousand years in southern Africa. Quat Sci Rev 22:2311-2326

Hu C, Henderson GM, Huang J, Xie S, Sun Y, Johnson KR (2008) Quantification of Holocene Asian monsoon rainfall from spatially separated cave records. Earth Planet Sci Lett 266:221-232

Huguet C, Kim J-H, Damsté JSS, Schouten S (2006) Reconstruction of sea surface temperature variations in the Arabian Sea over the last 23 kyr using organic proxies (TEX(86) and U(37)(K')). Paleoceanography 21:PA3003

Huntley B, Prentice IC (1988) July temperatures in Europe from pollen data, 6000 years before present. Science 241:687-690

Ijiri A, Wang L, Oba T, Kawahata H, Huang CH-Y, Huang C-Y (2005) Paleoenvironmental changes in the northern area of the East China Sea during the past 42,000 years. Palaeogeogr, Palaeoclimatol, Palaeoecol 219:239-261

Isono D, Yamamoto M, Irino T, Oba T, Murayama M, Nakamura T, Kawahata H (2009) The 1500-year climate oscillation in the midlatitude North Pacific during the Holocene. Geology 37:591-594

Jacob R, Schafer C, Foster I, Tobis M, Anderson J (2001) Computational design and performance of the Fast Ocean Atmosphere Model: Version 1. The 2001 International Conference on Computational Science, pp 175–184

Jaeschke A, Rühlemann C, Arz H, Heil G, Lohmann G (2007) Coupling of millennial-scale changes in sea surface temperature and precipitation off northeastern Brazil with high-latitude climate shifts during the last glacial period. Paleoceanography 22

Johns TC, Royer J-F, Höschel I, Heubener H, Roeckner E, Manzini E, May W, Dufresne J-L, Otterå OH, van Vuuren DP, Salas y Melia D, Giorgetta MA, Denvil S, Yang S, Fogli PG, Körper J, Tjiputra JF, Stehfest E, Hewitt CD (2011) Climate change under aggressive mitigation: The ENSEMBLES multi-model experiment. Clim Dyn 37:1975-2003

Jones CD, Hughes JK, Bellouin N, Hardiman SC, Jones GS, Knight J, Liddicoat S, O'Connor FM, Andres RJ, Bell C, Boo K-O, Bozzo A, Butchart N, Cadule P, Corbin KD, Doutriaux-Boucher M, Friedlingstein P, Gornall J, Gray L, Halloran PR, Hurtt G, Ingram WJ, Lamarque J-F, Law RM, Meinshausen M, Osprey S, Palin EJ, Chini LP, Raddatz T, Sanderson MG, Sellar AA, Schurer A, Valdes P, Wood N, Woodward S, Yoshioka M, Zerroukat M (2011) The HadGEM2-ES implementation of CMIP5 centennial simulations. Geosci Model Dev 4:543-570

Jungclaus JH, Keenlyside N, Botzet M, Haak H, Luo J-J, Latif M, Marotzke J, Mikolajewicz U, Roeckner E (2006) Ocean circulation and tropical variability in the coupled model ECHAM5/MPI-OM. J Clim 19:3952–3972

K-1-Model-Developers (2004) K-1 Coupled GCM (Miroc Description)1, pp 34

Kaiser J, Schefuß E, Lamy F, Mohtadi M, Hebbeln D (2008) Glacial to Holocene changes in sea surface temperature and coastal vegetation in north central Chile: high versus low latitude forcing. Quat Sci Rev 27:2064-2075

Kanner LC, Burns SJ, Cheng H, Edwards RL (2012) High-latitude forcing of the South American summer monsoon during the last glacial. Science 335:570-573

Kaufman A, Wasserburg GJ, Porcelli D, Bar-Matthews M, Ayalon A, Halicz L (1998) U-Th systematics from the Soreq cave, Israel and climatic correlations. Earth Planet Sci Lett 156:141-155

Kim JH, Rimbu N, Lorenz SJ, Lohmann G, Nam S-I, Schouten S, Rühlemann C, Schneider RR (2004) North Pacific and North Atlantic sea-surface temperature variability during the Holocene. Quat Sci Rev 23:2141-2154

Kienast M, Kienast SS, Calvert SE, Eglinton TI, Mollenhauser G, François R, Mix AC (2006) Eastern Pacific cooling and Atlantic overturning circulation during the last deglaciation. Nature 443:846-849

Koutavas A, Sachs JP (2008) Northern timing of deglaciation in the eastern equatorial Pacific from alkenone paleothermometry. Paleoceanography 23

Kucera M, Weinelt M, Kiefer T, Pflaumann U, Hayes A, Weinelt M, Chen M-T, Mix AC, Barrows TT, Cortijo E, Duprat J, Juggins S, Waelbroeck C (2005) Reconstruction of sea-surface temperatures from assemblages of

576 planktonic foraminifera: multi-technique approach based on geographically constrained calibration datasets
577 and its application to glacial Atlantic and Pacific Oceans. Quat Sci Rev 24:951-998
578 Lachniet MS, Asmerom Y, Burns SJ, Patterson WP, Polyak VJ, Seltzer GO (2004) Tropical response to the 8200 yr
579 B.P. cold event? Speleothem isotopes indicate a weakened early Holocene monsoon in Costa Rica. Geology
580 32:957-960
581 Lachniet MS (2009) Climatic and environmental controls on speleothem oxygen-isotope values. Quat Sci Rev
582 28:412-432
583 Lamy F, Kaiser J, Arz HW, Hebbeln D, Ninnemann U, Timm O, Timmermann A, Toggweiler JR (2007)
584 Modulation of the bipolar seesaw in the southeast Pacific during Termination 1. Earth Planet Sci Lett
585 259:400-413
586 Lauritzen SE, Lundberg J (1999) Calibration of the speleothem delta function: an absolute temperature record for
587 the Holocene in northern Norway. Holocene 9:659-669
588 Leduc G, Vidal L, Tachikawa K, Rostek F, Sonzogni C, Beaufort L, Bard E (2007) Moisture transport across
589 Central America as a positive feedback on abrupt climatic changes. Nature 445:908-911
590 Leduc G, Schneider R, Kim J-H, Lohmann G (2010) Holocene and Eemian sea surface temperature trends as
591 revealed by alkenone and Mg/Ca paleothermometry. Quat Sci Rev 29:989-1004
592 Lee-Thorp JA, Holmgren K, Lauritzen S-E, Linge H, Moberg A, Partridge TC, Stevenson C, Tyson PD (2001)
593 Rapid climate shifts in the southern African interior throughout the mid to late Holocene. Geophys Res Lett
594 28:4507-4510
595 Levi C, Labeyrie L, Bassinot F, Guichard F, Cortijo E, Waelbroeck C, Caillon N, Duprat J, de Garidel-Thoron T,
596 Elderfield H (2007) Low-latitude hydrological cycle and rapid climate changes during the last deglaciation.
597 Geochem Geophy Geosy 8
598 Linge H, Lauritzen S-E, Anderson C, Hansen JK, Skoglund RØ, Sundqvist HS (2009) Stable isotope records for the
599 last 10 000 years from Okshola cave (Fauske, northern Norway) and regional comparisons. Clim Past
600 5:667-682
601 Luetscher M, Hoffmann DL, Frisia S, Spötl C (2011) Holocene glacier history from alpine speleothems, Milchbach
602 cave, Switzerland. Earth Planet. Sci Lett 302:95-106
603 MARGO Project Members (2009) Constraints on the magnitude and patterns of ocean cooling at the Last Glacial
604 Maximum. Nat Geosci 2:127-132
605 Marsland SJ, Haak H, Jungclaus JH, Latif M, Röske F (2003) The Max-Planck-Institute global ocean/sea ice model
606 with orthogonal curvelinear coordinates. Ocean Model 5:91-127
607 Marti O, Braconnot P, Bellier J, Benshila R, Bony S, Brockmann P, Cadule P, Caubel A, Denvil S, Dufresne JL,
608 Fairhead L, Filiberti M-A, Fichefet T, Foujois M-A, Friedlingstein P, Grandpeix J-Y, Hourdin F, Krinner
609 G, Lévy C, Madec G, Musat I, De Noblet N, Polcher J, Talandier C (2005) The new IPSL climate system
610 model: IPSL-CM4. Tech Rep, Inst Pierre Simon Laplace, Paris. http://igcmg.ipsl.jussieu.fr/Doc/IPSLCM4
611 Marti O, Braconnot P, Dufresne J-F, Bellier J, Benshila R, Bony S, Brockmann P, Cadule P, Caudel A, Codron F, de
612 Noblet N, Denvil S, Fairhead L, Fichefet T, Foujols M-A, Friedlingstein P, Goosse H, Grandpeix J-Y,

Guilyardi E, Hourdin F, Idelkadi A, Kageyama M, Krinner G, Lévy C, Madec G, Mignot J, Musat I, Swingedouw D, Talandier C (2010) Key features of the IPSL ocean atmosphere model and its sensitivity to atmospheric resolution. Clim Dyn 35:1-26. doi: 10.1007/s00382-009-0640-6

Martin GM, Bellouin N, Collins WJ, Culverwell ID, Halloran PR, Hardiman SC, Hinton TJ, Jones CD, McDonald RE, McLaren AJ, O'Connor FM, Roberts MJ, Rodriguez JM, Woodward S, Best MJ, Brooks ME, Brown AR, Butchart N, Dearden C, Derbyshire SH, Dharssi I, Doutriaux-Boucher M, Edwards JM, Falloon PD, Gedney N, Gray LJ, Hewitt HT, Hobson M, Huddleston MR, Hughes J, Ineson S, Ingram WJ, James PM, Johns TC, Johnson CE, Jones A, Jones CP, Joshi MM, Keen AB, Liddicoat S, Lock AP, Maidens AV, Manners JC, Milton SF, Rae JGL, Ridley JK, Sellar A, Senior CA, Totterdell IJ, Verhoef A, Vidale PL, Wiltshire A (2011) HadGEM2 family of Met Office unified model climate configurations. Geosci Model Dev Discuss 4:765–841

Maslanik M, Stroeve J, Fowler C, Emery W (2011) Distribution and trends in Arctic sea ice age through spring 2011. Geophys Res Lett 38, L13502, doi:10.1029/2011GL047735

McDermott F, Frisia S, Huang Y, Longinelli A, Spiro B, Heaton THE, Hawkesworth CJ, Borsato A, Keppens E, Fairchild IJ, van der Berg K, Verheyden S, Selmo E (1999) Holocene climate variability in Europe: Evidence from $\delta^{18}O$, textural and extension-rate variations in three speleothems. Quat Sci Rev 18:1021-1038

McDermott F, Mattey DP, Hawkesworth CJ (2001) Centennial-scale Holocene climate variability revealed by a high-resolution speleothem $\delta^{18}O$ record from SW Ireland. Sci 294:1328-1331

McGarry S, Bar-Matthews M, Matthews A, Vaks A, Schilman B, Ayalon A (2004) Constraints on hydrological and paleotemperature variations in the Eastern Mediterranean region in the last 140 ka given by the δD values of speleothem fluid inclusions. Quat Sci Rev 23:919-934

Mikolajewicz U, Vizcaíno M, Jungclaus J, Schurgers G (2007) Effect of ice sheet inter- actions in anthropogenic climate change simulations. Geophys Res Lett 34:L18706

Minoshima K, Kawahata H, Ikehara K (2007) Changes in biological production in the mixed water region (MWR) of the northwestern North Pacific during the last 27 kyr. Palaeogeogr, Palaeoclimatol, Palaeoecol 254:430-447

Moreno A, Stoll H, Jiménez-Sánchez M, Cacho I, Valero-Garcés B, Ito E, Edwards RL (2010) A speleothem record of glacial (25-11.6 kyr BP) rapid climatic changes from northern Iberian Peninsula. Glob Planet Chang 71:218-231

Niggemann S, Mangini A, Richter DK, Wurth G (2003) A paleoclimate record of the last 17,600 years in stalagmites from the B7 cave, Sauerland, Germany. Quat Sci Rev 22:555-567

Notaro M, Liu Z, Gallimore R, Vavrus SJ, Kutzbach JE (2005) Simulated and observed preindustrial to modern vegetation and climate changes. J Clim 18:3650–3671

Onac BP, Constantin S, Lundberg J, Lauritzen SE (2002) Isotopic climate record in a Holocene stalagmite from Ursilor Cave, Romania. J Quat Sci 17:319-327

