A comprehensive history of climate and habitat stability of the last 800,000 years

Mario Krapp[1], Robert Beyer[1], Stephen L. Edmundson[1,2], Paul J. Valdes[3], and Andrea Manica[1]

[1]Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, United Kingdom
[2]Department of Earth Sciences, Utrecht University, Budapestlaan 4, 3584 CD Utrecht, The Netherlands
[3]School of Geographical Sciences, University of Bristol, BS8 1SS Bristol, United Kingdom

Correspondence: Mario Krapp (mariokrapp@gmail.com)

Abstract. A detailed and accurate reconstruction of the past climate is essential in understanding the interactions between ecosystems and their environment through time. We know that climatic drivers have shaped the distribution and evolution of species, including our own, and their habitats. Yet, spatially-detailed climate reconstructions that continuously cover the Quaternary do not exist. This is mainly because no paleoclimate model can reconstruct regional-scale dynamics over geolog-

5   ical time scales. Here we develop a statistical emulator, the Global Climate Model Emulator (GCMET), which reconstructs the climate of the last 800,000 years with unprecedented spatial detail. GCMET captures the temporal dynamics of glacial-interglacial climates as an Earth System Model of Intermediate Complexity would whilst resolving the local dynamics with the accuracy of a Global Climate Model. It provides a new, unique resource to explore the climate of the Quaternary, which we use to investigate the long-term stability of major habitat types. We identify a number of stable pockets of habitat that have

10   remained unchanged over the last 800 thousand years, acting as potential long-term evolutionary refugia. Thus, the highly detailed, comprehensive overview of climatic changes through time delivered by GCMET provides the needed resolution to quantify the role of long term habitat change and fragmentation in an ecological and anthropological context.

1   Introduction

Current patterns of diversification within and between species, such as our own (Scerri et al., 2018), and the structuring of

15   whole ecosystems can only be studied in the context of past climatic changes that have shaped them through time (Doebeli and Dieckmann, 2003). A detailed understanding of such processes has become an urgent necessity in order to predict responses to global change. However, whilst predictions of climate change and their impacts over the next few tens or hundreds of years are based on comprehensive Global Climate Models (GCMs) that resolve processes at high temporal and spatial resolution, such as those used in the latest IPCC Assessment Report (Solomon et al., 2007), reconstructions back in time are challenging

20   as they have to span a much longer period. GCMs can provide snapshots for a specific time or short transients in the order of a few thousands of years, whilst periods of tens or hundreds of thousands of years can only be covered with Earth System Models of Intermediate Complexity (EMICs) (Ganopolski and Calov, 2011; Timmermann et al., 2013), at the cost of low spatial resolution and a simplified representation of the climate system (Claussen et al., 2002). Neither of those two types of

models is intentionally designed for paleo-ecology or species evolution, disciplines that require appropriate temporal scales of

25 up to hundreds of thousands of years and spatial scales down to tens of kilometres.

Here, we fill this gap for a long-term reconstruction of climate that resolves regional-scale dynamics by reconstructing the last 800 thousand years (ka) at an unprecedented spatial resolution of approximately 1°. Unlike previous emulator approaches which aimed at capturing climate dynamics at a global level (Araya-Melo et al., 2015; Lord et al., 2017), we explicitly focus on the local emulation of climate. We critically evaluate our reconstructed 800 ka of climate history against proxy records,

30 demonstrating that our Global Climate Model Emulator (GCMET) provides a high quality reconstruction, equivalent to that provided by global climate model simulations, for the last 800 ka.

The aim of this paper is twofold: 1) We present the technical details of our emulator approach (Sect. 2), which, in this setup, can be thought of as an extension of the HadCM3 snapshot simulations into the deeper past. We validate GCMET (as an emulator of HadCM3) with longer-term Quaternary climate proxies (Sect. 3). 2) We regard the resulting reconstructed climate

35 history of the last 800 ka as a high-value data set relevant for a wide range of applications in research areas that deal with long-term past climate changes. As a case study, we reconstructed ecosystems throughout the last 800 ka (Sect. 4) and analysed the long-term stability of human habitats through time in terms of their spatial fragmentation.

2 Methods and data

Our emulator approach consists of two steps (Fig. 1): a reconstruction of the global climate at moderate spatial resolution

40 followed by a more detailed representation of local dynamics using multiple snapshot simulations from the family of HadCM3 climate models (Valdes et al., 2017). In the first step, we use 72 simulations covering the past 120 ka from the HadCM3 climate model (Singarayer and Valdes, 2010; Davies-Barnard et al., 2017), with a resolution of 3° (2.5° × 3.75°), and build a set of linear regression models that are the basis of GCMET. In the second step, we increase the resolution of our GCMET reconstructions to about 1° (1.25°×0.83°) using high resolution HadAM3H (Hadley Centre Atmospheric Model 3, High reso-

45 lution) simulations covering the period of the last deglaciation. To do so we computed high-resolution difference maps between equivalent HadAM3H and GCMET simulations.

2.1 HadCM3 and HadAM3H

HadCM3 is a fully coupled global climate model with an atmospheric component, HadAM3, which has a horizontal resolution of 2.5° × 3.75°, with 19 vertical levels, and a time step of 30 min. The ocean and sea-ice component of HadCM3 has a

50 horizontal resolution of 1.25° × 1.25° and 20 vertical levels. In this paper, we use 72 available HadCM3 simulations covering the last 120,000 years in 2,000-year intervals from 120,000 to 24,000 ka before present (BP) and in 1,000-year intervals from 22,000 to present-day (Singarayer and Valdes, 2010; Davies-Barnard et al., 2017) (https://www.paleo.bristol.ac.uk/ummodel/ data/tdwza/standard_html/tdwza.html, last accessed on 05 Oct 2018).

HadAM3H is the higher resolution version of the atmosphere model HadAM3. It has a horizontal resolution of 1.25° by

55 0.83° with 30 vertical levels and a time step of 10 min. HadAM3H uses the surface fluxes from the associated fully coupled

Climate
of the Past

Discussions

HadCM3 simulations. HadAM3H is available for the time since the last deglaciation, 21 ka BP, i.e., 21, 18, 15, 12, 10, 8, 6, 3, and 0 ka BP.

2.2 The global climate model emulator GCMET

GCMET relies on a set of linear regression models, one for each individual time series of every HadCM3 model grid box
60   with the following independent variables: atmospheric CO_2 concentrations (as a major greenhouse gas), and three variables reflecting the orbital forcing (Berger and Loutre, 1991). These are based on obliquity ε and two combinations of eccentricity e and precession ω: $e\sin\omega$, henceforth referred to as precession index I, and $e\cos\omega$ (precession index II). These are a generally accepted set of orbital forcings as they reflect insolation at any location and any time (Araya-Melo et al., 2015; Lord et al., 2017). Atmospheric CO_2 concentrations are the same as in the respective HadCM3 time slice simulation, e.g., 280 ppmv for
65   0 ka BP.

Our approach interpolates the HadCM3 model output at each grid box, i.e., a time series, to the parameter settings for which HadCM3 has been evaluated, i.e., the time series of CO_2 concentrations and the three orbital parameters. This is unlike state-of-the-art emulators that extrapolate to settings which have not been evaluated. However, it ensures that the predictions by GCMET are well behaved. As a consequence, GCMET is only valid within the boundary conditions for which it has been
70   parameterised; in our case, we build valid regression coefficients for the Quaternary. These limitation could be easily alleviated by any additional snapshot simulation with a new, expanded parameter setting.

The dependent variables, i.e, the predictands, are climate variables such as temperature T, precipitation P, or specific humidity Q. The independent variables, i.e., the predictors, are applied as normalised forcings. Thus, the resulting regression coefficients, denoted as β coefficients, can be compared across different climate variables, i.e., temperature and precipitation,
75   and across each other (Fig. 2).

Let us assume $Y(x,y,t)$ is a time series of a climate variable at a specific location (x,y) at time t. To explain variations of Y around a mean value $\overline{Y}$, i.e, $Y' = Y - \overline{Y}$, we run a multiple linear regression model for the anomalies Y':

$$Y'(x,y,t) = \underbrace{\beta_\varepsilon(x,y)\varepsilon'(t) + \beta_{e\cos\omega}(x,y)(e\cos\omega)'(t) + \beta_{e\cos\omega}(x,y)(e\cos\omega)'(t)}_{\text{orbital forcing}}$$

$$+ \underbrace{\beta_{CO_2}(x,y)CO_2'(t)}_{\text{greenhouse gas forcing}} + \underbrace{\beta_M(x,y)M(x,y,t)}_{\text{variable land–sea mask effect}} \tag{1}$$

80   Here, the βs are the regression coefficients for the respective predictor (see Fig. 2 for maps of β coefficients). We also consider changes in surface type, i.e., ocean (0), land (1), and ice (2) which have been masked as $M(x,y,t) \in [0,1,2]$. For example, around coastlines, land grid boxes can turn into ocean grid boxes when sea level is high. Similarly, expanding ice sheets turn land grid boxes into ice-covered grid boxes, and the climate variable Y may respond to different surface types in different ways. To make the linear regression statistically well-conditioned, all independent variables have been normalised, i.e.,
85   the mean has been subtracted and the data has then been divided by their standard deviation. To prevent our linear regression

model from predicting negative precipitation values, we apply a logarithmic transformation which maps values from $[0,+\infty]$ to $[-\infty,+\infty]$. For bounded variables such as precipitation this is a common procedure. In the case of precipitation, the linear regression coefficients predict the response in terms of anomalies in the exponent.

The units of specific humidity are [kg/kg] and its values fall in the range between 0 and 1. For this reason, we transform specific humidity using the logit function, $\mathrm{logit}(x) = \log(\frac{x}{1-x})$, which maps values from [0,1] to $[-\infty,+\infty]$. The decomposition of temperature T, precipitation P, and specific humidity Q into anomalies, i.e., the Y' on the left hand side of Eq. 1 is:

$$T = \overline{T} + \underbrace{T'}_{\triangleq Y'} \tag{2}$$

$$\log(P) = \overline{\log(P)} + \underbrace{\log(P)'}_{\triangleq Y'} \tag{3}$$

$$\mathrm{logit}(Q) = \overline{\mathrm{logit}(Q)} + \underbrace{\mathrm{logit}(Q)'}_{\triangleq Y'} \tag{4}$$

In contrast to existing emulator approaches (Araya-Melo et al., 2015; Lord et al., 2017; Rangel et al., 2018), our reconstructions are local-scale reconstructions which are in reasonable agreement with existing paleo-climate proxies as shown in our comprehensive model–data comparison below. Furthermore, because the parameter sampling is based on realistic glacial cycle snapshot simulations, the obtained regression coefficients are good enough approximations to predict previous Quaternary glacial–interglacial climate states well.

2.2.1 Training and test data

In order to make useful predictions and to evaluate the skill of our model, we divide the HadCM3 snapshots into and training and a test data set. A sensible choice is to use 80% of the HadCM3 snapshots to train the linear regression model and 20% to test it. For a 80/20 division of the 72 time slices into training and test data, i.e., 14 or 58 out of 72, there are $\binom{n}{k} = \binom{72}{58} \approx 3 \times 10^{14}$ possible combinations. But instead of randomly dividing the data into the training/test data, we follow an approach with the aim to preserve as much variance as possible in the training data, i.e. maximise the variance of the predictors. This is best illustrated by the phase plots of the predictors (Fig. 3). The training data set covers the edges of each phase plane and thus maximises the phase space covered by the linear regression model. This choice of training data ensures that the linear regression model interpolates within the phase space and does not need to extrapolate for the test data.