649 Otto-Bliesner BL, Brady EC, Clauzet G, Tomas R, Levis S, Kothavala Z (2006) Last Glacial Maximum and
650 Holocene Climate in CCSM3. J Clim 19:2526–2544
651 Overpeck JT, Webb III T, Prentice IC (1985) Quantitative interpretation of fossil pollen spectra: dissimilarity
652 coefficients and the method of modern analogs. Quat Res 23:87-108
653 Pahnke K, Sachs JP (2006) Sea surface temperatures of southern midlatitudes 0-160 kyr BP. Paleoceanogr 21.
654 doi: 10.1029/2005PA001191
655 Partin JW, Cobb KM, Adkins JF, Clark B, Fernandez DP (2007) Millennial-scale trends in west Pacific warm pool
656 hydrology since the Last Glacial Maximum. Nat 449:452-455
657 Pausata FSR, Battisti DS, Nisancioglu KH, Bitz CM (2011) Chinese stalagmite $\delta^{18}O$ controlled by changes in the
658 Indian monsoon during a simulated Heinrich event. Nat Geosci 4:474-480
659 Peck VL, Hall IR, Zahn R, Elderfield H (2008) Millennial-scale surface and subsurface paleothermometry from the
660 northeast Atlantic, 55-8 ka BP. Paleoceanogr 23 doi: 10.1029/2008PA001631
661 Peltier WR (2004) Global glacial isostasy and the surface of the ice-age Earth: the ICE-5G (VM2) model and
662 GRACE. Annu Rev Earth Planet Sci 32:111
663 Peterse F, Prins MA, Beets CJ, Troelstra SR, Zheng H, Gu Z, Schouten S, Sinninghe Damsté JS (2011) Decoupled
664 warming and monsoon precipitation in East Asia over the last deglaciation. Earth Planet Sci Lett 301:256-
665 264
666 Phipps SJ (2006) The CSIRO Mk3L Climate System Model, Technical Report No. 3, Antarctic Climate &
667 Ecosystems CRC, Hobart, Tasmania, Australia, pp 236
668 Phipps SJ, Rotstayn LD, Gordon HB, Roberts JL, Hirst AC, Budd WF (2011) The CSIRO Mk3L climate system
669 model version 1.0 - Part 1: Description and evaluation. Geosci Model Dev 4:483-509
670 Prentice IC, Cramer W, Harrison SP, Leemans R, Monserud RA, Solomon AM (1992) A global biome model based
671 on plant physiology and dominance, soil properties and climate. J Biogeogr 19:117-13
672 Prentice IC, Sykes MT, Cramer W (1993) A simulation model for the transient effects of climate change on forest
673 landscapes. Ecol Model 65:51-70
674 Quigley MC, Horton T, Hellstrom JC, Cupper ML, Sandiford M (2010) Holocene climate change in arid Australia
675 from speleothem and alluvial records. Holocene 20:1093-1104
676 Railsback LB, Liang F, Romaní JRV, Grandal-d'Anglade A, Rodríguez MV, Fidalgo LS, Mosquera DF, Cheng H,
677 Edwards RL (2011) Petrographic and isotopic evidence for Holocene long-term climate change and shorter-
678 term environmental shifts from a stalagmite from the Serra do Courel of northwestern Spain, and
679 implications for climatic history across Europe and the Mediterranean. Palaeogeogr, Palaeoclimatol,
680 Palaeoecol 305:172–184
681 Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A (2003) Global
682 analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth
683 century. J Geophys Res 108:D14, 4407

Renssen H, Goosse H, Fichefet T, Brovkin V, Driesschaert E, Wolk F (2005) Simulating the Holocene climate evolution at northern high latitudes using a coupled atmosphere-sea ice-ocean-vegetation model. Clim Dyn 24:23–43

Roeckner E, Bäuml G, Bonaventura L, Brokopf R, Esch M, Giorgetta M, Hagemann S, Kirchner I, Kornblueh L, Manzini E, Rhodin A, Schlese U, Schulzweida U, Tompkins A (2003) The atmospheric general circulation model ECHAM5. MPI Report No. 349

Rotstayn L, Collier MA, Dix MR, Feng Y, Gordon HB, O'Farrell SP, Smith IN, Syktus J (2010) Improved simulation of Australian climate and ENSO-related climate variability in a GCM with an interactive aerosol treatment. Int J Clim 30:1067-1088

Royer JF (2008) ENSEMBLES STREAM2 CNRM-CM33 20C3M run1, daily values. World Data Center for Climate. CERA-DB "ENSEMBLES2_CNCM33_20C3M_1_D" http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=ENSEMBLES2_CNCM33_20C3M_1_D

Sachs JP, Anderson RF, Lehman SJ (2001) Glacial surface temperatures of the southeast Atlantic Ocean. Sci 293:2077-2079

Sawada K, Handa N (1998) Variability of the path of the Kuroshio ocean current over the past 25,000 years. Nat 392:592-595

Schefuss E, Schouten S, Schneider RR (2005) Climatic controls on central African hydrology during the past 20,000 years. Nat 437:1003-1006

Schmidt MW, Spero HJ, Lea DW (2004) Links between salinity variation in the Caribbean and North Atlantic thermohaline circulation. Nat 428:160-163

Schmidt GA, Ruedy R, Hansen JE, Aleinov I, Bell N, Bauer M, Bauer S, Cairns B, Canuto V, Cheng Y, Del Genio A, Faluvegi G, Friend AD, Hall TM, Hu Y, Kelley M, Kiang NY, Koch D, Lacis AA, Lerner J, Lo KK, Miller RL, Nazarenko L, Oinas V, Perlwitz J, Perlwitz J, Rind D, Romanou A, Russell GL, Sato M, Shindell DT, Stone PH, Sun S, Tausnev N, Thresher D, Yao M-S (2006) Present day atmospheric simulations using GISS ModelE: Comparison to in-situ, satellite and reanalysis data. J Clim 19:153-192

Schmittner A, Urban NM, Shakun JD, Mahowald NM, Clark PU, Bartlein PJ, Mix AC, Rosell-Melé A (2011) Climate sensitivity estimated from temperature reconstructions of the Last Glacial Maximum. Sci 334:1385-1388

Scholz D, Frisia S, Borsato A, Spötl C, Fohlmeister J, Mudelsee M, Miorandi R, Mangini A (2012) Holocene climate variability in North-Eastern Italy: potential influence of the NAO and solar activity recorded by speleothem data. Clim Past Discuss 8:909–952

Shakun JD, Burns SJ, Fleitmann D, Kramers J, Matter A, Al-Subary A (2007) A high-resolution, absolute-dated deglacial speleothem record of Indian Ocean climate from Socotra Island, Yemen. Earth Planet Sci Lett 259:442-456

Southon J, Noronha AL, Cheng H, Edwards LR, Wang Y (2012) A high-resolution record of atmospheric ^{14}C based on Hulu Cave speleothem H82. Quat Sci Rev 33:32-41

720 Springer GS, Rowe HD, Hardt B, Edwards RL, Cheng H (2008) Solar forcing of Holocene droughts in a stalagmite
721 record from West Virginia in east-central North America. Geophys Res Lett 35:L17703
722 Steinke S, Kienast M, Groeneveld J, Lin L-C, Chen M-T, Rendle-Bühring R (2008) Proxy dependence of the
723 temporal pattern of deglacial warming in the tropical South China Sea: toward resolving seasonality. Quat
724 Sci Rev 27:688-700
725 Stenni B, Masson-Delmotte V, Selmo E, Oerter H, Meyer H, Röthlisberger R, Jouzel J, Cattani O, Falourd S,
726 Fischer H, Hoffmann G, Lacumin P, Johnsen SJ, Minster B, Udisti R (2010) The deuterium excess records
727 of EPICA Dome C and Dronning Maud Land ice cores (East Antarctica). Quat Sci Rev 29:146-159
728 Stott L, Timmermann A, Thunell R (2007) Southern hemisphere and deep-sea warming led deglacial atmospheric
729 CO_2 rise and tropical warming. Sci 318:435-438
730 Stríkis NM, Cruz FW, Cheng H, Karmann I, Edwards RL, Vuille M, Wang X, de Paula MS, Novello VF, Auler AS
731 (2011) Abrupt variations in South American monsoon rainfall during the Holocene based on a speleothem
732 record from central-eastern Brazil. Geol 39:1075-1078
733 Talma AS, Vogel JC (1992) Late Quaternary paleotemperatures derived from a speleothem from Cango Caves,
734 Cape Province, South-Africa. Quat Res 37:203-213
735 Tămaş, T, Onac BP, Bojar AV (2005) Lateglacial-Middle Holocene stable isotope records in two coeval stalagmites
736 from the Bihor Mountains, NW Romania. Geol Quart 49:185-194
737 Tarboton DG, Luce CH (1996) Utah Energy Balance Snow Accumulation and Melt Model (UEB): Computer model
738 technical description and users guide. Utah Water Research Laboratory, Utah State University.
739 Tierney JE, Russell JM, Huang Y, Sinninghe Damsté JS, Hopmans EC, Cohen AS (2008) Northern hemisphere
740 controls on tropical southeast African climate during the past 60,000 years. Sci 322:252-255
741 Tran L, Duckstein L (2002) Comparison of fuzzy numbers using a fuzzy distance measure. Fuzzy Sets Syst
742 130:331-341
743 Tremaine DM, Froelich PN, Wang Y (2011) Speleothem calcite farmed in situ: Modern calibration of $\delta^{18}O$ and $\delta^{13}C$
744 paleoclimate proxies in a continuously-monitored natural cave system. Geochim Cosmochim Ac 75:4929-
745 4950
746 Vaks A, Bar-Matthews M, Ayalon A, Matthews A, Frumkin A, Dayan U, Halicz L, Almogi-Labin A, Schilman B
747 (2006) Paleoclimate and location of the border between Mediterranean climate region and the Saharo-
748 Arabian Desert as revealed by speleothems from the northern Negev Desert, Israel. Earth Planet Sci Lett
749 249:384-399
750 van Beynen PE, Schwarcz HP, Ford DC (2004) Holocene climatic variation recorded in a speleothem from McFail's
751 cave, New York. Cave and Karst Stud 66:20-27
752 van Breukelen MR, Vonhof HB, Hellstrom JC, Wester WCG, Kroon D (2008) Fossil dripwater in stalagmites
753 reveals Holocene temperature and rainfall variation in Amazonia. Earth Planet Sci Lett 275:54-60
754 Verheyden S, Nader FH, Cheng HJ, Edwards LR, Swennen R (2008) Paleoclimatic reconstruction in the Levant
755 region from the geochemistry of a Holocene stalagmite from the Jeita cave, Lebanon. Quat Res 70:368-381

Voldoire A, Sanchez-Gomez E, Salas y Mélia D, Decharme B, Cassou C, Sénési S, Valcke S, Beau I, Alias A, Chevallier M, Déqué M, Deshayes J, Douville H, Fernandez E, Madec G, Maisonnave E, Moine M-P, Planton MS, Saint-Martin D, Szopa S, Tyteca S, Alkama R, Belamari S, Braun A, Coquart L, Chauvin F (2012) The CNRM-CM5.1 global climate model: description and basic evaluation. Clim Dyn. doi:10.1007/s00382-011-1259-y

Vollweiler N, Scholz D, Mühlinghaus C, Mangini A, Spötl C (2006) A precisely dated climate record for the last 9 kyr from three high alpine stalagmites, Spannagel Cave, Austria. Geophys Res Lett 33:L20703. doi:10.1029/2006GL027662

Waelbroeck C, Duplessy JC, Michel E, Labeyrie L, Paillard D, Duprat J (2001) The timing of the last deglaciation in North Atlantic climate records. Nat 412:724-727

Wagner JDM, Cole JE, Beck JW, Patchett PJ, Henderson GM, Barnett HR (2010) Moisture variability in the southwestern United States linked to abrupt glacial climate change. Nat Geosci 3:110-113

Wang X, Auler AS, Edwards RL, Cheng H, Ito E, Wang Y, Kong X, Solheid M (2007) Millennial-scale precipitation changes in southern Brazil over the past 90,000 years. Geophys Res Lett 34:L23701

Wang YJ, Cheng H, Edward RL, An ZS, Wu JY, Shen C-C, Dorale JA (2001) A high-resolution absolute-dated late Pleistocene monsoon record from Hulu Cave, China. Sci 294:2345-2348

Wang Y, Cheng H, Edwards RL, He Y, Kong X, An Z, Wu J, Kelly MJ, Dykoski CA, Li X (2005) The Holocene Asian monsoon: Links to solar changes and North Atlantic climate. Sci 308:854-857

Wang Y, Cheng H, Edwards RL, Kong X, Shao X, Chen S, Wu J, Jiang X, Wang X, An Z (2008) Millennial- and orbital-scale changes in the East Asian monsoon over the past 224,000 years. Nat 451:1090-1093

Watanabe S, Hajima T, Sudo K, Nagashima T, Takemura T, Okajima H, Nozawa T, Kawase H, Abe M, Yokohata T, Ise T, Sato H, Kato E, Takata K, Emori S, Kawamiya M (2011) MIROC-ESM: model description and basic results of CMIP5-20c3m experiments. Geosci Mod Dev 4:845-872

Webb T III, Bryson RA (1972) Late- and postglacial climatic change in the northern Midwest, U.S.A.: quantitative estimates derived from fossil pollen spectra by multivariate statistical analysis. Quaternary Research 2**:**70-115

Wei GJ, Deng WF, Liu Y, Li XH (2007) High-resolution sea surface temperature records derived from foraminiferal Mg/Ca ratios during the last 260 ka in the northern South China Sea. Palaeogeogr, Palaeoclimatol, Palaeoecol 250:126-138

Weldeab S, Schneider RR, Kölling M, Wefer G (2005) Holocene African droughts relate to eastern equatorial Atlantic cooling. Geol 33:981-984

Weldeab S, Schneider RR, Kölling M (2006) Deglacial sea surface temperature and salinity increase in the western tropical Atlantic in synchrony with high latitude climate instabilities. Earth Planet Sci Lett 241:699-706

Weldeab S, Lea DW, Schneider RR, Andersen N (2007) 155,000 years of West African monsoon and ocean thermal evolution. Sci 316:1303-1307

Whittaker TE, Hendy CH, Hellstrom JC (2011) Abrupt millennial-scale changes in intensity of Southern Hemisphere westerly winds during marine isotope stages 2-4. Geol 39:455-458

Williams PW, King DNT, Zhao J-X, Collerson KD (2004) Speleothem master chronologies: combined Holocene ^{18}O and ^{13}C records from the North Island of New Zealand and their palaeoenvironmental interpretation. Holocene 14:194-208

Williams PW, King DNT, Zhao J-X, Collerson KD (2005) Late Pleistocene to Holocene composite speleothem ^{18}O and ^{13}C chronologies from South Island, New Zealand - did a global Younger Dryas really exist? Earth Planet Sci Lett 230:301-317

Wu TW, Yu RC, Zhang F (2008) A modified dynamic framework for atmospheric spectral model and its application. J Atmos Sci 65:2235−2253

Wu TW, Yu R, Zhang F, Wang Z, Dong M, Wang L, Jin X, Chen D, Li L (2010) The Beijing Climate Center atmospheric general circulation model: description and its performance for the present-day climate. Clim Dyn 34:123−147

Wurth G, Niggemann S, Richter DK, Mangini A (2004) The younger Dryas and Holocene climate record of a stalagmite from Hölloch Cave (Bavarian Alps, Germany). J Quat Sci 19:291-298

Xia Q, Zhao J-X, Collerson KD (2001) Early-Mid Holocene climatic variations in Tasmania, Australia: multi-proxy records in a stalagmite from Lynds Cave. Earth Planet Sci Lett 194:177-187

Xu J, Holbourn A, Kuhnt W, Jian Z, Kawamura H (2008) Changes in the thermocline structure of the Indonesian outflow during Terminations I and II. Earth Planet Sci Lett 273:152-162

Yu YQ, Yu RC, Zhang XH, Liu HL (2002) A flexible coupled ocean-atmosphere general circulation model. Adv Atmos Sci 19:169–190

Yu YQ, Zhang XH, Guo YF (2004) Global coupled ocean-atmosphere general circulation models in Lasg/Iap. Adv Atmos Sci 21:444–455

Yuan D, Cheng H, Edwards RL, Dykoski CA, Kelly MJ, Zhang M, Qing J, Lin Y, Wang Y, Wu J, Dorale JA, An Z, Cai Y (2004) Timing, duration, and transitions of the last interglacial Asian monsoon. Sci 304:575-578

Yukimoto S, Noda A, Kitoh A, Hosaka M, Yoshimura H, Uchiyama T, Shibata K, Arakawa O, Kusunoki S (2006) Present-day climate and climate sensitivity in the Meteorological Research Institute Coupled GCM Version 2.3 (Mri-Cgcm2.3). J Meteorol Soc Japan 84:333–363

Yukimoto S, Yoshimura H, Hosaka M, Sakami T, Tsujino H, Hirabara M, Tanaka TY, Deushi M, Obata A, Nakano H, Adachi Y, Shindo E, Yabu S, Ose T, Kitoh A (2011) Meteorological Research Institute-Earth System Model v1 (MRI-ESM1) – Model Description. Tech Rep Meteor Res Inst 64, pp 88

Zanchetta G, Drysdale RN, Hellstrom JC, Fallick AE, Isola I, Gagan MK, Pareschi MT (2007) Enhanced rainfall in the Western Mediterranean during the deposition of sapropel S1: stalagmite evidence from Corchia cave (Central Italy). Quat Sci Rev 26:279-286

Ziegler M, Nürnberg D, Karas C, Tiedemann R, Lourens LJ (2008) Persistent summer expansion of the Atlantic Warm Pool during glacial abrupt cold events. Nat Geosci 1:601-605

Figure S1. Maps showing the location of speleothem records which met the selection criteria for reliability and are therefore used to provide either qualitative or quantitative estimates of temperature and/or precipitation for the LGM and MH. Qualitative and quantitative estimates of changes in (**A**) LGM (21ka) temperature, (**B**) LGM (21 ka) precipitation, (**C**) MH (6 ka) temperature and (**D**) MH (6 ka) precipitation, have been combined into a three part scale: cooler/no (significant) change/warmer for temperature, wetter/no (significant) change/drier for precipitation.

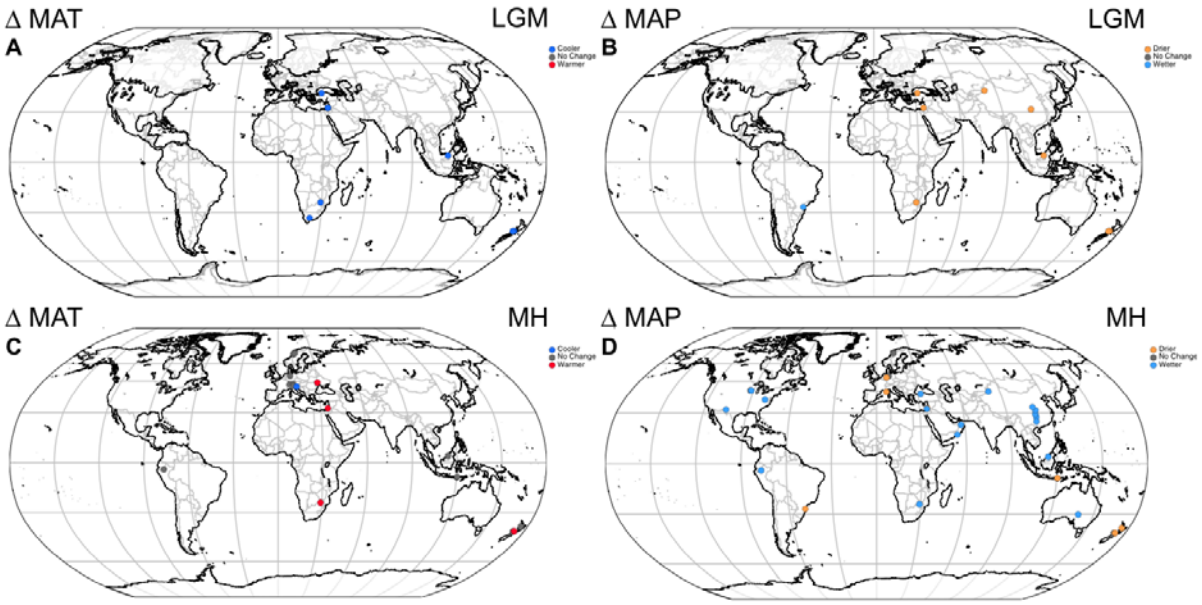


Figure S2. Quantitative reconstructions of changes in (**A**) mean annual temperature (MAT), (**B**) mean temperature of the coldest month (MTCO), (**C**) mean temperature of the warmest month (MTWA), (**D**) growing degree days above a base of 5°C (GDD5), (**E**) mean annual precipitation (MAP) and (**F**) α, the ratio of actual to equilibrium evapotranspiration (α) at the LGM.

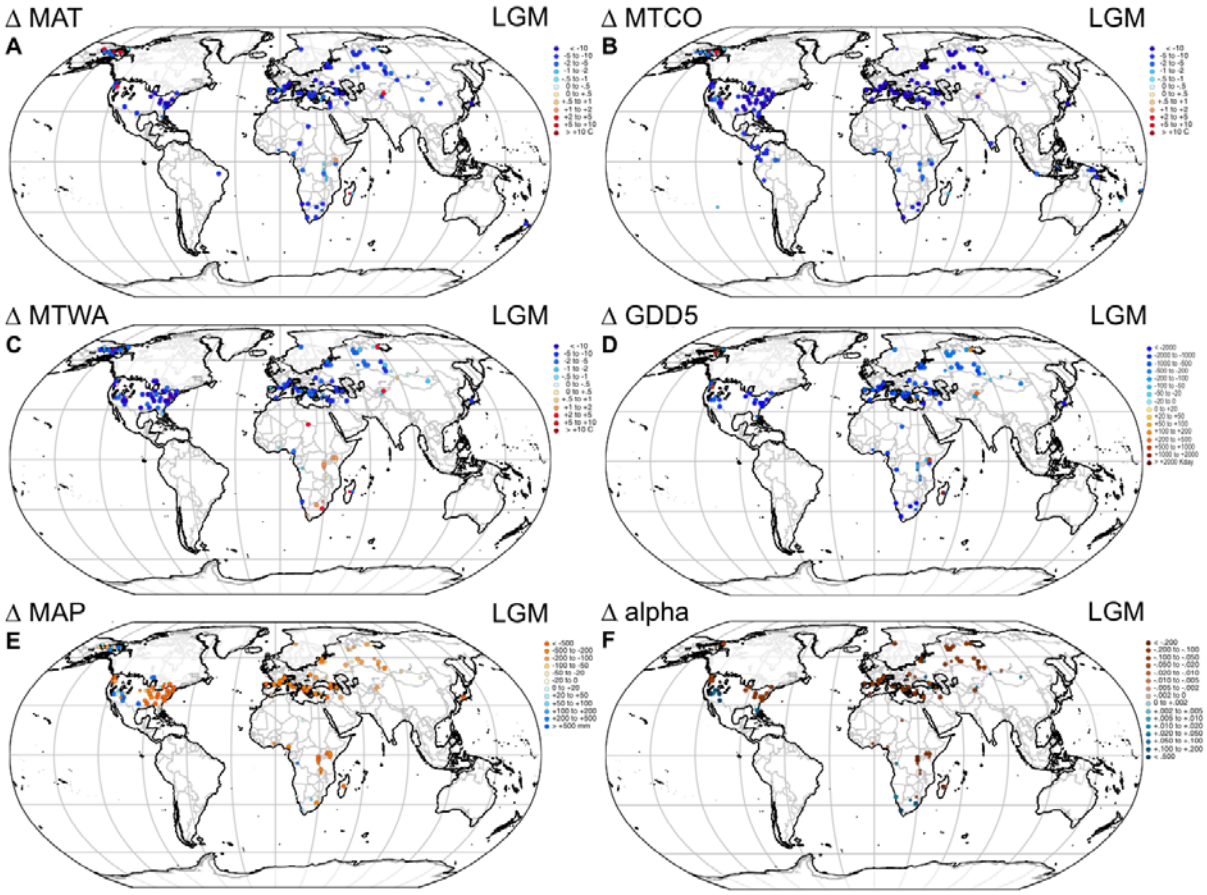


Figure S3. Quantitative reconstructions of (**A**) mean annual temperature (MAT), (**B**) mean temperature of the coldest month (MTCO), (**C**) mean temperature of the warmest month (MTWA), (**D**) growing degree days above a base of 5°C (GDD5), (**E**) mean annual precipitation (MAP) and (**F**) α, the ratio of actual to equilibrium evapotranspiration (α) at the MH.

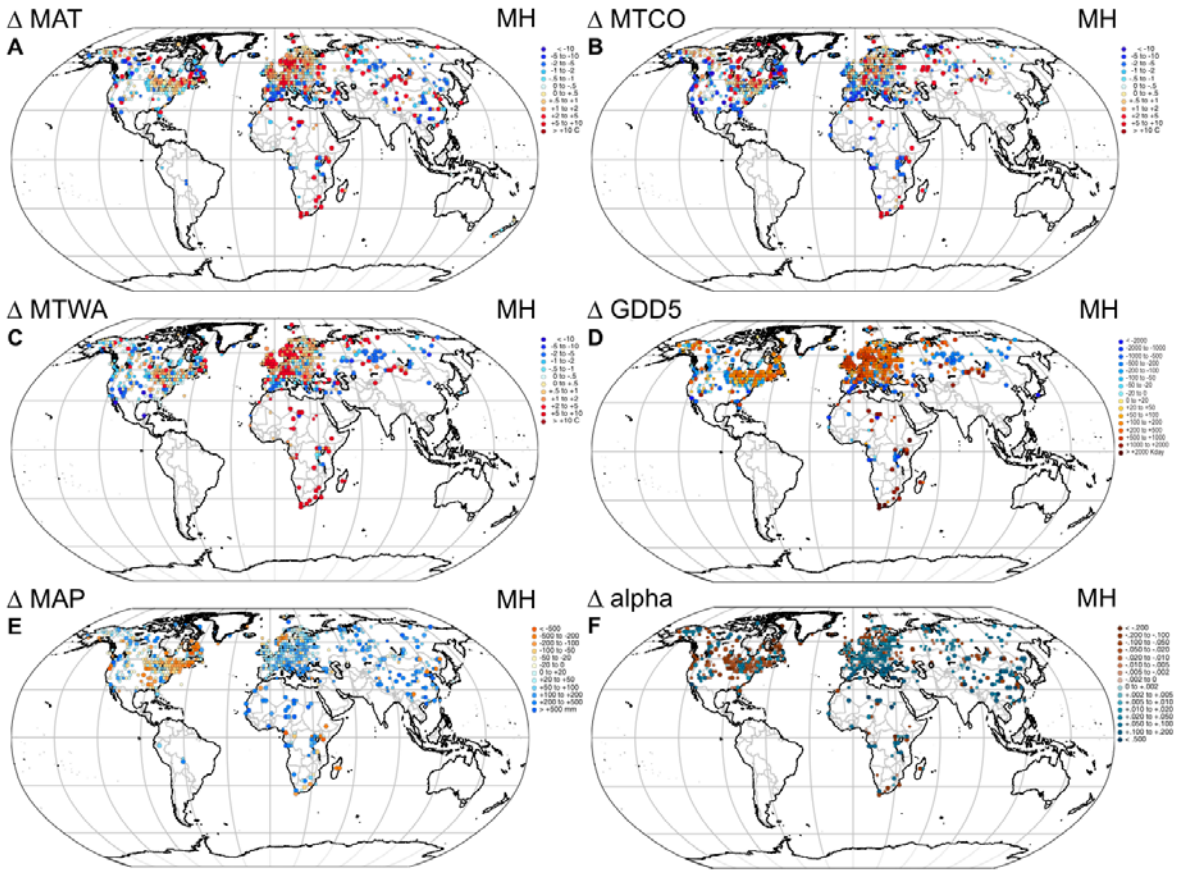
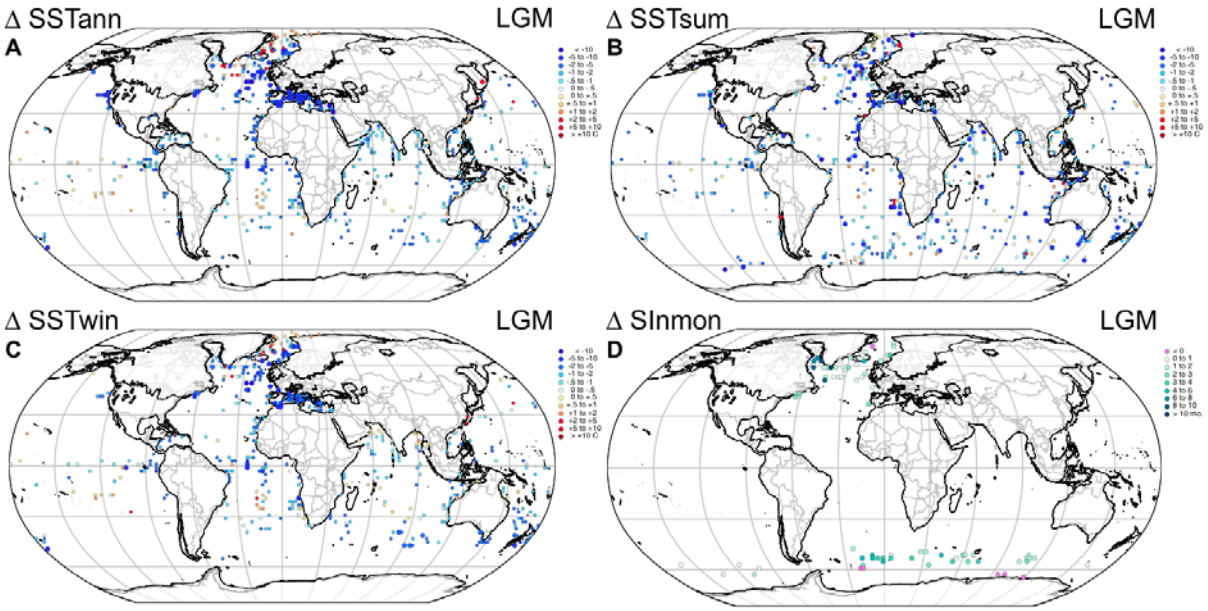
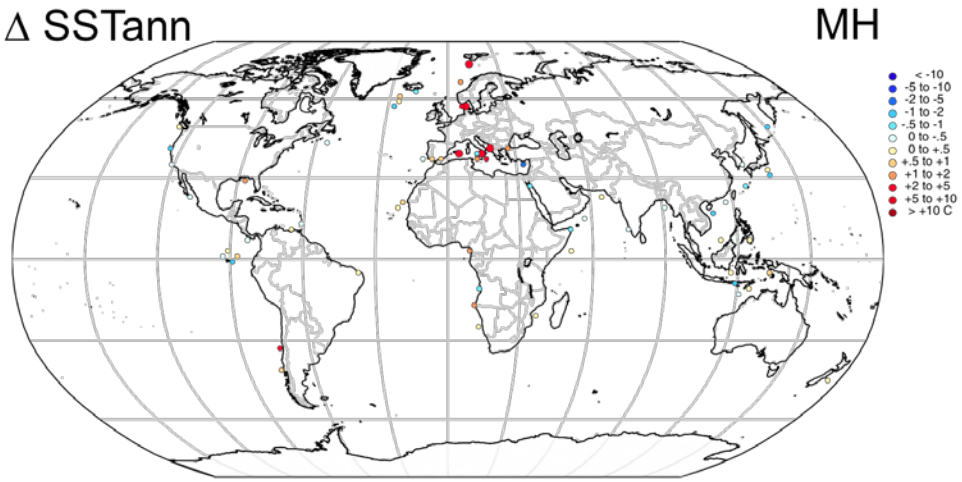