The procedure is as follows. We calculate the covariance matrix of the full parameter set (n=72) and its eigenvalues. Then, we randomly create a training data set (k=58) for which we compute the covariance. If the covariance of this sample training set is larger than the full covariance matrix, i.e., the eigenvalues of the covariance matrix are larger than the eigenvalues of the covariance matrix of the full parameter set, this sample parameter set is marked as a candidate for the final training set. After several iterations (N=10,000), we sum up how many times each time slice has appeared within a candidate training set. We then rank all time slices according to this number. In the final step, we pick the 80% top-ranked time slices as training data.

2.2.2 Validation

For the validation of GCMET, we use the proportion of variance in the response explained by the regression (R^2, a goodness of fit estimator of the training data; Fig. 4), and the root mean squared error (RMSE), an estimator of the goodness of the model for the prediction of the test data. Overall, our linear model is a better predictor for temperature than for precipitation.

Temperature responds more directly to local forcings because it is determined by the energy balance of downward and upward longwave and shortwave radiation and turbulent heat fluxes. The downward shortwave radiation depends on incoming solar radiation that is determined by orbital variations, whereas downward longwave radiation is determined by greenhouse gases such as CO_2 and water vapour, as well as cloud cover. Large-scale atmospheric circulation changes have a much smaller effect on temperature. Locally, it is therefore locally far better constrained by global CO_2 and orbital variations. This increases the predictive skill of our linear regression model substantially resulting in higher R^2 values and lower RMSEs.

The matter is more complicated for precipitation because it is a product of the hydrological cycle, which itself depends mainly on large-scale atmospheric dynamics, such as monsoonal systems in the tropics and subtropics, or mid-latitude storm systems. Local interactions between the atmosphere and the surface, such as evaporation and transpiration over the ocean, or deep convection over the tropics, matter to a lesser extent. Processes and circulation features like moisture transport or the atmospheric Hadley cell dynamics determine the non-local response of precipitation to CO_2 or orbital variations to a much larger extent. Because of the larger dynamical component of the hydrological cycle, compared to temperature, precipitation is much less constrained by external forcings than temperature. Thus, the linear regression model has less predictive skill for precipitation than for temperature. However, we find that the predictive skill for precipitation can be improved by using temperature and specific humidity as predictors instead of the orbital parameters and CO_2. In this way, the RMSEs can be substantially reduced, especially over land (Fig. 4E).

For the improved precipitation model we used temperature T and specific humidity Q as independent variables, i.e.,

$$\log(P)' = \beta_T T' + \beta_Q \mathrm{logit}(Q)'. \tag{5}$$

For the prediction of climate before 120 ka BP this implies that we first need to reconstruct T and Q, and then use the obtained β coefficients for T and Q to reconstruct P.

2.2.3 The regression coefficients

To assess how reliable our predictor estimates are, we calculate the p-values for each of the predictors, i.e., the significance of our β coefficients. Here, the p-value tests the null hypothesis whether the coefficient is equal to zero, i.e., the specific predictor has no effect. If the p-value is below a certain threshold—in our case below the 5% significance level: $p < 0.05$—the null hypothesis can be rejected. That means that the specific predictor is a meaningful addition to our linear regression model and any changes in the associated predictor are related to changes in the corresponding climate variable. Regions for which the null hypothesis cannot be rejected are displayed as shaded and hatched in Fig. 2.

2.3 Dynamic downscaling to higher resolution

Using nine high-resolution HadAM3H simulations covering the deglaciation since 21 ka BP (21, 18, 15, 12, 10, 8, 6, 3, and 0 ka BP), we are able to increase the spatial resolution from $3°$, which is the spatial resolution of GCMET after the linear regression step (and the same as the coarse resolution of the original HadCM3 snapshots), to ca. $1°$. We do so by calculating

150   the difference between equivalent coarse- and high-resolution snapshots. For example, the difference at 10 ka BP is $\Delta_{10\,ka\,BP} = HadAM3H_{10\,ka\,BP} - HadCM3_{10\,ka\,BP}$. We choose to interpolate the differences linearly according to their CO_2 levels, e.g., 231 ppm at 10 ka BP, because any statistical model with more than one variable would require more snapshots to adequately predict the differences. Thus, we simply assume that the differences between a coarse- and high-resolution climate can be explained as a function of the CO_2 forcing, i.e., $\Delta_{10\,ka\,BP} = \Delta_{231\,ppm}$. For any period in the past, e.g., 300 ka BP, we add the

155   high-resolution difference, i.e., the Δ, which corresponds to the respective CO_2 level, to the coarse-resolution reconstruction. Note that the downscaling approach captures the regional-scale dynamics of the GCM in this step, which change over time. This is in contrast to the commonly used "delta method" for downscaling of climate model data which assumes a constant difference between simulated and observed data.

     To illustrate the importance of higher spatial variability, we compared the high-resolution version of GCMET, the origi-

160   nal resolution version denoted as GCMET-LO, and LOVECLIM, an EMIC with a horizontal resolution of ca $5.5°\times5.5°$, to present-day observations (ERA-20C re-analysis 1961–1990 average (Poli et al., 2016)). Due to their lower spatial resolution LOVECLIM and GCMET-LO cannot capture the observed continental climate patterns, whereas GCMET resolves those spatial features well (Fig. 5B).

2.4 Boundary conditions: CO_2, global sea-level, and Northern Hemisphere ice sheets

165   For realistic high-resolution reconstructions the model boundary conditions for the last 800 ka need to be known: atmospheric CO_2 levels and orbital parameters (Fig. 3A), global sea level (for the land-sea mask), and the extent of Northern Hemisphere (NH) ice sheets. The longest, quasi-continuous record of past CO_2 levels is the 800,000 years long CO_2 record from the EPICA Dome C ice core in Antarctica (Bereiter et al., 2015). The orbital parameters are from the same data set as before (Berger and Loutre, 1991). For ice sheet extents, we use model outputs from CLIMBER-2/SICOPOLIS simulations (Ganopolski and Calov,

170   2011) for which NH ice sheet extents and heights are available for the last 800 ka in 1 ka-year intervals, from which we use the period from 800–123 ka BP . For 122–0 ka BP, we use the ice sheet configurations from the ICE-6G data set (Peltier et al., 2014) (http://www.atmosp.physics.utoronto.ca/~peltier/data.php, last accessed 09.11.2018). Changes in the coast lines affecting the land–sea mask are derived from a global sea-level record (Spratt and Lisiecki, 2016) which has been added on top of present-day coast lines while preserving inland lakes.

3 Comparing GCMET to climate proxies

Despite the increasing number of available paleoclimate proxies, only a small fraction can be used for a quantitative comparison to climate models as translating sediment core data into actual climate variables remains a difficult task. Marine sediment cores are valuable archives of past sea surface temperature (SST) records. Because their associated biogeochemistry is relatively straightforward, marine proxies can be utilised as so-called paleo-thermometers and are thus well suited for a direct proxy–model comparison. For these proxies, we make a direct comparison between mean annual temperature (MAT) and SST, quantified both in terms of correlation between the predicted and observed time series and the RMSE. We note that MAT and SST are not the same climatological quantities; SST is the temperature of the ocean surface and has a lower limit of about -1.8°C, the freezing point of saltwater. While we expect MAT and SST to co-vary in low and mid-latitudes, at higher latitudes, seasonal or perennial sea ice cover could make a comparison between both variables problematic.

For terrestrial proxies, for which a translation into climatic variables is not straightforward, we simply quantify the correlation between the two time series. However, the interpretation of terrestrial proxies from a climate perspective can also be problematic. For example, pollen-based vegetation reconstructions are suggested to be less reliable as climate proxies, particularly for interglacials (Herzschuh et al., 2016). Other land-based proxies such as dust deposits integrate long-term climatic changes over large regions and hence do not necessarily capture climatic effects at their specific location.

For the comparison of GCMET against proxy reconstruction, we assembled long-term marine SST and terrestrial climate proxy reconstructions (Figs. 7,8) which cover a period of at least 150 ka BP during the last 800 ka (Tables 1 and 2). We assessed the goodness of the long-term GCMET climate reconstructions by cross-comparing those SST proxy reconstructions to the MAT reconstruction by GCMET and to surface temperature data from the LOVECLIM climate model (Timmermann et al., 2013) (http://apdrc.soest.hawaii.edu/projects/paleomodeling/800k.php, last accessed 08/11/2018).

Reconstructions by GCMET are in good agreement with a number of marine records (Fig. 7B-E), with a mean RMSE of 1.5 K for all SST proxies and a mean correlation of 0.5. GCMET captures long-term temporal dynamics of glacial–interglacial climates and its performance in that respect is on average as good as LOVECLIM's (r=0.5, Fig. 5B). Despite the diverse nature of the terrestrial proxies (e.g. speleothems, loess, pollen), GCMET performance was as good as for marine proxies (r=0.5, Fig. 5B and Fig. 8B,C).

4 Ecosystem reconstructions and habitat stability over the last 800 ka

In the final part of this paper, we highlight the potential of a comprehensive, long-term climate data set. As an example we investigate ecosystem stability (Fig. 9) over the last 800 ka, focusing on 14 major terrestrial habitats defined by the WWF *Global 200* (Olson et al., 2001) (Fig. 9A). The motivation for this is that habitat fragmentation, which is closely related to ecosystem (in)stabilities, may have played a key role in the evolution of our species, *homo sapiens* (Scerri et al., 2018).

We use a random forest classifier (Breiman, 2001) that is trained on a set of four climate variables from GCMET (minimum and mean annual temperature, and minimum and mean annual precipitation) to reconstruct the present-day distribution of the 14 ecoregions. The required present-day data was split into a training (80%) and a test data set (20%). The classification factors,

i.e., the analogue of predictors in a linear model, from this training data set were then applied to predict ecosystem changes of the last 800,000 years.

210   The goodness of the predictions by the random forest classifier can be estimated by the so-called receiver operating characteristic (ROC, Fig. 9D). A ROC curve displays the true positive rate against the false positive rate and the closer that curve is to the upper left corner, the better the prediction for a specific ecosystem is. For example, the point at coordinate (0,1) represents the best possible prediction with 100% sensitivity (i.e., no false negatives) and 100% specificity (i.e., no false positives). The diagonal line corresponds to a prediction by random guessing.

215   The random forest classification is very close to perfect classification for the average of all ecosystem types and the area under the curve is an estimator for the goodness of the classification (numbers are given in legend of Fig. 9). Except for a few instances, such as for "Tropical & Subtropical Coniferous Forests" and "Mangroves", values are always larger than 0.9 (average 0.98).

The spatially detailed reconstructions provided by GCMET allow us to explore the effect of climate on habitats over time.

220   As can be seen in Fig. 9C, stability depends on location, with sparsely vegetated regions such as deserts among the most stable habitats in the world, the others being the core tropical rainforests along the equator. Large parts of Eurasia and North America are rendered unstable by the advancing and retreating NH ice sheets with ecosystems alternating between vast forests during the warm interglacials and large tundras during the cold glacials (an animated version of the habitat changes throughout the last 800 ka is available as *Supplementary Video*). However, a few fragmented core boreal forest habitats remain.