Figure S4. Quantitative reconstructions of changes in (**A**) mean annual sea-surface temperature (SSTann), (**B**) summer sea-surface temperature (SSTsum), (**C**) winter sea-surface temperature SSTwin, and (**D**) number of months with sea ice cover (SInmon) at the LGM.



858 **Figure S5.** Quantitative reconstructions of mean annual sea-surface temperature (SSTann) at the MH.

859



860

861

Figure S6. Comparison of simulated and reconstructed zonal mean annual sea-surface temperature (SSTann) at the LGM. The panels show the simulated and reconstructed values in 15° latitude bands from S to N. The median value of the observations is shown as a black vertical line, the IQR by dark grey shading and the 5-9 percentile limits by light grey shading. The models are color-coded to show whether they are PMIP2 or CMIP5 simulations, and whether they are ocean-atmosphere (OA), ocean-atmosphere-vegetation (OAV) or OA carbon-cycle (OAC) models. The simulated median for each model is shown by a vertical line, the box represents the IQR and the whiskers the 95% limits. The final panel shows SSTann by latitude for all models and all ensembles compared to the reconstructions.

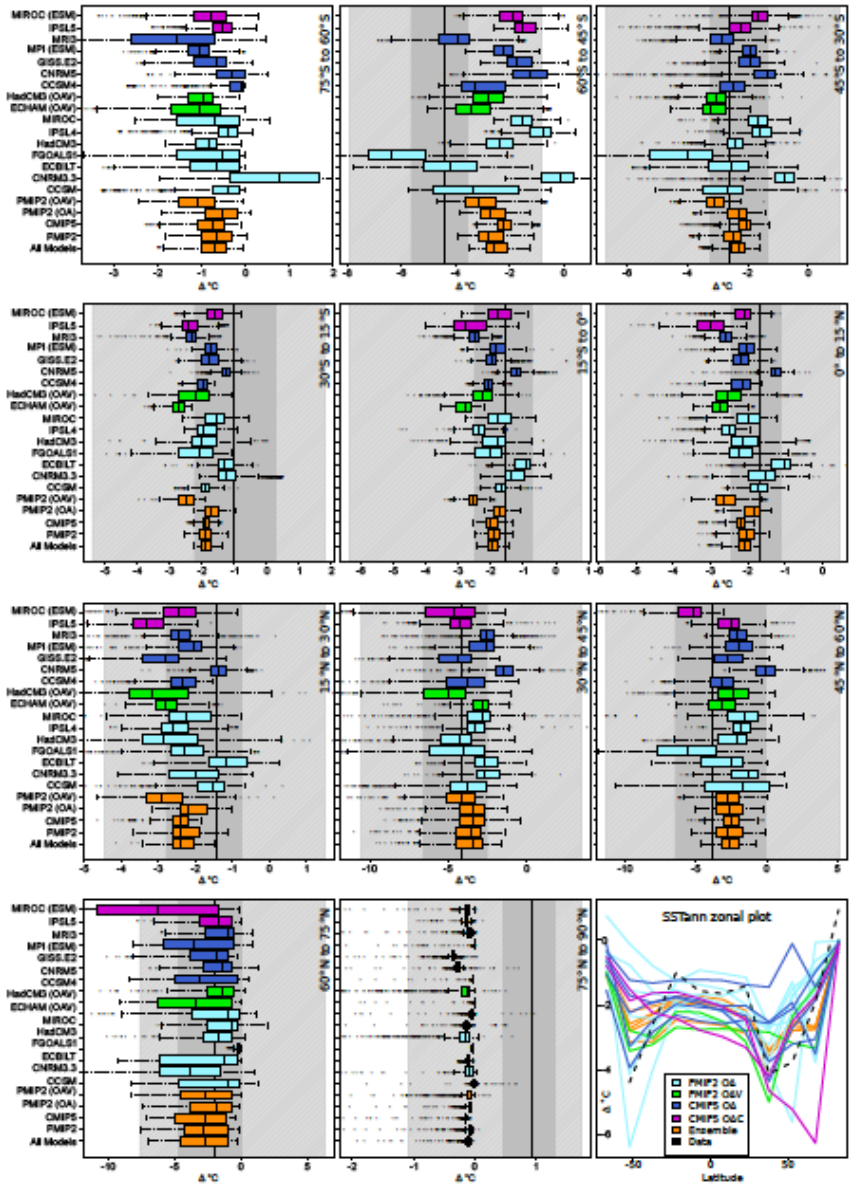


Figure S7. Comparison of simulated and reconstructed zonal sea-surface temperatures at the MH. The panels show the simulated and reconstructed values in 15° latitude bands from S to N. The median value of the observations is shown as a black vertical line, the IQR by dark grey shading and the 5 to 95 percentile limits by light grey shading. The models are color-coded to show whether they are PMIP2 or CMIP5 simulations, and whether they are ocean-atmosphere (OA), ocean-atmosphere-vegetation (OAV) or OA carbon-cycle (OAC) models. The simulated median for each model is shown by a vertical line, the box represents the IQR and the whiskers the 5 to 95 percentile limits.

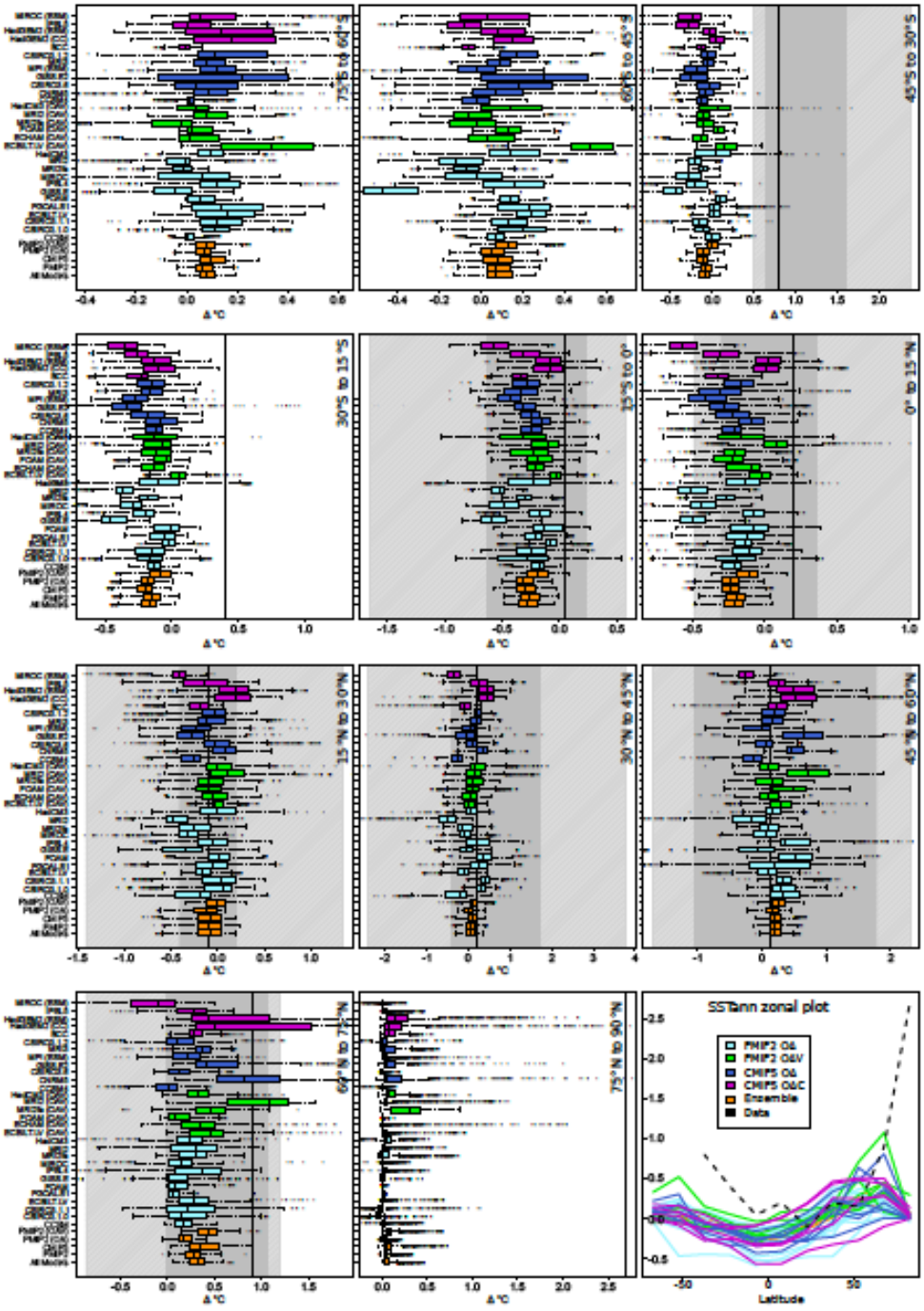


Table S1. Summary of data used for temperature reconstructions from ice core records.

Site	Latitude (°)	Longitude (°)	Modern ^{18}O (‰)	LGM ^{18}O (‰)	Anomaly ^{18}O (‰)	Sea water correction (‰)	Temperature anomaly (°C)	Error bar slope	LGM-present elevation change (m)	Elevation correction (°C)	Total temperature change (°C)
Vostok	-78.47	106.80	-57.03	-61.12	4.08	5.21	-6.51	1.653	125	1.25	-7.76
EDC	-75.10	123.35	-50.82	-56.01	5.20	6.33	-7.91	2.009	125	1.25	-9.16
EDML	-75.00	0.07	-44.91	-50.13	5.22	6.36	-7.95	2.018	100	1.00	-8.95
TALDICE	-72.82	159.18	-36.46	-41.46	5.00	6.15	-7.69	1.952	100	1.00	-8.69
Dome F	-77.32	39.67	-55.10	-59.53	4.42	5.55	-6.94	1.762	125	1.25	-8.19

Table S2. Sites from Schmittner et al. (2011) that are not included in other data sets, and have therefore been incorporated into the gridded data sets of quantitative climate variables used here for model evaluation.