225   **5   Summary and discussion**

The global climate model emulator, GCMET, is an effective and computationally efficient tool to reconstruct the climate of the past over long periods while based on only a relatively small number of paleoclimate simulations. In contrast to other emulators, our grid box by grid box approach assumes that the spatial correlations are sufficiently well preserved by HadCM3 and thus are well represented in the GCMET reconstructions as can be seen in the β coefficient maps in Fig. 2. The glacial and interglacial

230   climate states of the Quaternary with its respective glacial and interglacial boundary conditions are an ideal testbed to show how well GCMET emulates the HadCM3 climate response without explicitly running any HadCM3 simulation for earlier times. GCMET can in principle be applied to any other global climate model for which a reasonable number of simulations exists.

Our approach has several limitations. For example, we are unable to account for climate variability on time scales that are not simulated in the underlying HadCM3 snapshots. Originally, these have been set up to examine climate changes on

235   orbital timescales and they are not fully transient (Singarayer and Valdes, 2010). This has implications for the ecosystem reconstructions and the inferred habitat stability because we can only account for climate and ecosystem variability on time scales longer than 1000 years and more. However, the assumption of a equilibrium climate in GCMET is not too far fetched given the relatively good level of agreement with climate proxies, most of which have a temporal resolution of more than 1000 years, and as such can be thought of as reflecting the average long-term rather than an instantaneous climatic state.

240   Another limitation is that we cannot link the regression coefficients for CO_2 and the orbital parameters to true geophysical processes. GCMET is a purely statistical model and as such any physical interpretation of β coefficients would be flawed. However, the imposed CO_2 forcing from the Antarctic ice core with its distinct late Quaternary, 100,000 year cyclicity already accounts for the orbital effects on the Earth system. A different but interesting question would be whether it is possible to disentangle orbital changes from the natural variations in CO_2; however, this is beyond the scope of this paper.

245   Because GCMET is a statistical emulator, we can use it to assess the uncertainty of the predictors (i.e., boundary conditions) and their respective coefficient estimates. We did this for the global mean temperature comparison (Fig. 5A) but the uncertainty in the CO_2 forcing is fairly small as are the standard errors for the temperature regression (not shown), and thus, the uncertainty is hardly visible (blue shaded). In this specific example we also see how much GCMET depends on the underlying model performance. GCMET temperature reconstructions closely follow the HadCM3 curve for the last 120 ka simulations but differ

250   from the proxy record during certain periods of the past. For example, HadCM3 underestimates the global mean temperature 120 ka ago compared to the proxy record and so does GCMET. For earlier periods, GCMET can be used to detect discrepancies between the GCM and long-term paleo records as we did in the model–proxy comparison in Sect. 3. The performance of an emulator strongly depends on the representation of the climate system in the underlying GCM: Realistic climate reconstructions rely on both a realistic climate model and a good statistical model, i.e., emulator.

255 **6   Conclusions**

A major advantage of GCMET is that it is computationally inexpensive. GCMET can produce high-quality reconstructions of the last 800 ka that compare well with proxy records of the past. This is the equivalent of hundreds of GCM snapshots, a prohibitive endeavour for the foreseeable future. A way to understand the fit of GCMET predictions against climate proxy time series is that our approach captures the quasi-equilibrium state of the climate system, thus allowing us to efficiently describe

260   the behaviour over the longer, millennial, time scales. In turns, this implies that the glacial-interglacial climate of the Middle and Late Pleistocene responded in a consistent manner to orbital forcings and CO_2. It will be interesting to test whether this approximation is also valid for the Early Pleistocene with its faster ice age cyclicity of 41 ka; for this endeavour, we currently lack quasi-continuous CO_2 estimates before the Mid-Pleistocene Transition, however GCMET is fully capable of covering the appropriate time periods if estimates became available. For the moment, we can offer a detailed, coherent reconstruction of the

265   past 800 thousand years, which allowed us to pinpoint long-term potential refugia that have been characterised by the same habitat, and we expect that this will open up new ways to study the impact of past climate in a number of disciplines such as ecology and anthropology.

Code and data availability.   The generated climate reconstructions for the last 800,000 years are publicly available in the project's *Open Science Framework* repository at this address: https://bit.ly/2XWrGvF. This data set comes in two version, one is the output after the dynamical

270   downscaling step and one that has been bias corrected afterwards using the ERA-20C climatology from 1961–1990 (Poli et al., 2016). The

variables listed in Table 3 are available in 1000 year intervals and in 1°, the HadAM3H resolution, and 1.125° horizontal, i.e, the ERA-20C grid resolution under the project's `data` directory. The individual proxy time series and their respective GCMET counterpart as shown in Fig. 7 and 8 are available as Excel spreadsheet (`output_150ka.xlsx`) in the same directory.

Model code for GCMET as well as the code for the analysis and visualisation of figures is also publicly available at the same address

275 under the project's `code` directory (`gcmet.tar.gz` and required input files `inputs.tar.gz`). NetCDF files have been processed using `cdo` (Schulzweida, 2019). The linear regression for GCMET is based on the Python library `statsmodels` (Seabold and Perktold, 2010). The random forest classifier has been implemented using the Python library `scikit-learn` (Pedregosa et al., 2011). All visualisations are made with `matplotlib` (Hunter, 2007) and `cartopy` (Met Office, 2010).

Video supplement. Movies of the habitat evolution over the last 800,000 years can be found in the *Open Science Framework* repository

280 (https://bit.ly/2XWrGvF) under the project's `movies` directory. The 14 major WWF habitats (`ecosystems_800ka.mp4`) have been aggregated into the four categories: snow and ice, barren and sparsely vegetated, open habitat, and forests (`habitats_800ka.mp4`). The temporal evolution of the atmospheric CO_2 concentrations are shown below the map.

Author contributions. AM and MK devised the project. MK devised and implemented the emulator with input from RB and SLE. PJV provided additional *HadCM3* snapshot simulations. MK and AM wrote a first draft of the paper which was improved by input from all other

285 authors. MK prepared the figures and videos.

Competing interests. The authors declare no conflict of interest.

References

290 Alonso-Garcia, M., Sierro, F., Kucera, M., Flores, J., Cacho, I., and Andersen, N.: Ocean circulation, ice sheet growth and interhemispheric coupling of millennial climate variability during the mid-Pleistocene (ca 800–400ka), Quaternary Science Reviews, 30, 3234–3247, https://doi.org/10.1016/j.quascirev.2011.08.005, http://linkinghub.elsevier.com/retrieve/pii/S0277379111002435, 2011.

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

295 Bar-Matthews, M., Ayalon, A., Gilmour, M., Matthews, A., and Hawkesworth, C. J.: Sea–land oxygen isotopic relationships from planktonic foraminifera and speleothems in the Eastern Mediterranean region and their implication for paleorainfall during interglacial intervals, Geochimica et Cosmochimica Acta, 67, 3181–3199, https://doi.org/10.1016/S0016-7037(02)01031-1, http://www.sciencedirect.com/science/article/pii/S0016703702010311, 2003.

Bard, E., Rostek, F., and Sonzogni, C.: Interhemispheric synchrony of the last deglaciation inferred from alkenone palaeothermometry, Nature, 385, 707–710, https://doi.org/10.1038/385707a0, https://www.researchgate.net/publication/246503195_Interhemispheric_synchrony_of_the_last_deglaciation_inferred_from_alkenone_palaeothermometry, 1997.

Beck, J. W., Zhou, W., Li, C., Wu, Z., White, L., Xian, F., Kong, X., and An, Z.: A 550,000-year record of East Asian monsoon rainfall from 10Be in loess, Science, 360, 877–881, https://doi.org/10.1126/science.aam5825, http://science.sciencemag.org/content/360/6391/877, 2018.

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

Berger, A. and Loutre, M. F.: Insolation values for the climate of the last 10 million years, Quaternary Science Reviews, 10, 297–317, https://doi.org/10.1016/0277-3791(91)90033-Q, http://www.sciencedirect.com/science/article/pii/027737919190033Q, 1991.

310 Breiman, L.: Random Forests, Machine Learning, 45, 5–32, https://doi.org/10.1023/A:1010933404324, https://doi.org/10.1023/A:1010933404324, 2001.

Caley, T., Kim, J.-H., Malaizé, B., Giraudeau, J., Laepple, T., Caillon, N., Charlier, K., Rebaubier, H., Rossignol, L., Castañeda, I. S., Schouten, S., and Sinninghe Damsté, J. S.: High-latitude obliquity as a dominant forcing in the Agulhas current system, Clim. Past, 7, 1285–1296, https://doi.org/10.5194/cp-7-1285-2011, https://www.clim-past.net/7/1285/2011/, 2011.

315 Carolin, S. A., Cobb, K. M., Lynch-Stieglitz, J., Moerman, J. W., Partin, J. W., Lejau, S., Malang, J., Clark, B., Tuen, A. A., and Adkins, J. F.: Northern Borneo stalagmite records reveal West Pacific hydroclimate across MIS 5 and 6, Earth and Planetary Science Letters, 439, 182–193, https://doi.org/10.1016/j.epsl.2016.01.028, http://linkinghub.elsevier.com/retrieve/pii/S0012821X16000509, 2016.

Cheng, H., Zhang, P. Z., Spötl, C., Edwards, R. L., Cai, Y. J., Zhang, D. Z., Sang, W. C., Tan, M., and An, Z. S.: The climatic cyclicity in semiarid-arid central Asia over the past 500,000 years, Geophysical Research Letters, 39, https://doi.org/10.1029/2011GL050202, 320 https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011GL050202, 2012.

Cheng, H., Edwards, R. L., Sinha, A., Spötl, C., Yi, L., Chen, S., Kelly, M., Kathayat, G., Wang, X., Li, X., Kong, X., Wang, Y., Ning, Y., and Zhang, H.: The Asian monsoon over the past 640,000 years and ice age terminations, Nature, 534, 640–646, https://doi.org/10.1038/nature18591, http://www.nature.com/nature/journal/v534/n7609/full/nature18591.html?foxtrotcallback=true, 00000, 2016.

325 Claussen, M., Mysak, L., Weaver, A., Crucifix, M., Fichefet, T., Loutre, M.-F., Weber, S., Alcamo, J., Alexeev, V., Berger, A., Calov, R., Ganopolski, A., Goosse, H., Lohmann, G., Lunkeit, F., Mokhov, I., Petoukhov, V., Stone, P., and Wang, Z.: Earth system models of intermediate complexity: closing the gap in the spectrum of climate system models, Climate Dynamics, 18, 579–586, https://doi.org/10.1007/s00382-001-0200-1, https://link.springer.com/article/10.1007/s00382-001-0200-1, 2002.

Crundwell, M., Scott, G., Naish, T., and Carter, L.: Glacial–interglacial ocean climate variability from planktonic foraminifera during the
330 Mid-Pleistocene transition in the temperate Southwest Pacific, ODP Site 1123, Palaeogeography, Palaeoclimatology, Palaeoecology, 260, 202–229, https://doi.org/10.1016/j.palaeo.2007.08.023, http://www.sciencedirect.com/science/article/pii/S0031018207005974, 2008.