Site	Latitude (°)	Longitude (°)	Elevation	Indicator	Reference
NA 87-22	55.5	-14.7	-2161	foram assemblages	Waelbroeck et al. (2001)
MD01-2461	51.8	-12.9	-1153	Mg/Ca	Peck et al. (2008)
CH 69-09	41.8	-47.4	-4100	foram assemblages	Waelbroeck et al. (2001)
PC-6	40.4	143.5	-2215	UK'37	Minoshima et al. (2007)
SU81-18	37.8	-10.2	-3135	foram assemblages	Waelbroeck et al. (2001)
MD01-2421	36.0	141.8	-2224	UK'37	Isono et al. (2009)
KT92-17 St. 14	32.6	138.6	-3252	UK'37	Sawada and Handa (1998)
KNR140-51GGC	32.6	-76.3	-1790	Mg/Ca	Carlson et al. (2008)
GeoB 7702-3	31.7	34.1	-562	TEX86	Castañeda et al. (2010)
MD98-2195	31.6	129.0	-746	UK'37	Ijiri et al. (2005)
MD02-2575	29.0	-87.1	-847	Mg/Ca	Ziegler et al. (2008)
GeoB 5844-2	27.7	34.7	-963	UK'37	Arz et al. (2003)
EN32-PC6	27.0	-91.3	-2280	Mg/Ca	Flower et al. (2004)
ODP 658C	20.8	-18.6	-2263	foram assemblages	deMenocal et al. (2000)
ODP 1144	20.1	117.6	-2037	Mg/Ca	Wei et al. (2007)
74KL	14.3	57.3	-3212	UK'37	Huguet et al. (2006)
74KL	14.3	57.3	-3212	TEX86	Huguet et al. (2006)
VM28-122	11.6	-78.4	-3623	Mg/Ca	Schmidt et al. (2004)
NIOP-905	10.8	51.9	-1567	UK'37	Huguet et al. (2006)
NIOP-905	10.8	51.9	-1567	TEX86	Huguet et al. (2006)
MD02-2529	8.2	-84.1	-1619	UK'37	Leduc et al. (2007)
ME0005A-43JC	7.9	-83.6	-1368	Mg/Ca	Benway et al. (2006)
MD01-2390	6.6	113.4	-1545	UK'37	Steinke et al. (2008)
MD01-2390	6.6	113.4	-1545	Mg/Ca	Steinke et al. (2008)
MD03-2707	2.5	9.4	-1295	Mg/Ca	Weldeab et al. (2007)
GeoB 4905	2.5	9.4	-1328	Mg/Ca	Weldeab et al. (2005)
ME0005A-24JC	0.0	-86.5	-2941	UK'37	Kienast et al. (2006)
V21-30	-1.2	-89.7	-617	UK'37	Koutavas and Sachs (2008)
V19-28	-2.4	-84.7	-2720	UK'37	Jaeschke et al. (2007)
GeoB 3910	-4.2	-36.3	-2362	UK'37	Jaeschke et al. (2007)
GeoB 3129	-4.6	-36.6	-830	Mg/Ca	Weldeab et al. (2006)
MD98-2176	-5.0	133.4	-2382	Mg/Ca	Stott et al. (2007)
GeoB 6518-1	-5.6	11.2	-962	UK'37	Schefuss et al. (2005)
MD98-2165	-9.7	118.4	-2100	Mg/Ca	Levi et al. (2007)
MD98-2170	-10.6	125.4	-832	Mg/Ca	Xu et al. (2008)
MD01-2378	-13.1	121.8	-1783	Mg/Ca	Farmer et al. (2005)
ODP 1084B	-25.5	13.0	-1992	Mg/Ca	Farmer et al. (2005)
KNR159-5-36GGC	-27.5	-46.5	-1268	Mg/Ca	Carlson et al. (2008)
GeoB 7139-2	-30.2	-72.0	-3270	UK'37	Kaiser et al. (2008)
MD03-2611	-36.7	136.7	-2420	UK'37	Calvo et al. (2007)
MD97-2121	-40.4	178.0	-3014	UK'37	Pahnke and Sachs (2006)
ODP 1233	-41.0	-74.5	-838	UK'37	Lamy et al. (2007)
TN057-21-PC2	-41.1	7.8	-4981	UK'37	Sachs et al. (2001)
TN057-21	-41.1	7.8	-4981	Mg/Ca	Barker et al. (2009)
SO136-GC11	-43.5	167.9	-1556	UK'37	Barrows et al. (2007)
MD97-2120	-45.5	174.9	-1210	UK'37	Pahnke and Sachs (2006)
Section MS2008E	34.90	113.30	200	MBT/CBT	Peterse et al. (2011)
Lake Tanganyika	-6.70	29.60	773	TEX86	Tierney et al. (2008)

Table S3. Cave records providing information for the LGM and MH but lacking information from modern speleothems about present day conditions. These sites are not used for subsequent analyses.

Site	Latitude (°)	Longitude (°)	Sample Id	Data type	References
Lynds Cave	-41.97	146.42	LYN	$\delta^{18}O$, $\delta^{13}C$, growth rate	Xia et al. (2001)
Hollywood Cave	-41.95	171.41	HW3	$\delta^{18}O$	Whittaker et al. (2011)
Jintanwan Cave	29.48	109.53	J1	$\delta^{18}O$, $\delta^{13}C$	Cosford et al. (2010)
Xianshui Caves, Yaoba Don	21.25; 28.8	110.92; 109.83	X3, YB1	$\delta^{18}O$	Cosford et al. (2008)
Hulu Cave	32.50	119.17		$\Delta^{18}O$	Wang et al. (2001); Yuan et al. (2004); Southon et al. (2012)
Okshola Cave	67.00	15.00	Oks82	$\delta^{18}O$, $\delta^{13}C$	Linge et al. (2009)
B7 Cave	51.37	7.67	B7-1, 5	$\delta^{18}O$, $\delta^{13}C$, fabric	Niggemann et al. (2003)
V11 Cave	46.60	22.70		$\delta^{18}O$, $\delta^{13}C$	Tămaș et al. (2005)
Milchbach Cave	46.62	8.08	MB-2, 3, 5, 6	$\delta^{18}O$, $\delta^{13}C$, fabric	Luetscher et al. (2011)
Villars	45.43	0.77	vil-stm 11	$\delta^{13}C$	Genty et al. (2006)
Chau	44.23	4.26	Chau-stm6	$\delta^{13}C$	Genty et al. (2006)
Buca della Renella	44.00	10.00	RL4	$\delta^{18}O$	Drysdale et al. (2006)
Pindal Cave	43.38	-4.50	CAN	$\delta^{18}O$, $\delta^{13}C$, trace elements	Moreno et al. (2010)
Ursilor Cave	46.55	22.57	PU-2	$\delta^{18}O$, $\delta^{13}C$	Onac et al. (2002)
Antro del Corchia	43.98	10.22	CC26	$\delta^{18}O$, $\delta^{13}C$	Zanchetta et al. (2007)
Jeita Cave	32.93	35.63	JeG-stm-1	$\delta^{18}O$, $\delta^{13}C$, growth rate	Verheyden et al. (2008)
Soreq Cave	32.58	35.03		$\delta^{18}O$, $\delta^{13}C$	Bar-Matthews and Ayalon (2010)
Tzavoa	31.25	35.20	TZ-14	$\delta^{18}O$	Vaks et al. (2006)
Jerusalem West Cave	31.78	35.15	AF12	$\delta^{18}O$	Frumkin et al. (1999)
Moomi Cave	12.50	54.00	M1-5	$\delta^{18}O$	Shakun et al. (2007)
McFail's Cave	38.37	-90.65	MF1	$\delta^{18}O$, $\delta^{13}C$	van Beynen et al. (2004)
Fort Stanton Cave	33.30	-105.30	FS2	$\delta^{18}O$	Asmeron et al. (2010)
Cave of the Bells	31.75	-109.25		$\delta^{18}O$	Wagner et al. (2010)
Cueva del Diablo	18.18	-99.92	CBD-2	$\delta^{18}O$	Bernal et al. (2011)
Venado Cave	10,60	-84.80	V1	$\delta^{18}O$, $\delta^{13}C$	Lachniet et al. (2004)
Lapa Grande	-13.58	-43.63	LG3, LG11	$\delta^{18}O$	Stríkis et al. (2011)
Pacupahuain Cave	-11.24	-75.82	p09-PH2	$\delta^{18}O$	Kanner et al. (2012)
La Mine Cave	35.00	9.50	Min-stm 1	$\delta^{13}C$	Genty et al. (2006)

Table S4. Cave records providing information on relative and/or absolute changes in precipitation and/or temperature at 21000 yr BP (21k) and/or at 6000 yr BP (6k). Data type is indicated by a numerical code, where 1 is $\delta^{18}O$, 2 is $\delta^{13}C$, 3 is growth rate, 4 is trace elements, 5 is fabrics, 6 is fluid inclusions, 7 is clumped isotopes, 8 is grey scale and 9 is calibration against other records. Length of record (ka) indicates the length of the record covering our periods of interest (21000 yr BP, 6000 yr BP) to the present-day or to a present hiatus.

Cave	Lat (°)	Long (°)	Sample Id	Data type	Length of record (ka)	No of dates 22-20ka	No of dates 6.5-5.5ka	Sample resoln (yrs)	T 6k (°C)	T 21k (°C)	P 6k (mm)	P 21k (mm)	T 6k Warm/ Cold	T 21k Warm/ Cold	P 6k Wet/ Dry	P 21k Wet/ Dry	References
Liang Luar Cave	-8.53	120.43	LR06-B1, B3	1,4,6	12.7	0	6	~8	-0.2 +/-1.5				No signif change		Dry		Griffiths et al. (2009); Griffiths et al. (2010a,b)
Nettlebed Cave	-41.08	172.67	MD3	1	31	1	0	~150					Warm	Cold	Dry	Dry	Hellstrom et al. (1998)
Babylon, Twin Forks, Wezpetti Caves	-41.31	171.9	BN1,2; TF2; WP1; WN4;	1,2	23.4	1	2	~41					No signif change	Cold	No change	Dry	Williams et al. (2005)
Gardners Gut	-38.25	175.10	GG1,2; RK-1	1,2	11.8	0	1	60-160					No signif change		Dry		Williams et al. (2004)
Yudnamutana Cave	-30.19	139.42	Yudnamutana	1,2	11.6-5.2	0	3	~32							Wet		Quigley et al. (2010)
Gunung Buda	4.00	114.00	BA04, SCH02, SSC01	1	33	2	1	56-72					No signif change	Cold	Wet	Dry	Partin et al. (2007)
Heshang, Dongge Caves	30.45; 25.00	110.42; 109.00	HS-4, DA	1,9	9.4	0	10	~16			100				Wet		Hu et al. (2008)
Dongge Cave	25.00	109.00	D3,4; DA	1	16	0	12	~5-20							Wet		Yuan et al. (2004); Dykoski et al. (2005); Wang et al. (2005)
Lianhua Cave	29.48	109.53	A1	1,2,4	6.6	0	1	~8							Wet		Cosford et al. (2008); Cosford et al. (2009)
Sanbao Cave	31.67	110.43	Multiple	1	18	0	6	<40-70							Wet	Dry	Wang et al. 2008; Dong et al. (2010)
Jiuxian Cave	33.57	109.10	C966-1, 2	1	19	0	2	4-112							Wet		Cai et al. (2010)
Kesang Cave	42.87	81.75	KS08-1-H, 2-H; KS06-A-H-A, -A-H-B	1	23	1	0								Wet	Dry	Cheng et al. (2012)
Soyle-grotta	66.33	13.55	SG93	1,9	10		0	~20	0.6 +/- 1.8				No signif change				Lauritzen and Lundbeg (1999)
Okshola Cave	67.00	15.00	FM3	1,2	7.2		3	~29					No signif change		No change		Linge et al. (2009)
Crag Cave	52.25	9.43	CC3	1	10.1		1	~6					No signif change		No change		McDermott et al. (2001)
B7 Cave	51.37	7.67	B7-7	1,2,5	17.6		1								Dry		Niggemann et al. (2003)
Hölloch Cave	46.97	8.78	HÖL-1	1,2	14.2		0						No signif change/Cold				Wurth et al. (2004)
Poleva	48.00	30.00	PP9,10	1	~60-0	0	2						Warm				Constantin et al. (2007)
Spannagel Cave	47.05	11.40	COMNISPA	1	9		4						No signif change				Vollweiler et al. (2006)
Grotta Savi	45.62	13.89	SV1	1	17		1	18-125					Cold				Frisia et al. (2005)
Grotta di Ernesto	45.98	11.66	ER76	1,2, 3,5	8		1	~50					No signif change				McDermott et al. (1999); Scholz et al. (2012)
Grotto de Clamouse	42.71	3.61	CL26	1,5	11		1	~40					Warm				McDermott et al. (1999)
Cova da Arcoia	42.61	7.09	ESP-03	1,2,5	9.3		0								Dry		Railsback et al. (2011)

Soreq Cave	32.58	35.03	12-Z-B	7						-6.5 +/- 0.5				Cold			Affek et al. (2008)
Soreq Cave	32.58	35.03	12-Z-C	1,6						- 7+/- 2				Cold			McGarry et al. (2004)
Peqiin Cave + comp. Soreq Cave	32.97	35.33	Multi	1,2	250	0	1				-20				Wet		Bar-Matthews et al. (2003)
Sofular Cave	41.42	31.93	So-1, So-2	1,2	50	2	2	5-50						Cold	Wet/Dry	Dry	Fleitmann et al. (2009a); Göktürk et al. (2011)
Soreq Cave	32.58	35.03	Multi	1,2,4	60	3	1						Warm	Cold	Wet	Dry	Bar-Matthews et al. (1999); Kaufman et al. (1998)
Hoti Cave	23.08	57.35	H12	1	6.3		6	1.5-10							Wet		Fleitmann et al. (2003); Fleitmann et al. (2007); Fleitmann et al. (2009b)
Qunf Cave	17.17	54.30	Q5	1	10.6		2	1.5-10							Wet		Fleitmann et al. (2007); Fleitmann et al. (2009b)
Cold Water Cave	43.47	-91.97	CWCIS	1,2	7.8										Wet		Denniston et al. (1999)
Spring Valley, Mystery Cave	43.53	-92.52	SV1, MC28	1,2	8.5										No change/ Dry		Denniston et al. (1999)
Buckeye Creek Cave	37.98	-79.59	BCC-2, 4,6	1,4	7.8			~60							Wet		Springer et al. (2008); Hardt et al. (2010)
Pink Panther Cave	32,.8	-104.83	PP1	1	12.3		2	~17							Wet		Asmeron et al. (2007)
Cueva del Tigre Perdido	-4.06	-76.69	NC-B	1, 6	13.3		2		-0.5+/- 1.5		75+/- 25		No signif change		Wet		van Breukelen et al. (2008)
Boutverá Cave	-26.78	-48.84	Bt2	1,4	116										Dry	Wet	Cruz et al. (2007); Cruz et al. (2005); Wang et al. (2007)
Cold Air Cave	-23.98	29.18	T7, T8	1,2,8	24.4	1	2	6-500					Warm	Cold	Wet	Dry	Lee-Thorp et al. (2001); Holmgren et al. (2003)
Cango Cave	-33.38	22.22	V3	1,2	50	1	0							Cold			Talma and Vogel (1992)

Table S5. The boundary conditions used for the PMIP 2 and CMIP5 paleo-experiments. Boundary conditions that are different between different sets of simulations are highlighted in yellow; blue highlighting shows boundary conditions not included in the PMIP2 experiments (see PMIP website for more details of the protocols: http://pmip3.lsce.ipsl.fr/).

		PMIP2	**CMIP5/PMIP3**
Mid Holocene	Insolation	**eccentricity** = 0.018682 **obliquity** = 24.105° **perihelion-180°** = 0.87°	**eccentricity** = 0.018682 **obliquity** = 24.105° **perihelion-180°** = 0.87°
	Solar constant	As in PI simulation	As in PI simulation
	Trace gases	**CO_2** = 280 ppm **CH_4** = 650 ppb **N_2O** = 270 ppb **CFC** = 0 **O_3** = not considered	**CO_2** = 280 ppm **CH_4** = 650 ppb **N_2O** = 270 ppb **CFC** = 0 **O_3** = same as in CMIP5 PI
	Aerosols	As in PI simulation	As in PI simulation
	Ice sheets and land-sea distribution	As in PI simulation	As in PI simulation
	Land-sea mask	As in PI simulation	As in PI simulation
	Vegetation and land surface	Either prescribed modern vegetation or computed using a dynamic vegetation module	Computed or prescribed as in PI, with phenology computed for models with active carbon cycle or prescribed from data
	Carbon cycle	Not considered	Either not considered or interactive, with atmospheric concentration prescribed and ocean and land carbon fluxes diagnosed as recommended in CMIP5
Last Glacial Maximum	Insolation	**eccentricity** = 0.018994 **obliquity** = 22.949° **perihelion-180°** = 114.42°	**eccentricity** = 0.018994 **obliquity** = 22.949° **perihelion-180°** = 114.42°
	Solar constant	As in PI simulation	As in PI simulation
	Trace gases	**CO_2** = 185 ppm **CH_4** = 350 ppb **N_2O** = 200 ppb **CFC** =0 **O_3** = same as in PI	**CO_2** = 185 ppm **CH_4** = 350 ppb **N_2O** = 200 ppb **CFC** =0 **O_3** = same as in PI
	Aerosols	As in PI simulation	As in PI simulation
	Ice sheet	Peltier et al. (2004)	Blended ice sheet (Braconnot et al. 2012)
	Land-sea mask	Prescribed following Peltier et al. (2004) land-sea mask	Prescribed from blended ice-sheet land-sea mask
	Freshwater	Excess freshwater added to ocean in 3 different regions	Excess freshwater added to ocean in 3 different regions
	River runoff	As in PI or river pathway modified (no formal protocol)	As in PI or river pathway modifier according to PMIP protocol
	Mean ocean salinity	Not considered	+1 PSU everywhere
	Carbon cycle	Not considered	Interactive, with atmospheric concentration prescribed and ocean and land carbon fluxes diagnosed following CMIP5 recommendations

Table S6. Details of the models and paleo-simulations from the PMIP2 and CMIP5 archives. The model attributes given here are derived from the archived netcdf files, but these are not always complete or accurate, and thus there may be some inconsistencies or missing information about specific models. Abbreviated codes used to refer to specific models in the text and in other tables are given here.

Archive	Model name	Code	Type	Resolution (number of gridcells: latitude, longitude)			Year length	Simulations		Reference
				Atmosphere	Ocean	Sea Ice		MH	LGM	
PMIP2	CCSM	CCSM	OA	64, 128	395, 320	395, 320	365	X	X	Otto-Bliesner et al. (2006)
PMIP2	CNRM-CM33	CNRM3.3	OA	64, 128	170, 180	170, 180	365		X	Royer (2008); Johns et al. (2011)
PMIP2	CSIRO-Mk3L-1.0	CSIRO3.1.0	OA	56, 64	56, 64	56, 64	365	X		Phipps (2006)
PMIP2	CSIRO-Mk3L-1.1	CSIRO3.1.1	OA	56, 64	112, 128	56, 64	365	X		Phipps (2006)
PMIP2	ECBILTCLIO	ECBILT	OA	32, 64	72, 144	72, 144	360		X	de Vries and Weber (2005)
PMIP2	ECBILTCLIOVECODE	ECBILT.LV	OA	32, 64	60, 120	60, 120	360	X		Renssen et al. (2005)
PMIP2	ECHAM5-MPIOM1	ECHAM	OA	48, 96	180, 360	48, 96	365-366	X		Jungclaus et al. (2006)
PMIP2	FGOALS-1.0g	FGOALS1	OA	60, 128	170, 360	170, 360	365	X	X	Yu et al. (2002); Yu et al. (2004)
PMIP2	FOAM	FOAM	OA	40, 40	128, 128	128, 128	360	X		Jacob et al. (2001)
PMIP2	GISSmodelE	GISS.E	OA	46, 72	46, 72	46, 72	365	X		Schmidt et al. (2006)
PMIP2	HadCM3M2	HadCM3	OA	73, 96	144, 288	144, 288	360		X	Gordon et al. (2000)
PMIP2	IPSL-CM4-V1-MR	IPSL4	OA	72, 96	170, 180	170, 180	360	X	X	Marti et al. (2005)
PMIP2	MIROC3.2	MIROC	OA	64, 128	192, 256	192, 256	360	X	X	K-1-Model-Developers (2004)
PMIP2	MRI-CGCM2.3.4fa	MRI2fa	OA	64, 128	111, 144	111, 144	365	X		Yukimoto et al. (2006)
PMIP2	MRI-CGCM2.3.4nfa	MRI2	OA	64, 128	111, 144	111, 144	365	X		Yukimoto et al. (2006)
PMIP2	UBRIS-HadCM3M2	HadCM3	OA	73, 96	144, 288	144, 288	360	X		Gordon et al. (2000)
PMIP2	ECBILTCLIOVECODE	ECBILT.LV (OAV)	OAV	32, 64	60, 120	60, 120	360	X		Renssen et al. (2005)
PMIP2	ECHAM53-MPIOM127-LPJ	ECHAM (OAV)	OAV	48, 96	180, 360	180, 360	365-366		X	Roeckner et al. (2003); Marsland et al. (2003); Mikolajewicz et al. (2007)
PMIP2	FOAM	FOAM (OAV)	OAV	40, 40	128, 128	128, 128	360	X		Notaro et al. (2005)
PMIP2	HadCM3M2	HadCM3 (OAV)	OAV	73, 96	144, 288	144, 288	360		X	Gordon et al. (2000); Essery et al. (2003)
PMIP2	MRI-CGCM2.3.4fa	MRI2fa (OAV)	OAV	64, 128	111, 144	111, 144	365	X		Yukimoto et al. (2006)
PMIP2	MRI-CGCM2.3.4nfa	MRI2 (OAV)	OAV	64, 128	111, 144	111, 144	365	X		Yukimoto et al. (2006)
PMIP2	UBRIS-HadCM3M2	HadCM3 (OAV)	OAV	73, 96	144, 288	144, 288	360	X		Gordon et al. (2000)
CMIP5	CCSM4	CCSM4	OA	192, 288	320,384	320,384	365	X	X	Gent et al. (2011)
CMIP5	CNRM-CM5	CNRM5	OA	128, 256	292, 362	292, 362	365-366	X		Voldoire et al. (2012)
CMIP5	CSIRO-Mk3-6-0	CSIRO3.6	OA	96, 192	189, 192	96, 192	365	X		Rotstayn et al. (2010)
CMIP5	CSIRO-Mk3L-1-2 (UNSW)	CSIRO3.1.2	OA	56, 64	128, 225	56, 64	365	X		Phipps et al. (2011); Wu et al. (2008)
CMIP5	EC-EARTH (KNMI)	EARTH	OA	160, 320	362, 292	362, 292	365-366	X		Hazeleger et al. (2012)
CMIP5	FGOALS-s2 (LASG/IAP)	FGOALS2	OA	108, 128	196, 360	196, 360	365	X		No publication
CMIP5	FGOALS-g2 (LASG/IAP)	FGOALSG2	OA	60, 128	196, 360	196, 360	365	X		No publication
CMIP5	GISS-E2-R	GISS.E2	OA	90, 144	90, 144	90, 144	365	X	X	No publication
CMIP5	MPI-ESM-P	MPI (ESM)	OA	96, 192	220, 256	220, 256	365-366	X	X	No publication
CMIP5	MRI-CGCM3	MRI3	OA	160, 320	360, 368	360, 368	365	X		Yukimoto et al. (2011)
CMIP5	BCC-CSM1-1	BCC	OAC	64, 128	232, 360	232, 360	365	X		Wu et al. (2010); Marti et al. (2010)
CMIP5	COSMOS-ASO	COSMOS	OAC	96, 48	120, 101	120, 101	360		X	Budich et al. (2010)
CMIP5	IPSL-CM5A-LR	IPSL5	OAC	96, 96	149, 182	149, 182	365	X	X	Watanabe et al. (2011)
CMIP5	MIROC-ESM	MIROC (ESM)	OAC	64, 128	192, 256	192, 256	365	X	X	Collins et al. (2011)
CMIP5	HadGEM2-CC	HadGEM2 (CC)	OAC	145, 192	216, 360	216, 360	360	X		Martin et al. (2011); Jones et al. (2011)
CMIP5	HadGEM2-ES	HadGEM2 (ESM)	OAC	145, 192	216, 360	216, 360	360	X		Martin et al. (2011); Jones et al. (2011)

Table S7. Ensemble values of global metrics for the Last Glacial Maximum (LGM) simulations. The ensembles are all of the models (ALL), all of the PMIP2 models (PMIP2), all of the CMIP5 models (CMIP5), all of the PMIP2 ocean-atmosphere models (PMIP2 OA), all of the PMIP2 ocean-atmosphere-vegetation models (PMIP2 OAV), all of the CMIP5 ocean-atmosphere models (CMIP5 OA) and all of the CMIP5 ocean-atmosphere-carbon cycle models (CMIP5 OAC).

		All		PMIP2		CMIP5		PMIP2OA		PMIP2OAV		CMIP5OA		CMIP5OAC	
Variable	Metric	mean	median	mean	median	mean	median	mean	median	mean	median	mean	median	mean	median
SSTann	Median Bias	-0.37	-0.32	-0.37	-0.38	-0.34	-0.28	-0.19	-0.16	-0.91	-0.91	-0.19	-0.21	-0.68	-0.68
	IQR Ratio	0.27	0.29	0.36	0.38	0.23	0.27	0.33	0.37	0.29	0.29	0.18	0.23	0.53	0.53
	1–τ	0.68	0.72	0.68	0.70	0.77	0.82	0.68	0.72	0.77	0.77	0.79	0.85	0.78	0.78
	Fuzzy Distance	1.20	1.19	1.18	1.20	1.26	1.23	1.16	1.18	1.42	1.42	1.21	1.22	1.35	1.19
SSTsum	Median Bias	-0.44	-0.35	-0.51	-0.38	-0.38	-0.36	-0.34	-0.13	-0.93	-0.93	-0.28	-0.32	-0.58	-0.58
	IQR Ratio	0.32	0.29	0.39	0.38	0.27	0.29	0.39	0.32	0.35	0.35	0.27	0.26	0.52	0.52
	1–τ	0.97	0.97	0.96	0.96	0.99	0.99	0.96	0.97	0.97	0.97	1.00	1.01	0.99	0.99
	Fuzzy Distance	1.46	1.44	1.52	1.51	1.43	1.45	1.50	1.43	1.63	1.63	1.41	1.45	1.55	1.55
SSTwin	Median Bias	-0.32	-0.23	-0.40	-0.32	-0.28	-0.20	-0.20	-0.07	-0.80	-0.80	-0.15	-0.15	-0.48	-0.48
	IQR Ratio	0.31	0.28	0.38	0.39	0.27	0.30	0.39	0.35	0.40	0.40	0.24	0.29	0.45	0.45
	1–τ	0.98	0.98	0.98	0.98	0.99	1.00	0.98	1.00	0.97	0.97	0.99	0.99	0.98	0.98
	Fuzzy Distance	1.34	1.31	1.37	1.31	1.32	1.35	1.37	1.33	1.44	1.44	1.31	1.35	1.42	1.42
SInmon	Median Bias	1.52	-0.69	2.07	0.82	0.60	-0.69	2.15	0.82	0.82	0.82	0.12	-0.69	-0.69	-0.69
	IQR Ratio	1.10	1.27	1.21	1.48	1.15	0.84	1.06	1.27	1.69	1.69	1.14	0.63	1.06	1.06
	1–τ	0.73	0.84	0.74	0.76	0.75	0.84	0.71	0.72	0.77	0.77	0.78	0.94	0.75	0.75
	Fuzzy Distance	0.97	1.02	1.17	1.10	0.83	1.00	1.33	0.98	0.95	0.95	0.70	1.00	1.05	1.05
MAT	Median Bias	0.83	1.12	1.06	1.26	0.93	1.19	1.45	1.52	0.15	0.15	1.68	1.83	0.12	0.43
	IQR Ratio	0.74	0.67	0.64	0.66	0.84	0.66	0.59	0.61	0.92	0.92	0.75	0.61	0.56	0.45
	1–τ	0.69	0.67	0.70	0.68	0.68	0.67	0.66	0.66	0.78	0.78	0.71	0.68	0.66	0.67
	Fuzzy Distance	2.35	2.44	2.40	2.49	2.58	2.55	2.58	2.50	3.25	3.25	2.80	2.75	2.52	2.24
MTCO	Median Bias	3.92	4.16	4.21	4.61	3.65	3.75	4.77	5.04	3.21	3.21	4.46	4.82	2.51	2.81
	IQR Ratio	0.69	0.74	0.57	0.63	0.79	0.78	0.56	0.58	0.58	0.58	0.75	0.77	0.54	0.45
	1–τ	0.68	0.69	0.69	0.70	0.67	0.67	0.66	0.63	0.76	0.76	0.69	0.69	0.62	0.62
	Fuzzy Distance	4.16	4.30	4.42	4.42	4.17	4.20	4.27	3.90	4.69	4.69	4.71	4.45	3.70	4.35
MTWA	Median Bias	-0.47	-0.62	-0.58	-0.58	-0.47	-0.33	-0.51	-0.49	-1.47	-1.47	0.10	0.36	-1.07	-1.59
	IQR Ratio	0.43	0.48	0.54	0.52	0.28	0.43	0.57	0.52	0.57	0.57	0.36	0.35	0.50	0.52
	1–τ	0.69	0.72	0.70	0.75	0.74	0.70	0.71	0.73	0.72	0.72	0.85	0.83	0.71	0.73
	Fuzzy Distance	2.81	2.95	2.77	3.07	2.53	2.65	2.77	2.81	3.69	3.69	3.32	3.05	3.12	3.04
GDD5	Median Bias	-282.19	-301.42	-259.56	-271.06	-283.12	-308.10	-197.04	-259.31	-416.01	-416.01	-178.11	-245.74	-510.64	-481.68
	IQR Ratio	0.48	0.55	0.49	0.52	0.51	0.62	0.47	0.50	0.58	0.58	0.45	0.59	0.70	0.77
	1–τ	0.62	0.61	0.61	0.56	0.63	0.64	0.61	0.57	0.64	0.64	0.61	0.62	0.66	0.66
	Fuzzy Distance	349.25	390.89	333.48	392.09	379.51	383.58	311.58	387.40	437.89	437.89	394.25	394.03	521.62	546.02
MAP	Median Bias	123.96	132.87	157.30	145.68	50.51	59.52	158.18	127.15	167.15	167.15	75.83	84.97	31.83	42.48
	IQR Ratio	0.61	0.59	0.71	0.76	0.56	0.65	0.76	0.77	0.72	0.72	0.56	0.66	0.67	0.72
	1–τ	1.03	1.03	1.11	1.12	0.90	0.90	1.09	1.08	1.15	1.15	0.93	0.91	0.85	0.87
	Fuzzy Distance	196.73	189.32	184.02	200.34	187.72	177.07	177.52	192.38	225.30	225.30	169.64	167.67	207.86	199.11
α	Median Bias	0.03	0.03	0.05	0.06	0.00	0.01	0.05	0.05	0.09	0.09	0.02	0.01	-0.01	0.00
	IQR Ratio	2.60	2.45	2.18	2.36	2.98	2.61	2.02	2.23	2.73	2.73	2.68	2.29	2.97	2.98
	1–τ	0.76	0.75	0.76	0.79	0.76	0.76	0.75	0.76	0.84	0.84	0.74	0.74	0.79	0.77
	Fuzzy Distance	0.11	0.11	0.10	0.11	0.11	0.11	0.09	0.10	0.13	0.13	0.11	0.10	0.13	0.11