Davies-Barnard, T., Ridgwell, A., Singarayer, J., and Valdes, P.: Quantifying the influence of the terrestrial biosphere on glacial–interglacial climate dynamics, Clim. Past, 13, 1381–1401, https://doi.org/10.5194/cp-13-1381-2017, https://www.clim-past.net/13/1381/2017/, 2017.

Doebeli, M. and Dieckmann, U.: Speciation along environmental gradients, Nature, 421, 259–264, https://doi.org/10.1038/nature01274,
335 https://www.nature.com/articles/nature01274, 2003.

Dyez, K. A. and Ravelo, A. C.: Late Pleistocene tropical Pacific temperature sensitivity to radiative greenhouse gas forcing, Geological Society of America, 41, 23–26, https://doi.org/10.1130/G33425.1, https://pubs.geoscienceworld.org/geology/article/41/1/23-26/130996, 2013.

Dyez, K. A., Ravelo, A. C., and Mix, A. C.: Evaluating drivers of Pleistocene eastern tropical Pacific sea surface temperature, Paleoceanog-
340 raphy, 31, 2015PA002 873, https://doi.org/10.1002/2015PA002873, http://onlinelibrary.wiley.com/doi/10.1002/2015PA002873/abstract, 2016.

Etourneau, J., Martinez, P., Blanz, T., and Schneider, R.: Pliocene–Pleistocene variability of upwelling activity, productivity, and nutrient cycling in the Benguela region, Geology, 37, 871–874, https://doi.org/10.1130/G25733A.1, https://pubs.geoscienceworld.org/gsa/geology/article/37/10/871/103830/pliocene-pleistocene-variability-of-upwelling, 2009.

345 Ganopolski, A. and Calov, R.: The role of orbital forcing, carbon dioxide and regolith in 100 kyr glacial cycles, Clim. Past, 7, 1415–1425, https://doi.org/10.5194/cp-7-1415-2011, http://www.clim-past.net/7/1415/2011/, 2011.

Garidel-Thoron, T. d., Rosenthal, Y., Beaufort, L., Bard, E., Sonzogni, C., and Mix, A. C.: A multiproxy assessment of the western equatorial Pacific hydrography during the last 30 kyr, Paleoceanography, 22, https://doi.org/10.1029/2006PA001269, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006PA001269, 2005.

350 Guo, Z. T., Berger, A., Yin, Q. Z., and Qin, L.: Strong asymmetry of hemispheric climates during MIS-13 inferred from correlating China loess and Antarctica ice records, Clim. Past, 5, 21–31, https://doi.org/10.5194/cp-5-21-2009, https://www.clim-past.net/5/21/2009/, 2009.

Hao, Q., Wang, L., Oldfield, F., Peng, S., Qin, L., Song, Y., Xu, B., Qiao, Y., Bloemendal, J., and Guo, Z.: Delayed build-up of Arctic ice sheets during 400,000-year minima in insolation variability, Nature, 490, 393–396, https://doi.org/10.1038/nature11493, http://www.nature.com/articles/nature11493, 2012.

355 Hayward, B. W., Scott, G. H., Crundwell, M. P., Kennett, J. P., Carter, L., Neil, H. L., Sabaa, A. T., Wilson, K., Rodger, J. S., Schaefer, G., Grenfell, H. R., and Li, Q.: The effect of submerged plateaux on Pleistocene gyral circulation and sea-surface temperatures in the Southwest Pacific, Global and Planetary Change, 63, 309–316, https://doi.org/10.1016/j.gloplacha.2008.07.003, http://www.sciencedirect.com/science/article/pii/S0921818108000775, 2008.

Hayward, B. W., Sabaa, A. T., Kolodziej, A., Crundwell, M. P., Steph, S., Scott, G. H., Neil, H. L., Bostock, H. C., Carter, L., and Grenfell,
360 H. R.: Planktic foraminifera-based sea-surface temperature record in the Tasman Sea and history of the Subtropical Front around New Zealand, over the last one million years, Marine Micropaleontology, 82-83, 13–27, https://doi.org/10.1016/j.marmicro.2011.10.003, http://www.sciencedirect.com/science/article/pii/S0377839811001150, 2012.

Climate
of the Past
Discussions

Herbert, T. D.: Collapse of the California Current During Glacial Maxima Linked to Climate Change on Land, Science, 293, 71–76, https://doi.org/10.1126/science.1059209, http://www.sciencemag.org/cgi/doi/10.1126/science.1059209, 2001.

365 Herbert, T. D., Peterson, L. C., Lawrence, K. T., and Liu, Z.: Tropical Ocean Temperatures Over the Past 3.5 Million Years, Science, 328, 1530–1534, https://doi.org/10.1126/science.1185435, http://www.sciencemag.org/cgi/doi/10.1126/science.1185435, 2010.

Herzschuh, U., Birks, H. J. B., Laepple, T., Andreev, A., Melles, M., and Brigham-Grette, J.: Glacial legacies on interglacial vegetation at the Pliocene-Pleistocene transition in NE Asia, Nature Communications, 7, 11 967, https://doi.org/10.1038/ncomms11967, https://www.nature.com/articles/ncomms11967, 2016.

370 Ho, S. L., Mollenhauer, G., Lamy, F., Martínez-Garcia, A., Mohtadi, M., Gersonde, R., Hebbeln, D., Nunez-Ricardo, S., Rosell-Melé, A., and Tiedemann, R.: Sea surface temperature variability in the Pacific sector of the Southern Ocean over the past 700 kyr, Paleoceanography, 27, https://doi.org/10.1029/2012PA002317, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012PA002317, 2012.

Horikawa, K., Murayama, M., Minagawa, M., Kato, Y., and Sagawa, T.: Latitudinal and downcore (0–750 ka) changes in n-alkane chain lengths in the eastern equatorial Pacific, Quaternary Research, 73, 573–582, https://doi.org/10.1016/j.yqres.2010.01.001, http://www.

375 sciencedirect.com/science/article/pii/S0033589410000025, 2010.

Hunter, J. D.: Matplotlib: A 2D graphics environment, Computing In Science & Engineering, 9, 90–95, https://doi.org/10.1109/MCSE.2007.55, 2007.

Jouzel, J., Masson-Delmotte, V., Cattani, O., Dreyfus, G., Falourd, S., Hoffmann, G., Minster, B., Nouet, J., Barnola, J. M., Chappellaz, J., Fischer, H., Gallet, J. C., Johnsen, S., Leuenberger, M., Loulergue, L., Luethi, D., Oerter, H., Parrenin, F., Raisbeck, G., Raynaud,

380 D., Schilt, A., Schwander, J., Selmo, E., Souchez, R., Spahni, R., Stauffer, B., Steffensen, J. P., Stenni, B., Stocker, T. F., Tison, J. L., Werner, M., and Wolff, E. W.: Orbital and Millennial Antarctic Climate Variability over the Past 800,000 Years, Science, 317, 793–796, https://doi.org/10.1126/science.1141038, http://science.sciencemag.org/content/317/5839/793, 2007.

K. Thomas, E., C. Clemens, S., Sun, Y., L. Prell, W., Huang, Y., Gao, L., Loomis, S., Chen, G., and Liu, Z.: Heterodynes dominate precipitation isotopes in the East Asian monsoon region, reflecting interaction of multiple climate factors, Earth and Planetary Science Letters,

385 455, https://doi.org/10.1016/j.epsl.2016.09.044, 2016.

Kathayat, G., Cheng, H., Sinha, A., Spötl, C., Edwards, R. L., Zhang, H., Li, X., Yi, L., Ning, Y., Cai, Y., Lui, W. L., and Breitenbach, S. F. M.: Indian monsoon variability on millennial-orbital timescales, Scientific Reports, 6, 24 374, https://doi.org/10.1038/srep24374, https://www.nature.com/articles/srep24374, 2016.

Landwehr, J. M., Sharp, W. D., Coplen, T. B., Ludwig, K. R., and Winograd, I. J.: The Chronology for the δ^{18}O Record from

390 Devils Hole, Nevada, Extended Into the Mid-Holocene, Tech. rep., US Geological Survey, https://pubs.usgs.gov/of/2011/1082/, 2011.

Lawrence, K. T., Herbert, T. D., Brown, C. M., Raymo, M. E., and Haywood, A. M.: High-amplitude variations in North Atlantic sea surface temperature during the early Pliocene warm period, Paleoceanography, 24, PA2218, https://doi.org/10.1029/2008PA001669, http://adsabs.harvard.edu/abs/2009PalOc..24.2218L, 2009.

Li, L., Li, Q., Tian, J., Wang, P., Wang, H., and Liu, Z.: A 4-Ma record of thermal evolution in the tropical western Pacific and its implications

395 on climate change, Earth and Planetary Science Letters, 309, 10–20, https://doi.org/10.1016/j.epsl.2011.04.016, http://www.sciencedirect.com/science/article/pii/S0012821X11002329, 2011.

Liu, Z.: Pleistocene climate evolution in the eastern Pacific and implications for the orbital theory of climate change, Ph.D., Brown University, United States – Rhode Island, https://search.proquest.com/docview/305227845/abstract/5386AB9B03324943PQ/1, 2004.

Liu, Z., Altabet, M. A., and Herbert, T. D.: Glacial-interglacial modulation of eastern tropical North Pacific denitrification over the last
400   1.8-Myr, Geophysical Research Letters, 32, https://doi.org/10.1029/2005GL024439, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.
        1029/2005GL024439, 2005.

Lord, N. S., Crucifix, M., Lunt, D. J., Thorne, M. C., Bounceur, N., Dowsett, H., O'Brien, C. L., and Ridgwell, A.: Emula-
        tion of long-term changes in global climate: application to the late Pliocene and future, Climate of the Past, 13, 1539–1571,
        https://doi.org/https://doi.org/10.5194/cp-13-1539-2017, https://www.clim-past.net/13/1539/2017/cp-13-1539-2017.html, 2017.

405 Martrat, B., Grimalt, J. O., Shackleton, N. J., Abreu, L. d., Hutterli, M. A., and Stocker, T. F.: Four Climate Cycles of Recurring Deep and
        Surface Water Destabilizations on the Iberian Margin, Science, 317, 502–507, https://doi.org/10.1126/science.1139994, http://science.
        sciencemag.org/content/317/5837/502, 2007.

Martínez-Garcia, A., Rosell-Melé, A., McClymont, E. L., Gersonde, R., and Haug, G. H.: Subpolar Link to the Emergence of the Mod-
        ern Equatorial Pacific Cold Tongue, Science, 328, 1550–1553, https://doi.org/10.1126/science.1184480, http://science.sciencemag.org/
410   content/328/5985/1550, 2010.

Martínez-Garcia, A., Rosell-Melé, A., Geibert, W., Gersonde, R., Masqué, P., Gaspari, V., and Barbante, C.: Links between iron supply,
        marine productivity, sea surface temperature, and CO2 over the last 1.1 Ma, Paleoceanography, 24, https://doi.org/10.1029/2008PA001657,
        https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008PA001657, 2009.