Table S8. Ensemble values of global metrics for the Mid-Holocene (MH) simulations. The ensembles are all of the models (ALL), all of the PMIP2 models (PMIP2), all of the CMIP5 models (CMIP5), all of the PMIP2 ocean-atmosphere models (PMIP2 OA), all of the PMIP2 ocean-atmosphere-vegetation models (PMIP2 OAV), all of the CMIP5 ocean-atmosphere models (CMIP5 OA) and all of the CMIP5 ocean-atmosphere-carbon cycle models (CMIP5 OAC).

		All Models		PMIP2		CMIP5		PMIP2OA		PMIP2OAV		CMIP5OA		CMIP5OAC	
Variable	Metric	mean	median	mean	median	mean	median	mean	median	mean	median	mean	median	mean	median
SSTann	Median Bias	-0.29	-0.29	-0.29	-0.27	-0.31	-0.28	-0.31	-0.31	-0.22	-0.21	-0.25	-0.29	-0.29	-0.24
	IQR Ratio	0.33	0.28	0.34	0.31	0.28	0.28	0.34	0.34	0.20	0.20	0.32	0.33	0.32	0.40
	1–τ	0.90	0.89	0.90	0.91	0.91	0.91	0.95	0.95	0.88	0.84	0.89	0.89	0.90	0.94
	Fuzzy Distance	0.48	0.49	0.51	0.50	0.45	0.48	0.50	0.50	0.44	0.44	0.49	0.53	0.49	0.50
MAT	Median Bias	0.12	0.12	0.22	0.18	-0.02	0.01	0.14	0.12	0.27	0.27	-0.02	0.00	0.09	0.10
	IQR Ratio	0.15	0.14	0.15	0.16	0.18	0.15	0.12	0.15	0.24	0.20	0.16	0.13	0.24	0.18
	1–τ	0.90	0.89	0.91	0.91	0.90	0.88	0.92	0.93	0.92	0.90	0.89	0.89	0.92	0.89
	Fuzzy Distance	1.08	1.09	1.08	1.09	1.09	1.09	1.09	1.08	1.14	1.07	1.08	1.09	1.08	1.07
MTCO	Median Bias	0.02	0.04	0.05	0.05	-0.07	-0.05	-0.06	-0.05	0.18	0.19	-0.06	-0.05	-0.07	-0.09
	IQR Ratio	0.12	0.12	0.12	0.13	0.16	0.17	0.12	0.12	0.18	0.15	0.13	0.16	0.27	0.26
	1–τ	0.91	0.90	0.93	0.93	0.91	0.91	0.96	0.97	0.92	0.91	0.89	0.88	0.95	0.92
	Fuzzy Distance	1.32	1.29	1.30	1.29	1.31	1.26	1.26	1.28	1.34	1.34	1.29	1.28	1.37	1.35
MTWA	Median Bias	1.33	1.30	1.43	1.39	1.20	1.18	1.39	1.34	1.47	1.40	1.19	1.19	1.21	1.14
	IQR Ratio	0.25	0.23	0.25	0.27	0.23	0.22	0.25	0.26	0.29	0.30	0.23	0.23	0.25	0.28
	1–τ	1.05	1.07	1.06	1.08	1.03	1.05	1.08	1.09	1.01	1.03	1.07	1.09	0.97	0.99
	Fuzzy Distance	1.75	1.70	1.80	1.78	1.68	1.63	1.73	1.67	1.77	1.73	1.56	1.56	1.67	1.61
GDD5	Median Bias	27.36	22.95	38.67	31.51	10.12	7.36	27.77	20.41	57.15	44.59	2.21	2.43	18.11	19.70
	IQR Ratio	0.14	0.11	0.13	0.12	0.15	0.12	0.13	0.12	0.20	0.17	0.14	0.14	0.20	0.19
	1–τ	0.98	1.01	1.01	1.02	0.95	0.99	1.03	1.05	0.98	0.96	1.01	1.06	0.91	0.89
	Fuzzy Distance	192.10	189.91	194.57	192.55	192.03	193.27	191.82	192.87	199.79	191.83	195.48	192.11	184.57	188.16
MAP	Median Bias	-50.19	-49.68	-46.95	-46.42	-53.91	-53.44	-47.82	-47.31	-43.39	-41.57	-53.08	-51.00	-53.17	-53.39
	IQR Ratio	0.13	0.13	0.15	0.14	0.14	0.15	0.15	0.14	0.17	0.14	0.14	0.16	0.22	0.20
	1–τ	0.87	0.87	0.91	0.91	0.85	0.86	0.93	0.93	0.87	0.86	0.85	0.84	0.87	0.93
	Fuzzy Distance	85.41	85.90	85.28	83.45	86.83	88.76	86.84	84.99	82.78	80.55	87.98	85.16	89.05	91.47
α	Median Bias	-0.04	-0.03	-0.03	-0.03	-0.04	-0.04	-0.03	-0.03	-0.03	-0.02	-0.04	-0.04	-0.04	-0.04
	IQR Ratio	0.57	0.57	0.53	0.51	0.64	0.67	0.57	0.57	0.54	0.49	0.63	0.64	0.73	0.70
	1–τ	0.93	0.94	0.95	0.96	0.91	0.92	0.96	0.96	0.93	0.90	0.90	0.89	0.95	0.99
	Fuzzy Distance	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05

Table S9. Summary of the metrics for the Last Glacial Maximum (LGM) simulations for each of the individual models, calculated for the complete range of variables.

		PMIP2 OA							PMIP2 OAV		CMIP5 OA					CMIP5 OAC		
Variable	Metric	CCSM	CNRM3.3	ECBILT	FGOALS1	HadCM3	IPSL4	MIROC	ECHAM (OAV)	HadCM3 (OAV)	CCSM4	CNRM5	GISS.E2	MPI (ESM)	MRI3	COSMOS	IPSL5	MIROC (ESM)
SSTann	Median Bias	-0.02	0.33	0.22	-0.62	-0.48	-0.49	-0.1	-0.93	-0.81	-0.46	0.61	-0.24	-0.19	-0.7	NA	-0.99	-0.25
	IQR Ratio	0.51	0.49	0.72	0.67	0.5	0.38	0.38	0.23	0.5	0.36	0.18	0.45	0.29	0.25	NA	0.44	0.68
	1-τ	0.77	0.88	0.72	0.74	0.79	0.84	0.9	0.92	0.77	0.85	0.96	0.83	0.9	0.76	NA	0.78	0.77
	Fuzzy Distance	1.19	1.30	1.18	1.36	1.25	1.38	1.30	1.48	1.39	1.27	1.32	1.22	1.28	1.41	NA	1.50	1.25
SSTsum	Median Bias	-0.18	0.38	0.11	-0.88	-0.51	-0.38	-0.08	-1.01	-0.81	-0.52	0.67	-0.52	-0.2	-0.73	NA	-1	-0.16
	IQR Ratio	0.58	0.43	0.97	1.11	0.42	0.47	0.31	0.31	0.49	0.54	0.25	0.47	0.28	0.32	NA	0.55	0.59
	1-τ	0.94	0.98	0.99	0.95	1	0.99	1	0.99	0.99	0.99	1.03	0.97	1	1.03	NA	0.97	1
	Fuzzy Distance	1.48	1.68	1.84	1.93	1.54	1.60	1.53	1.67	1.68	1.60	1.63	1.58	1.44	1.59	NA	1.69	1.59
SSTwin	Median Bias	0.07	0.66	-0.23	-0.65	-0.42	-0.14	0.1	-0.83	-0.76	-0.47	0.52	-0.02	-0.16	-0.62	NA	-0.55	-0.18
	IQR Ratio	0.58	0.53	0.86	0.93	0.47	0.6	0.31	0.33	0.51	0.38	0.33	0.35	0.37	0.37	NA	0.51	0.54
	1-τ	0.95	1	0.97	0.95	0.99	0.98	1.01	0.95	0.98	0.97	1.02	0.95	0.95	0.98	NA	0.96	0.97
	Fuzzy Distance	1.49	1.63	1.71	1.70	1.35	1.38	1.29	1.43	1.47	1.44	1.46	1.28	1.32	1.48	NA	1.50	1.42
SInmon	Median Bias	0.81	-0.69	2.31	2.31	0.31	NA	-0.69	-0.69	0.31	0.31	-0.69	-0.69	-0.69	1.31	NA	-0.69	-0.69
	IQR Ratio	2.53	0	2.11	4.22	1.27	NA	0	2.11	1.69	2.11	0	0	0.84	1.69	NA	1.16	0.84
	1-τ	0.72	0.95	0.71	0.88	0.78	NA	1.07	0.85	0.79	0.74	0.93	1.14	0.93	0.77	NA	0.74	0.85
	Fuzzy Distance	2.53	NA	2.11	4.22	1.27	NA	NA	2.11	1.69	2.11	NA	NA	0.84	1.69	NA	1.16	0.84
MAT	Median Bias	1.8	2.97	2.89	0.86	0.15	0.6	2.47	1.33	-1.05	0.88	3.77	0.54	1.68	2.01	0.82	0.43	-0.73
	IQR Ratio	0.74	0.27	0.61	0.72	1.1	0.68	0.42	0.44	1.45	0.86	0.33	1.18	0.73	0.51	0.44	0.44	1.35
	1-τ	0.57	1	0.85	0.7	0.75	0.68	0.6	0.89	0.76	0.7	0.67	0.77	0.71	0.71	0.73	0.65	0.73
	Fuzzy Distance	2.63	3.30	2.93	2.52	2.70	2.82	2.36	2.78	3.38	2.74	3.57	3.51	2.80	3.04	2.39	2.21	3.26
MTCO	Median Bias	5.17	5.78	7.29	5.62	3.37	4.14	6.23	4.53	2.9	4.61	5.9	4.11	4.36	4.52	3.12	3.23	2.44
	IQR Ratio	0.94	0.32	0.26	0.69	0.85	0.53	0.53	0.39	0.94	0.92	0.62	0.69	0.7	0.78	0.32	0.43	1.03
	1-τ	0.62	0.82	0.78	0.77	0.77	0.75	0.7	0.8	0.75	0.69	0.66	0.72	0.67	0.75	0.61	0.58	0.69
	Fuzzy Distance	4.96	5.52	6.24	5.05	4.43	4.26	4.91	4.33	4.75	4.34	4.59	5.19	4.47	3.61	4.32	4.22	4.37
MTWA	Median Bias	0.66	1.51	1.15	-2.68	-2.21	-1.43	0.21	0.26	-3.95	-1.41	4.74	-1.68	-0.13	1.16	-1	-0.96	-2.3
	IQR Ratio	0.52	0.29	0.43	0.8	0.77	0.7	0.62	0.38	0.71	0.5	0.58	0.86	0.53	0.3	0.49	0.54	0.69
	1-τ	0.81	0.9	0.81	0.92	0.66	0.64	0.69	1.04	0.63	0.72	1.18	0.97	0.7	1	0.81	0.72	0.67
	Fuzzy Distance	3.26	4.35	3.28	4.36	3.98	3.83	3.16	4.31	5.04	3.05	6.30	4.86	3.56	4.18	3.44	3.20	3.24
GDD5	Median Bias	-124.11	52.29	109.06	-340.42	-485.53	-410.81	-86.07	-90.72	-705.57	-312.7	393.01	-446.33	-274.62	-127.72	-592.37	-481.68	-365.75
	IQR Ratio	0.56	0.64	0.4	0.43	0.69	0.41	0.6	0.7	0.6	0.49	0.53	0.32	0.67	0.64	1.12	0.7	0.55
	1-τ	0.6	0.82	0.74	0.58	0.63	0.64	0.61	0.79	0.63	0.59	0.74	0.65	0.67	0.68	0.73	0.66	0.62
	Fuzzy Distance	357.33	633.98	464.76	507.93	509.07	528.69	463.43	478.09	621.61	467.06	671.54	610.52	399.02	479.12	637.89	521.39	546.36
MAP	Median Bias	123.03	89.11	54.85	334.24	171.3	172.36	141.2	143.17	202.87	132.44	106.81	90.07	109.85	33.53	50.88	40.32	9.74
	IQR Ratio	0.75	0.92	0.54	0.84	1.16	0.55	0.86	0.52	1.1	0.68	0.48	0.83	0.83	0.83	0.76	0.78	0.9
	1-τ	1.08	1.01	0.92	1.12	1.11	1.01	1.05	1.16	1.14	1.07	0.86	0.93	0.95	0.86	0.89	0.84	0.97
	Fuzzy Distance	199.35	174.60	149.05	387.37	205.58	211.32	203.95	208.64	228.00	206.18	135.91	162.74	219.73	169.63	188.58	166.17	209.67
α	Median Bias	0.02	0.04	-0.02	0.04	0.08	0.07	0.06	0.05	0.09	0.01	0.02	0.02	0.05	-0.01	0.01	0	-0.07
	IQR Ratio	2.11	3.07	2.29	2.54	2.82	1.55	2.77	2.05	3.45	2.34	2.18	4.19	2.6	2.6	3.19	2.85	3
	1-τ	0.77	0.76	0.83	0.81	0.83	0.72	0.79	0.86	0.88	0.75	0.74	0.74	0.79	0.81	0.81	0.73	0.83
	Fuzzy Distance	0.11	0.10	0.10	0.11	0.12	0.10	0.13	0.11	0.17	0.10	0.09	0.13	0.12	0.13	0.13	0.13	0.12