McClymont, E. L., Rosell-Melé, A., Giraudeau, J., Pierre, C., and Lloyd, J. M.: Alkenone and coccolith records of the mid-Pleistocene
415   in the south-east Atlantic: Implications for the U37K' index and South African climate, Quaternary Science Reviews, 24, 1559–1572,
        https://doi.org/10.1016/j.quascirev.2004.06.024, http://www.sciencedirect.com/science/article/pii/S0277379105000806, 2005.

Medina-Elizalde, M. and W Lea, D.: The Mid-Pleistocene Transition in the Tropical Pacific, Science (New York, N.Y.), 310, 1009–12,
        https://doi.org/10.1126/science.1115933, 2005.

Melles, M., Brigham-Grette, J., Minyuk, P. S., Nowaczyk, N. R., Wennrich, V., DeConto, R. M., Anderson, P. M., Andreev, A. A., Coletti,
420   A., Cook, T. L., Haltia-Hovi, E., Kukkonen, M., Lozhkin, A. V., Rosén, P., Tarasov, P., Vogel, H., and Wagner, B.: 2.8 Million Years
        of Arctic Climate Change from Lake El'gygytgyn, NE Russia, Science, 337, 315–320, https://doi.org/10.1126/science.1222135, http:
        //science.sciencemag.org/content/337/6092/315, 2012.

Met Office: Cartopy: a cartographic python library with a Matplotlib interface, http://scitools.org.uk/cartopy, 2010.

Nürnberg, D. and Groeneveld, J.: Pleistocene variability of the Subtropical Convergence at East Tasman Plateau: Evidence from plank-
425   tonic foraminiferal Mg/Ca (ODP Site 1172A), Geochemistry, Geophysics, Geosystems, 7, https://doi.org/10.1029/2005GC000984,
        https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005GC000984, 2006.

Nürnberg, D., Müller, A., and Schneider, R. R.: Paleo-sea surface temperature calculations in the equatorial east Atlantic from Mg/Ca ratios
        in planktic foraminifera: A comparison to sea surface temperature estimates from U37K', oxygen isotopes, and foraminiferal transfer
        function, Paleoceanography and Paleoclimatology, 15, 124–134, https://doi.org/10.1029/1999PA000370, https://agupubs.onlinelibrary.
430   wiley.com/doi/abs/10.1029/1999PA000370, 2000.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'amico, J. A., Itoua, I., Strand,
        H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., and Kassem,
        K. R.: Terrestrial Ecoregions of the World: A New Map of Life on EarthA new global map of terrestrial ecoregions provides an innovative
        tool for conserving biodiversity, BioScience, 51, 933–938, https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2, https:
435   //academic.oup.com/bioscience/article/51/11/933/227116, 2001.

Climate
of the Past
Discussions

Pahnke, K., Zahn, R., Elderfield, H., and Schulz, M.: 340,000-Year Centennial-Scale Marine Record of Southern Hemisphere Climatic Oscillation, Science, 301, 948–952, https://doi.org/10.1126/science.1084451, http://science.sciencemag.org/content/301/5635/948, 2003.

Past Interglacials Working Group of PAGES: Interglacials of the last 800,000 years, Rev. Geophys., 54, 2015RG000 482, https://doi.org/10.1002/2015RG000482, http://onlinelibrary.wiley.com/doi/10.1002/2015RG000482/abstract, 2016.

440 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Peltier, W. R., Argus, D. F., and Drummond, R.: Space geodesy constrains ice age terminal deglaciation: The global ICE-6G_C (VM5a) model, Journal of Geophysical Research: Solid Earth, 120, 450–487, https://doi.org/10.1002/2014JB011176, https://agupubs.onlinelibrary.
445 wiley.com/doi/full/10.1002/2014JB011176, 2014.

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G. H., Peubey, C., Thépaut, J.-N., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L., and Fisher, M.: ERA-20C: An Atmospheric Reanalysis of the Twentieth Century, J. Climate, 29, 4083–4097, https://doi.org/10.1175/JCLI-D-15-0556.1, https://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-15-0556.1, 2016.

Prokopenko, A. A., Hinnov, L. A., Williams, D. F., and Kuzmin, M. I.: Orbital forcing of continental climate during the Pleis-
450 tocene: a complete astronomically tuned climatic record from Lake Baikal, SE Siberia, Quaternary Science Reviews, 25, 3431–3457, https://doi.org/10.1016/j.quascirev.2006.10.002, http://www.sciencedirect.com/science/article/pii/S0277379106002812, 2006.

Rangel, T. F., Edwards, N. R., Holden, P. B., Diniz-Filho, J. A. F., Gosling, W. D., Coelho, M. T. P., Cassemiro, F. A. S., Rahbek, C., and Colwell, R. K.: Modeling the ecology and evolution of biodiversity: Biogeographical cradles, museums, and graves, Science, 361, eaar5452, https://doi.org/10.1126/science.aar5452, http://science.sciencemag.org/content/361/6399/eaar5452, 2018.

455 Rincón-Martínez, D. and Leduc, G.: Sea surface temperature calculated from alkenones for the last 285 ka with high-reolution Holocene of sediment core MD02-2529, Panama Basin, https://doi.org/https://doi.org/10.1594/PANGAEA.777473, https://doi.pangaea.de/10.1594/PANGAEA.777473, type: dataset, 2012.

Rodrigues, T., Voelker, A. H. L., Grimalt, J. O., Abrantes, F., and Naughton, F.: Iberian Margin sea surface temperature dur-ing MIS 15 to 9 (580–300 ka): Glacial suborbital variability versus interglacial stability, Paleoceanography, 26, PA1204,
460 https://doi.org/10.1029/2010PA001927, http://onlinelibrary.wiley.com/doi/10.1029/2010PA001927/abstract, 2011.

Rostek, F., Ruhlandt, G., Bassinot, F. C., Muller, P. J., Labeyrie, L. D., Lancelot, Y., and Bard, E.: Reconstructing sea surface temperature and salinity using $\delta^{18}O$ and alkenone records, Nature, 364, 319, https://doi.org/10.1038/364319a0, https://www.nature.com/articles/364319a0, 1993.

Ruddiman, W. F., Raymo, M. E., Martinson, D. G., Clement, B. M., and Backman, J.: Pleistocene evolution: Northern hemisphere ice
465 sheets and North Atlantic Ocean, Paleoceanography and Paleoclimatology, 4, 353–412, https://doi.org/10.1029/PA004i004p00353, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/PA004i004p00353, 1989.

Russon, T., Elliot, M., Sadekov, A., Cabioch, G., Corrège, T., and Deckker, P. D.: Inter-hemispheric asymmetry in the early Pleistocene Pacific warm pool, Geophysical Research Letters, 37, https://doi.org/10.1029/2010GL043191, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010GL043191, 2010.

470 Scerri, E. M. L., Thomas, M. G., Manica, A., Gunz, P., Stock, J. T., Stringer, C., Grove, M., Groucutt, H. S., Timmermann, A., Rightmire, G. P., d'Errico, F., Tryon, C. A., Drake, N. A., Brooks, A. S., Dennell, R. W., Durbin, R., Henn, B. M., Lee-Thorp, J., deMenocal, P., Petraglia, M. D., Thompson, J. C., Scally, A., and Chikhi, L.: Did Our Species Evolve in Subdivided Populations across Africa, and Why

15

Does It Matter?, Trends in Ecology & Evolution, 33, 582–594, https://doi.org/10.1016/j.tree.2018.05.005, http://www.sciencedirect.com/science/article/pii/S0169534718301174, 2018.

475 Schaefer, G., Rodger, J. S., Hayward, B. W., Kennett, J. P., Sabaa, A. T., and Scott, G. H.: Planktic foraminiferal and sea surface temperature record during the last 1 Myr across the Subtropical Front, Southwest Pacific, Marine Micropaleontology, 54, 191–212, https://doi.org/10.1016/j.marmicro.2004.12.001, http://linkinghub.elsevier.com/retrieve/pii/S0377839804001288, 2005.

Schefuß, E., Damsté, J. S. S., and Jansen, J. H. F.: Forcing of tropical Atlantic sea surface temperatures during the mid-Pleistocene transition, Paleoceanography, 19, https://doi.org/10.1029/2003PA000892, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2003PA000892, 480 2004.

Schmidt, M. W., Vautravers, M. J., and Spero, H. J.: Western Caribbean sea surface temperatures during the late Quaternary, Geochemistry, Geophysics, Geosystems, 7, https://doi.org/10.1029/2005GC000957, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005GC000957, 2006.

Schulzweida, U.: CDO User Guide, https://doi.org/10.5281/zenodo.2558193, https://doi.org/10.5281/zenodo.2558193, 2019.

485 Seabold, S. and Perktold, J.: Statsmodels: Econometric and statistical modeling with python, in: 9th Python in Science Conference, 2010.

Singarayer, J. S. and Valdes, P. J.: High-latitude climate sensitivity to ice-sheet forcing over the last 120kyr, Quaternary Science Reviews, 29, 43–55, https://doi.org/10.1016/j.quascirev.2009.10.011, 00092, 2010.

Snyder, C. W.: Evolution of global temperature over the past two million years, Nature, 538, 226–228, https://doi.org/10.1038/nature19798, https://www.nature.com/articles/nature19798, 2016.

490 Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Avery, K., Tignor, M., and Miller, H., eds.: IPCC: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2007.

Spratt, R. M. and Lisiecki, L. E.: A Late Pleistocene sea level stack, Clim. Past, 12, 1079–1092, https://doi.org/10.5194/cp-12-1079-2016, http://www.clim-past.net/12/1079/2016/, 2016.

495 Stoykova, D. A., Shopov, Y. Y., Garbeva, D., Tsankov, L. T., and Yonge, C. J.: Origin of the climatic cycles from orbital to sub-annual scales, Journal of Atmospheric and Solar-Terrestrial Physics, 70, 293–302, https://doi.org/10.1016/j.jastp.2007.08.018, http://adsabs.harvard.edu/abs/2008JASTP..70..293S, 2008.

Tierney, J. E., deMenocal, P. B., and Zander, P. D.: A climatic context for the out-of-Africa migration, The Geological Society of America, 45, 1023–1026, https://doi.org/10.1130/G39457.1, http://pubs.geoscienceworld.org/geology/article/45/11/1023/516677/A-climatic-context-for-the-outofAfrica-migration, 2017.

Timmermann, A., Friedrich, T., Timm, O. E., Chikamoto, M. O., Abe-Ouchi, A., and Ganopolski, A.: Modeling Obliquity and CO2 Effects on Southern Hemisphere Climate during the Past 408 ka, J. Climate, 27, 1863–1875, https://doi.org/10.1175/JCLI-D-13-00311.1, http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-13-00311.1, 2013.

Tzedakis, P. C., Hooghiemstra, H., and Pälike, H.: The last 1.35 million years at Tenaghi Philippon: revised chronostratigraphy and 505 long-term vegetation trends, Quaternary Science Reviews, 25, 3416–3430, https://doi.org/10.1016/j.quascirev.2006.09.002, http://www.sciencedirect.com/science/article/pii/S0277379106002708, 2006.

Vaks, A., Bar-Matthews, M., Ayalon, A., Matthews, A., Frumkin, A., Dayan, U., Halicz, L., Almogi-Labin, A., and Schilman, B.: Paleoclimate and location of the border between Mediterranean climate region and the Saharo–Arabian Desert as revealed by speleothems from the northern Negev Desert, Israel, Earth and Planetary Science Letters, 249, 384–399, https://doi.org/10.1016/j.epsl.2006.07.009, 510 http://www.sciencedirect.com/science/article/pii/S0012821X06005000, 2006.

Vaks, A., Bar-Matthews, M., Matthews, A., Ayalon, A., and Frumkin, A.: Middle-Late Quaternary paleoclimate of northern margins of the Saharan-Arabian Desert: reconstruction from speleothems of Negev Desert, Israel, Quaternary Science Reviews, 29, 2647–2662, https://doi.org/10.1016/j.quascirev.2010.06.014, http://www.sciencedirect.com/science/article/pii/S0277379110002088, 2010.

Valdes, P. J., Armstrong, E., Badger, M. P. S., Bradshaw, C. D., Bragg, F., Crucifix, M., Davies-Barnard, T., Day, J. J., Farnsworth, A.,
515   Gordon, C., Hopcroft, P. O., Kennedy, A. T., Lord, N. S., Lunt, D. J., Marzocchi, A., Parry, L. M., Pope, V., Roberts, W. H. G., Stone, E. J., Tourte, G. J. L., and Williams, J. H. T.: The BRIDGE HadCM3 family of climate models: HadCM3@Bristol v1.0, Geosci. Model Dev., 10, 3715–3743, https://doi.org/10.5194/gmd-10-3715-2017, https://www.geosci-model-dev.net/10/3715/2017/, 2017.

Waldmann, N., Torfstein, A., and Stein, M.: Northward intrusions of low- and mid-latitude storms across the Saharo-Arabian belt during past interglacials, Geology, 38, 567–570, https://doi.org/10.1130/G30654.1, https://pubs.geoscienceworld.org/gsa/geology/article/38/
520   6/567/130337/northward-intrusions-of-low-and-mid-latitude, 2010.

Yamamoto, M., Yamamuro, M., and Tanaka, Y.: The California current system during the last 136,000 years: response of the North Pacific High to precessional forcing, Quaternary Science Reviews, 26, 405–414, https://doi.org/10.1016/j.quascirev.2006.07.014, http://linkinghub.elsevier.com/retrieve/pii/S0277379106002435, 2007.

Table 1. Marine proxy records that have been used in the validation of the climate reconstruction, their coordinates, correlation coefficients, types, and respective references.

core/name	lon	lat	corr coeff	type	reference(s)
DSDP 594	175.0	-45.5	0.60	SST	PIGS, Schaefer et al. 2005
DSDP 607	-33.0	41.0	0.46	SST	PIGS, Ruddiman et al. 1989
GeoB 1105	-12.4	-1.7	0.62	SST	Nürnberg et al. 2000
GeoB 1112	-10.7	-5.8	0.53	SST	Nürnberg et al. 2000
HY04	-95.0	4.0	0.30	SST	PIGS, Horikawa et al. 2010
MD01-2444	-10.1	37.6	0.71	SST	Martrat et al. 2007
MD02-2529	-84.1	8.2	0.53	SST	Rincón-Martínez and Leduc 2012
MD03-2699	-10.7	39.0	0.61	SST	Rodrigues et al. 2011
MD06-2986	167.9	-43.4	0.70	SST	PIGS, Hayward et al. 2012
MD06-3018	166.2	-22.6	0.38	SST	Russon et al. 2010
MD85-668	46.0	0.0	0.48	SST	Bard et al. 1997
MD90-963	73.9	5.1	0.50	SST	Rostek et al. 1993
MD96-2048	36.0	-26.2	0.52	SST	Caley et al. 2011
MD97-2120	174.9	-45.5	0.73	SST	Pahnke et al. 2003
MD97-2140	141.5	2.0	0.55	SST	PIGS, Garidel-Thoron et al. 2005
ODP 1012	-118.4	32.3	0.59	SST	PIGS, Liu et al. 2005
ODP 1014	-118.9	32.8	0.75	SST	Yamamoto et al. 2007
ODP 1020	-126.4	41.0	0.50	SST	PIGS, Herbert 2001
ODP 1077b	10.4	-5.2	0.18	SST	Schefuß et al. 2004
ODP 1082	11.8	-21.1	0.43	SST	Etourneau et al. 2009
ODP 1087	15.3	-31.5	0.13	SST	McClymont et al. 2005
ODP 1090	8.9	-42.9	0.69	SST	PIGS, Martínez-Garcia et al. 2009
ODP 1123	-171.5	-41.8	0.35	SST	PIGS, Crundwell et al. 2008
ODP 1125	-178.2	-42.6	0.55	SST	Hayward et al. 2008
ODP 1143	113.3	9.4	0.61	SST	PIGS, Li et al. 2011
ODP 1146	116.3	19.5	0.53	SST	PIGS, Herbert et al. 2010
ODP 1172	149.9	-44.0	0.28	SST	Nürnberg and Groeneveld 2006
ODP 1239	-82.1	-0.7	0.53	SST	Dyez et al. 2016
ODP 306	-27.9	56.4	0.33	SST	Alonso-Garcia et al. 2011
ODP 722	59.8	16.6	0.45	SST	PIGS, Herbert et al. 2010
ODP 806b	159.4	0.3	0.55	SST	PIGS, Medina-Elizalde and W Lea 2005
ODP 846	-90.8	-3.1	0.49	SST	PIGS, Liu 2004
ODP 871	172.3	5.6	0.63	SST	Dyez and Ravelo 2013
ODP 882	167.6	50.4	0.11	SST	Martínez-Garcia et al. 2010
ODP 977A	0.0	37.5	0.63	SST	Martrat et al. 2007
ODP 982	-15.9	57.5	0.35	SST	PIGS, Lawrence et al. 2009
ODP 999	-78.7	12.8	0.21	SST	Schmidt et al. 2006
PS75034-2	-80.1	-54.4	0.79	SST	PIGS, Ho et al. 2012
RC09-166	48.8	12.5	0.35	SST	Tierney et al. 2017

PIGS refers to Past Interglacials Working Group of PAGES (2016)

Table 2. Terrestrial proxy records that have been used in the validation of the climate reconstruction, their coordinates, correlation coefficients, types, and respective references.

core/name	lon	lat	corr coeff	type	reference(s)
Baoji, China	107.1	34.4	0.62	rainfall	Beck et al. 2018
Bittoo	77.8	30.8	-0.43	$\delta^{18}O$	Kathayat et al. 2016
Chanwu	107.7	35.2	0.58	$\delta^{18}O$	Guo et al. 2009
Clearwater	114.9	4.1	-0.47	$\delta^{18}O$	Carolin et al. 2016
Dead Sea	35.0	30.5	-0.68	lake level	Waldmann et al. 2010
Devil's Hole	-116.3	36.4	0.67	$\delta^{18}O$	Landwehr et al. 2011
Duhlata	23.2	42.5	0.41	ODL	Stoykova et al. 2008
EPICA Dome C	123.4	-75.0	0.91	temperature	Jouzel et al. 2007
Kesang	81.8	42.9	-0.15	$\delta^{18}O$	Cheng et al. 2012
Lake Baikal	108.4	53.7	-0.16	Bio. sil.	Prokopenko et al. 2006
Lake El'gygytgyn	172.0	67.5	0.17	mag. susc.	Melles et al. 2012
Negev	34.8	30.6	-0.61	$\delta^{18}O$	Vaks et al. 2010
Peqiin	36.0	32.6	-0.67	$\delta^{18}O$	Bar-Matthews et al. 2003
Sanbao-Dongge	110.4	31.7	0.12	$\delta^{18}O$	Cheng et al. 2016
Soreq	36.0	31.4	-0.56	$\delta^{18}O$	Bar-Matthews et al. 2003
Tenaghi Philippon	24.2	41.0	0.65	arb. pollen	PIGS, Tzedakis et al. 2006
Tzavoa	35.2	31.2	-0.37	$\delta^{18}O$	Vaks et al. 2006
Weinan	109.6	34.4	0.48	temperature	K. Thomas et al. 2016
Xifeng loess	107.6	35.7	0.60	Fed/Fet	Guo et al. 2009
Yimaguan Luochuan	108.5	35.8	0.60	mag. susc.	PIGS, Hao et al. 2012

PIGS refers to Past Interglacials Working Group of PAGES (2016)

Table 3. List of file names for GCMET regression results and climate reconstructions of the last 800 ka as well as the Excel spreadsheet for the proxy–model comaprison. This data can be found in the *Open Science Framework* repository https://bit.ly/2XWrGvF under the project's `data` directory.

	variable	file name
regression results	MAT	`hadcm3_000-120_ann_regression_temp_co2-ecospre-esinpre-obl.nc`
	MAP	`hadcm3_000-120_ann_regression_prec_co2-ecospre-esinpre-obl.nc`
	MAQ	`hadcm3_000-120_ann_regression_shum_co2-ecospre-esinpre-obl.nc`
	MAP (MAT and MAQ)	`hadcm3_000-120_ann_regression_prec_shum-temp.nc`
	MINT	`hadcm3_000-120_min_regression_temp_co2-ecospre-esinpre-obl.nc`
	MINQ	`hadcm3_000-120_min_regression_shum_co2-ecospre-esinpre-obl.nc`
	MINP (MINT and MINQ)	`hadcm3_000-120_min_regression_prec_shum-temp.nc`
reconstructions	MAT	`temp_800ka_ann_hi_nobc.nc`
	MINT	`temp_800ka_min_hi_nobc.nc`
	MAP	`prec_800ka_ann_hi_nobc.nc`
	MINP	`prec_800ka_min_hi_nobc.nc`
	WWF 14 major habitats	`ecosystems_random_forests_wwf_tmean_tmin_pmean_pmin.nc`
	MAT (BC)	`temp_800ka_ann_hi_20c_bc.nc`
	MINT (BC)	`temp_800ka_min_hi_20c_bc.nc`
	MAP (BC)	`prec_800ka_ann_hi_20c_bc.nc`
	MINP (BC)	`prec_800ka_min_hi_20c_bc.nc`
	Proxy comparison	`output_150ka.xlsx`

MAT–mean annual temperature; MINT–minimum temperature; MAP–mean annual precipitation; MINP–minimum precipitation; MAQ–mean annual specific humidity; MINQ–minimum specific humidity

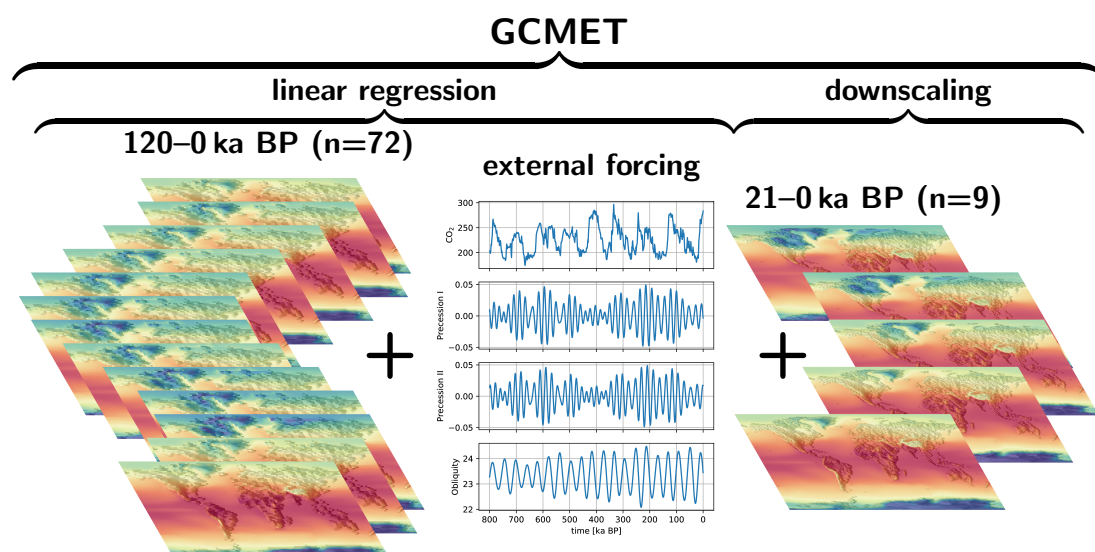BC–bias correction with ERA-20C 1961–1990 climatology