Table S10. Summary of the metrics for the simulations for Mid-Holocene (MH) for each of the individual models.

		PMIP2 OA										PMIP2 OA			PMIP2 OAV			
Variable	Metric	CCSM	CSIRO3.1.0	CSIRO3.1.1	ECBILT.LV	ECHAM	FGOALS1	FOAM	GISS.E	IPSL4	MIROC	MRI2fa	MRI2	HadCM3	ECBILT.LV (OAV)	ECHAM (OAV)	FOAM (OAV)	MRI2fa (OAV)
SSTann	Median Bias	-0.36	-0.17	-0.12	-0.22	NA	-0.19	-0.14	-0.53	-0.23	-0.46	-0.32	-0.59	-0.31	-0.12	-0.34	-0.21	-0.17
	IQR Ratio	0.37	0.47	0.5	0.24	NA	0.4	0.31	0.56	0.46	0.33	0.3	0.25	0.35	0.33	0.22	0.22	0.33
	1–τ	0.89	0.95	1.01	0.87	NA	0.98	1.02	0.9	1.07	0.96	0.93	0.78	1.12	0.86	0.82	0.88	0.89
	Fuzzy Distance	0.56	0.59	0.53	0.48	NA	0.57	0.55	0.74	0.55	0.61	0.58	0.75	0.50	0.45	0.48	0.59	0.53
MAT	Median Bias	-0.14	0.47	0.46	0.1	-0.02	-0.15	0.5	-0.01	0.74	-0.13	0	-0.55	0.04	0.33	0.02	0.19	0.13
	IQR Ratio	0.17	0.23	0.26	0.16	0.16	0.18	0.25	0.29	0.17	0.2	0.21	0.14	0.2	0.17	0.16	0.37	0.22
	1–τ	0.97	0.94	0.9	0.91	0.93	1.07	0.95	0.91	0.87	0.96	0.88	0.99	0.99	0.9	0.96	0.92	0.87
	Fuzzy Distance	1.16	1.13	1.11	1.09	1.10	1.21	1.22	1.12	1.22	1.14	1.10	1.28	1.15	1.11	1.13	1.22	1.05
MTCO	Median Bias	-0.31	0.4	0.71	-0.03	-0.13	-0.49	0.15	-0.43	0.48	-0.6	-0.1	-0.82	-0.08	0.35	0.13	-0.05	-0.08
	IQR Ratio	0.26	0.29	0.32	0.14	0.38	0.26	0.2	0.28	0.3	0.29	0.2	0.27	0.25	0.15	0.16	0.34	0.28
	1–τ	1.05	0.93	0.86	0.97	1.04	1.04	0.95	0.92	0.89	1.09	0.96	1	0.98	0.91	0.95	0.95	1.01
	Fuzzy Distance	1.46	1.51	1.44	1.40	1.47	1.61	1.43	1.46	1.43	1.59	1.43	1.70	1.44	1.36	1.38	1.64	1.36
MTWA	Median Bias	0.98	1.44	1.41	0.56	0.99	1.76	1.8	1.02	2	1.31	1.32	1.17	1.55	0.78	0.83	1.51	1.55
	IQR Ratio	0.3	0.4	0.39	0.38	0.29	0.55	0.48	0.44	0.28	0.36	0.27	0.32	0.39	0.37	0.28	0.39	0.32
	1–τ	1.09	1.07	1.05	1.01	1.12	1.11	1.06	1.05	1.06	1.09	0.99	1.01	1.14	1	1.12	1.04	1.02
	Fuzzy Distance	1.49	1.87	1.79	1.36	1.55	2.23	2.18	1.65	2.17	1.61	1.63	1.57	1.98	1.53	1.49	1.89	1.85
GDD5	Median Bias	-24.87	48.6	54.15	-23.9	-10	21.83	87.12	4.93	119.16	8.74	8.14	-46.24	28.96	20.2	-22.26	49.08	37.7
	IQR Ratio	0.15	0.28	0.29	0.17	0.21	0.27	0.26	0.32	0.19	0.21	0.19	0.26	0.19	0.16	0.18	0.3	0.23
	1–τ	0.98	1.02	0.99	0.92	1.09	1.19	1.01	1.02	0.91	1.04	0.86	1.07	1.09	0.89	1.08	0.96	0.94
	Fuzzy Distance	195.91	196.91	199.06	199.27	210.70	228.74	210.50	198.14	228.13	199.60	184.10	206.22	200.66	203.82	205.81	199.38	191.77
MAP	Median Bias	-52.45	-47.96	-42.63	-43.76	-38.47	-38.75	-47.46	-44.3	-36.38	-54.24	-57.48	-64.06	-43.27	-36.26	-41.53	-44.51	-53.42
	IQR Ratio	0.2	0.24	0.2	0.23	0.31	0.27	0.23	0.29	0.24	0.26	0.23	0.23	0.22	0.26	0.23	0.28	0.28
	1–τ	0.95	1.07	1.07	1.03	0.77	0.94	1.1	0.98	0.92	0.89	0.89	0.84	0.96	0.91	0.8	1.03	0.88
	Fuzzy Distance	92.99	98.62	86.63	95.02	76.03	82.74	93.73	94.23	81.53	91.75	92.75	92.35	85.29	84.01	77.54	95.40	87.92
α	Median Bias	-0.03	-0.04	-0.04	-0.02	-0.01	-0.02	-0.03	-0.04	-0.04	-0.04	-0.03	-0.04	-0.03	-0.01	-0.02	-0.02	-0.03
	IQR Ratio	0.5	0.94	0.84	0.57	0.52	0.63	0.82	0.84	0.76	0.75	0.77	0.74	0.8	0.6	0.48	0.8	0.85
	1–τ	1	1.03	1	0.99	0.88	0.92	1.09	0.99	0.86	0.96	0.96	0.93	0.98	0.91	0.88	1.06	0.92
	Fuzzy Distance	0.05	0.06	0.06	0.05	0.04	0.05	0.06	0.06	0.05	0.06	0.05	0.06	0.06	0.04	0.04	0.06	0.06

Table S10 continued.

		PMIP2 OAV		CMIP5 OA										CMIP5 OAC				
Variable	Metric	MRI2 (OAV)	HadCM3 (OAV)	CCSM4	CNRM5	CSIRO3.6	EARTH	FGOALSG 2	FGOALS2	GISS.E2	MPI (ESM)	MRI3	CSIRO3.1. 2	BCC	HadGEM2 (CC)	HadGEM2 (ESM)	IPSL5	MIROC (ESM)
SSTann	Median Bias	-0.07	-0.31	-0.39	-0.09	-0.21	NA	NA	NA	-0.28	-0.3	-0.27	-0.24	-0.41	-0.06	-0.08	-0.26	-0.59
	IQR Ratio	0.34	0.38	0.17	0.48	0.42	NA	NA	NA	0.4	0.45	0.46	0.33	0.41	0.42	0.49	0.45	0.26
	1–Tau	0.9	0.98	0.93	0.98	0.97	NA	NA	NA	0.84	0.84	0.86	1.03	0.94	0.89	1.03	0.9	0.95
	Fuzzy Distance	0.52	0.58	0.60	0.55	0.57	NA	NA	NA	0.59	0.57	0.52	0.63	0.64	0.49	0.56	0.52	0.71
MAT	Median Bias	0.56	0.27	-0.51	0.48	0.2	0.01	-0.75	0	-0.07	0.01	0.04	0.19	0.01	0.6	0.64	0.03	-0.55
	IQR Ratio	0.46	0.2	0.21	0.21	0.24	0.25	0.18	0.21	0.23	0.16	0.14	0.16	0.24	0.35	0.28	0.16	0.27
	1–Tau	0.94	0.9	0.91	0.88	0.9	0.94	0.94	0.93	0.9	0.95	0.93	0.93	0.92	0.91	0.93	0.91	0.94
	Fuzzy Distance	1.42	1.09	1.18	1.23	1.05	1.11	1.38	1.11	1.09	1.12	1.09	1.08	1.10	1.23	1.20	1.10	1.28
MTCO	Median Bias	0.4	0.15	-1.07	0.76	0.07	0.21	-0.91	-0.05	0.21	-0.12	-0.26	0.15	-0.18	0.76	0.51	-0.21	-0.79
	IQR Ratio	0.3	0.32	0.31	0.3	0.34	0.33	0.15	0.34	0.24	0.25	0.22	0.19	0.31	0.56	0.31	0.29	0.31
	1–Tau	0.95	0.9	1.1	0.87	0.92	0.86	0.99	0.89	0.91	0.92	1.08	0.94	0.95	0.92	0.95	0.96	1.06
	Fuzzy Distance	1.60	1.34	1.91	1.81	1.38	1.68	1.65	1.34	1.40	1.43	1.50	1.37	1.44	1.81	1.51	1.41	1.69
MTWA	Median Bias	2.27	1.57	1.24	1.69	1.38	0.92	0.6	1.12	0.78	1.08	1.08	1.27	1.28	0.97	1.79	1.28	0.82
	IQR Ratio	0.54	0.44	0.35	0.22	0.5	0.41	0.3	0.47	0.32	0.31	0.29	0.42	0.45	0.37	0.52	0.3	0.34
	1–Tau	1	1.03	1.04	1.06	1.05	1.02	1.11	1.07	1.08	1.03	1.18	1.05	1.05	0.97	0.93	0.99	1.06
	Fuzzy Distance	2.51	2.02	1.59	1.98	1.99	1.48	1.35	1.61	1.33	1.56	1.59	1.75	1.80	1.57	2.20	1.59	1.46
GDD5	Median Bias	148.31	55.19	-25.98	83.41	33.65	-37.33	-106.45	-15.68	-15.39	-2.43	3.77	22.46	5.51	59.4	88.77	19.2	-72.07
	IQR Ratio	0.57	0.2	0.21	0.17	0.29	0.27	0.26	0.22	0.31	0.22	0.15	0.23	0.25	0.34	0.36	0.18	0.27
	1–Tau	1.01	0.97	0.98	0.95	0.99	1	0.98	1	0.93	1.07	1.15	1.05	1	0.88	0.9	0.89	0.97
	Fuzzy Distance	278.94	193.92	209.11	229.71	200.83	190.37	223.19	201.54	199.29	201.84	209.63	195.33	197.42	204.64	231.52	188.24	215.77
MAP	Median Bias	-32.12	-44.79	-68.97	-40.32	-59.81	-47.27	-83.09	-46.92	-39.03	-45.81	-49.04	-46.56	-54.61	-48.66	-46.37	-44.54	-74.01
	IQR Ratio	0.38	0.29	0.26	0.24	0.26	0.27	0.21	0.19	0.38	0.35	0.22	0.23	0.17	0.39	0.35	0.23	0.34
	1–Tau	0.87	1.01	0.87	0.82	1	0.89	0.89	0.94	0.91	0.82	0.82	1.04	1.06	0.87	0.87	0.95	0.93
	Fuzzy Distance	76.57	94.75	101.22	84.20	99.70	90.40	109.71	83.18	84.46	80.80	84.60	93.57	96.91	87.95	86.15	85.76	99.27
α	Median Bias	-0.02	-0.03	-0.04	-0.02	-0.05	-0.03	-0.04	-0.03	-0.02	-0.03	-0.03	-0.04	-0.03	-0.03	-0.04	-0.03	-0.05
	IQR Ratio	1.02	0.94	0.86	0.66	1.15	0.81	0.7	0.52	0.78	0.82	0.58	0.9	0.67	0.89	1.3	0.71	0.9
	1–Tau	0.91	1	0.93	0.87	1	0.94	0.95	0.95	0.85	0.82	0.98	1	1.01	0.98	0.97	0.95	1
	Fuzzy Distance	0.05	0.05	0.06	0.04	0.07	0.05	0.06	0.05	0.04	0.05	0.05	0.06	0.05	0.06	0.06	0.05	0.07