Figure 1. Schematic of the GCMET components: A linear regression combines 72 HadCM3 snapshot simulations with the external forcings, i.e., CO_2 and the three orbital parameters, which provides the basis of the long-term climate reconstructions of the last 800 thousand (or 2 million) years. Using 9 high-resolution snapshots covering the last deglaciation provides the basis of the downscaling approach based on CO_2 which yields the final high-resolution long-term climate reconstructions of GCMET.
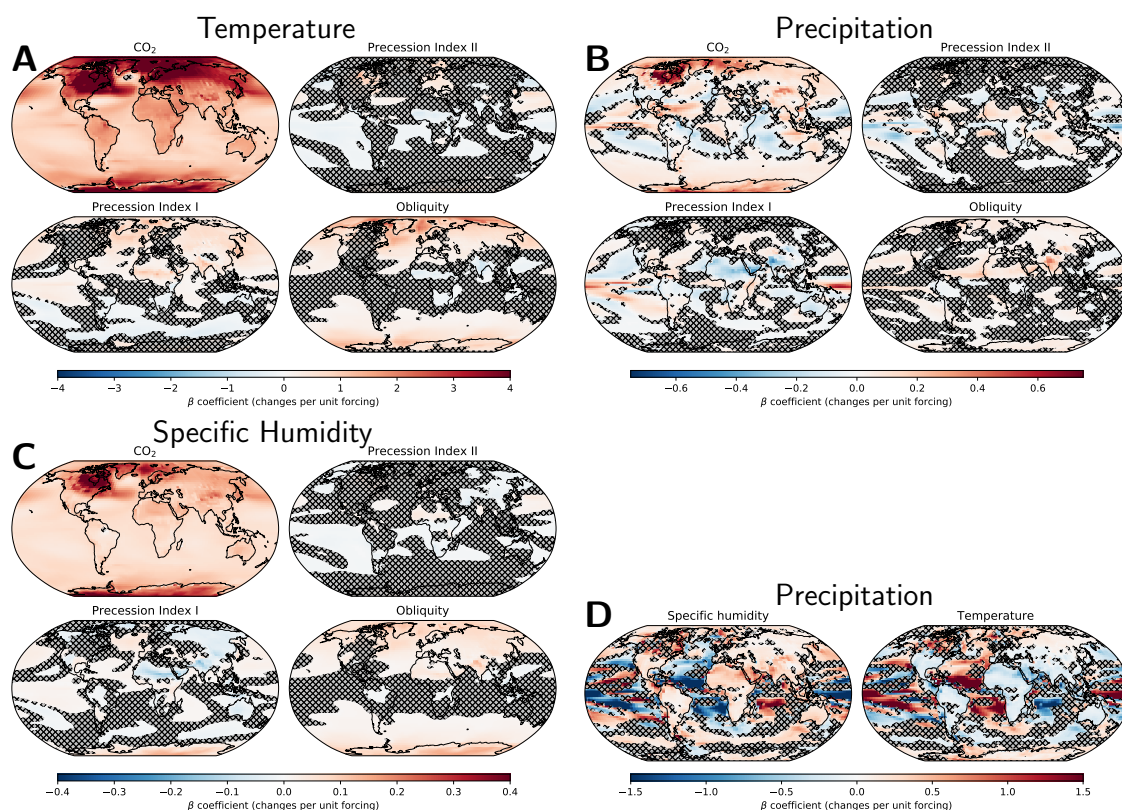
Figure 2. Regression coefficients, i.e., β coefficients, for (A) mean annual temperature, (B) precipitation, (C), specific humidity, and (D) the alternative model for precipitation—based on temperature and specific humidity. Regions where the respective coefficient is not statistically significant ($p < 0.05$) are hatched and shaded.
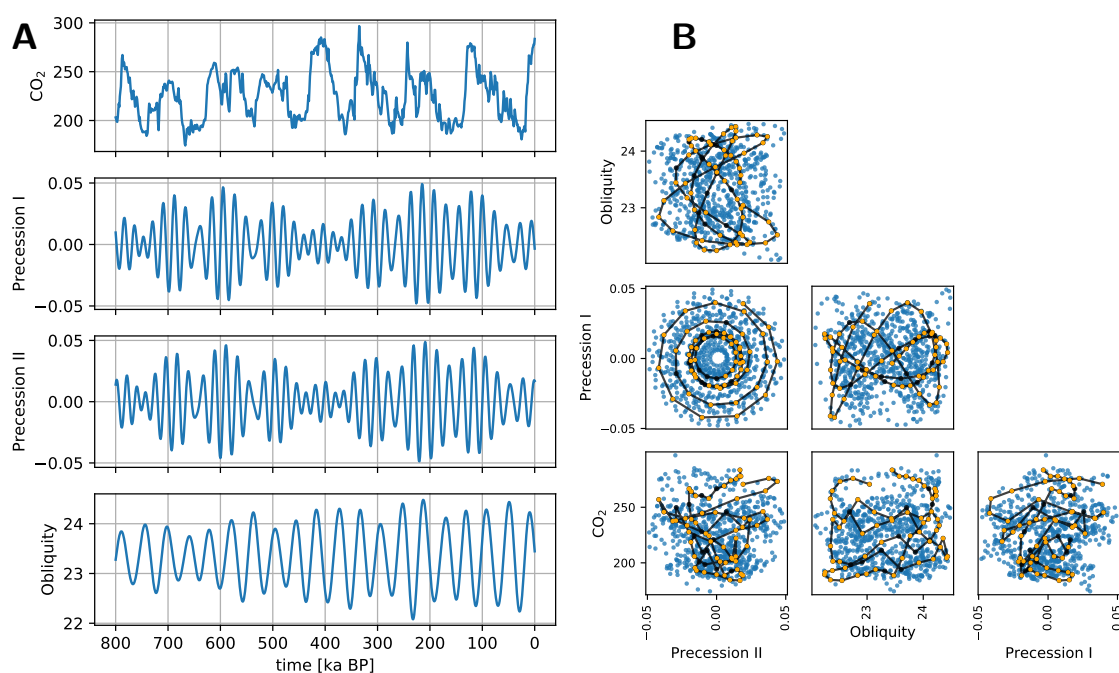
Figure 3. (A) Time series of the four external parameters: CO_2 and orbital parameters for the last 800 ka and (B) the associated parameter space as scatter plot matrix (blue dots). The continuous CO_2 record is from the EPICA Dome C ice core in Antarctica (Bereiter et al., 2015). The orbital parameters are numerical solutions for the Earth's orbit and rotation in terms of eccentricity, precession, and obliquity (Berger and Loutre, 1991). In (B), black lines with black dots represent the total 72 parameter sets. Orange dots highlight the parameter sets of the 58 HadCM3 snapshot simulations which we used as training data (80% of the total 72) for the linear regression model.
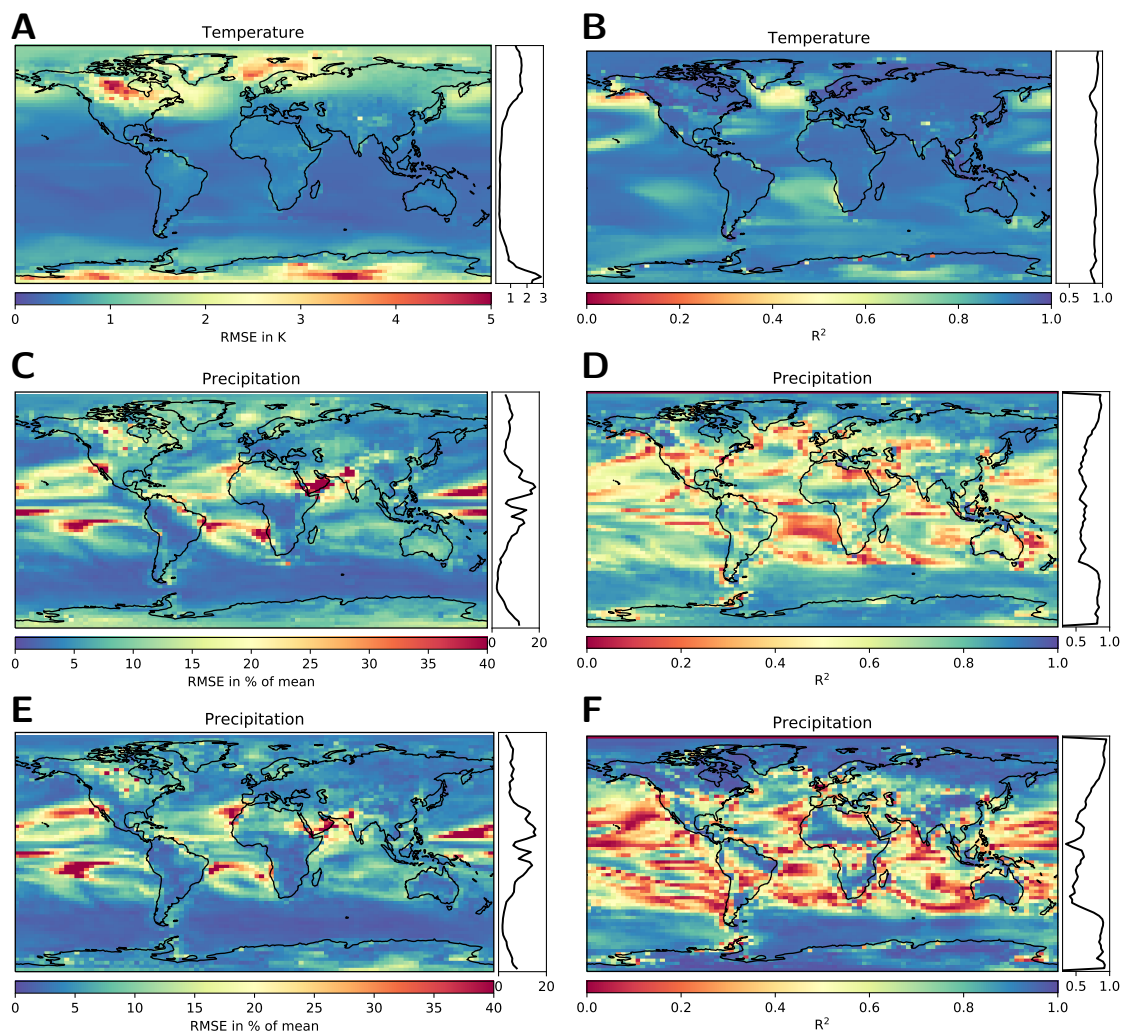
Figure 4. Left panel (A, C, E): Root mean square errors (RMSE) as estimators of the goodness of fit (lower is better) calculated using the test data. Right panel (B, D, F): R^2 values as estimator for the goodness of the model (higher is better) using the training data. Shown are the R^2 and RMSEs for (A,B) mean annual temperature, (C, D) precipitation, and (E, F) the alternative model for precipitation—based on temperature and specific humidity.
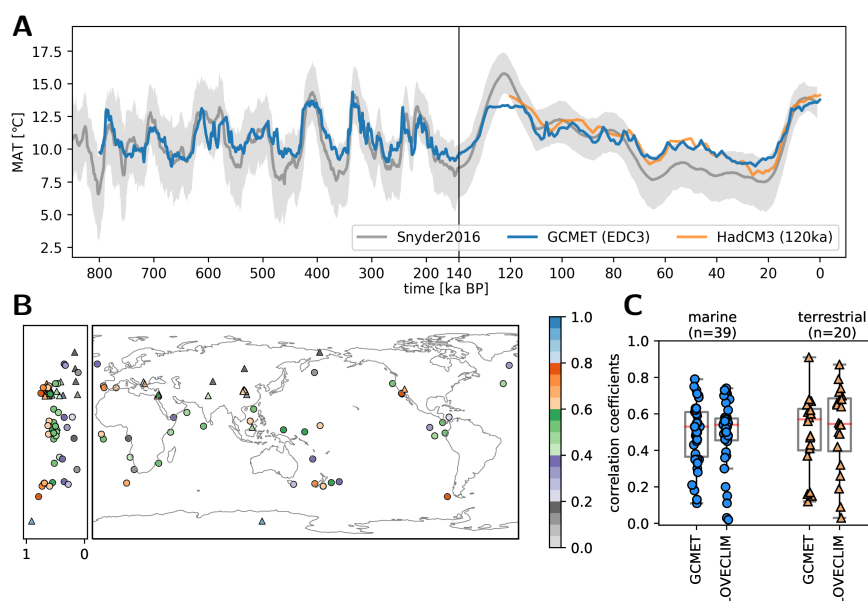
Figure 5. (*A*) Global mean temperature for the last 800 ka as predicted by GCMET based on different CO_2 records in comparison with a proxy-based global mean temperature reconstruction (Snyder, 2016). Furthermore, the time series from the 72 HadCM3 snapshots for the last 120 ka have also been added. Note the change in the spacing of the time axis at 140 ka BP. (*B*) Map of correlation coefficients between marine (in terms of as sea surface temperature) and terrestrial climate proxy time series and mean annual temperatures as reconstructed by GCMET-LO for the respective locations. The left panel shows the latitudinal distribution of the correlation coefficients. (*C*) Box plot of the correlation coefficients of GCMET and LOVECLIM with marine and terrestrial proxies.
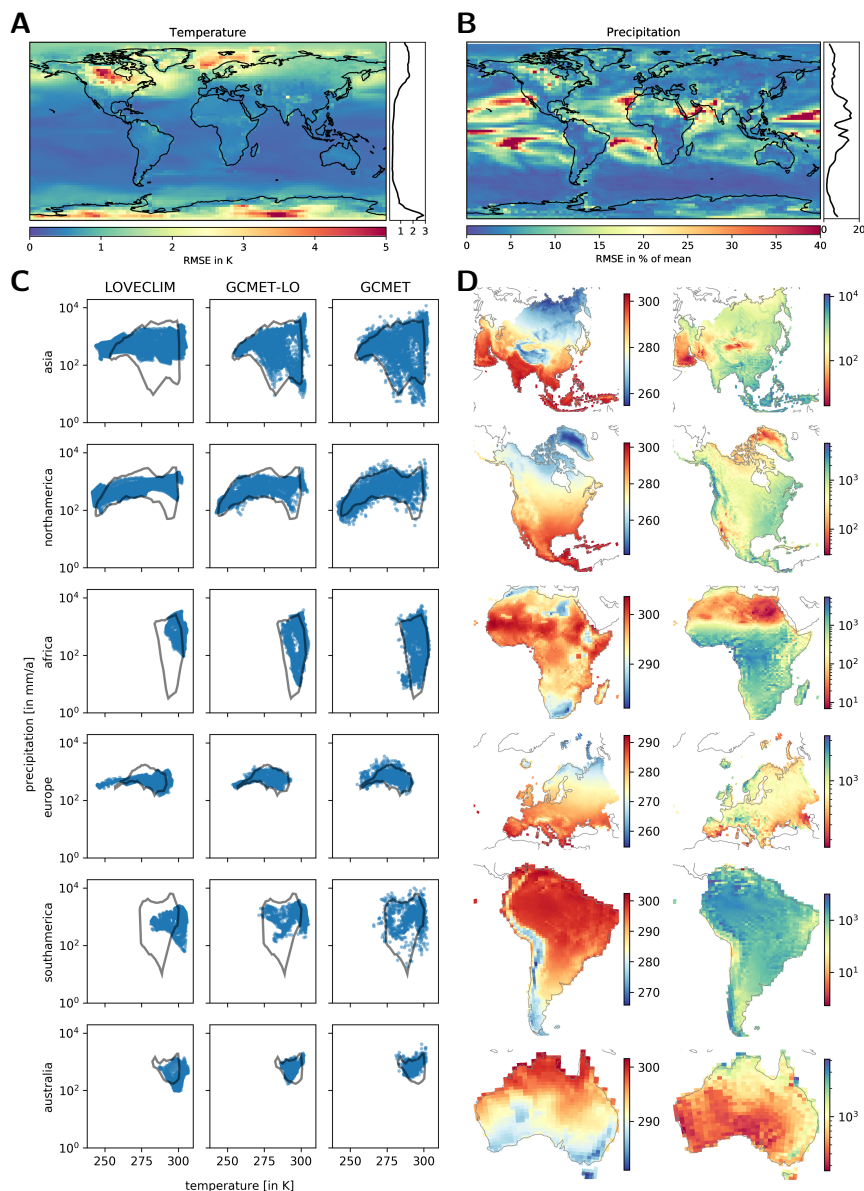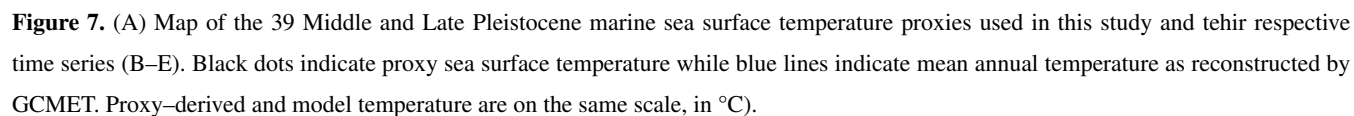
Figure 6. Root mean square error of the GCMET-LO predictions for the 14 HadCM3 snapshots for (*A*) MAT and (*B*) MAP (lower is better). (*C*) Present-day, i.e., 0 ka BP, temperature–precipitation phase diagram for Asia, North America, Africa, Europe, South America, and Australia, as modelled by LOVECLIM and reconstructed by GCMET-LO and GCMET and compared to observed multi-annual mean values (grey contours) for the period from 1961-1990 of the ERA-20C re-analysis data set (Poli et al., 2016). All model outputs have been bi-linearly interpolated onto the same grid, i.e. of the observational data, ERA-20C. (*D*) Maps of present-day temperature (in K) and precipitation (in mm/a) as reconstructed by GCMET for the six continents.

Climate
of the Past
Discussions

Figure 7. (A) Map of the 39 Middle and Late Pleistocene marine sea surface temperature proxies used in this study and tehir respective time series (B–E). Black dots indicate proxy sea surface temperature while blue lines indicate mean annual temperature as reconstructed by GCMET. Proxy–derived and model temperature are on the same scale, in °C.
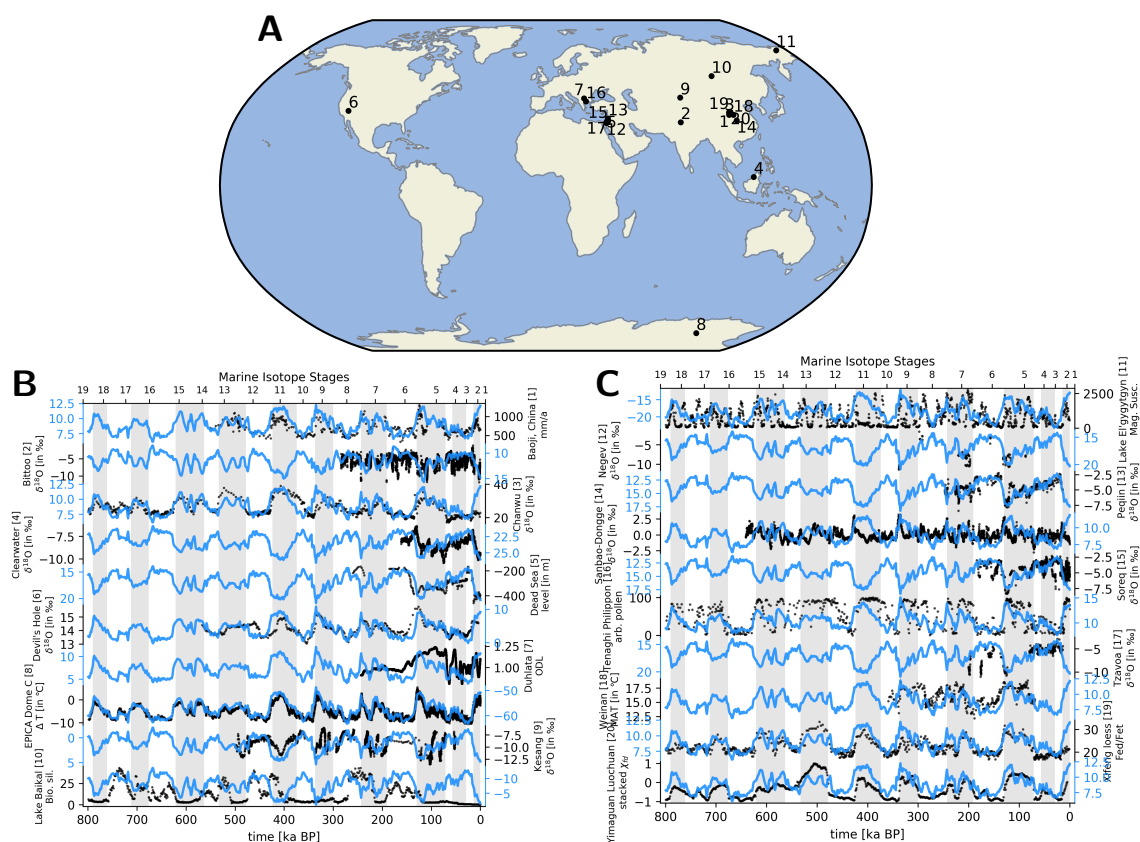
Figure 8. (A) Map of the 39 Middle and Late Pleistocene terrestrial climate proxies used in this study and their respective time series (B, C). Black dots indicate proxy variables (in different units) while blue lines indicate mean annual temperature as reconstructed by GCMET (in °C).
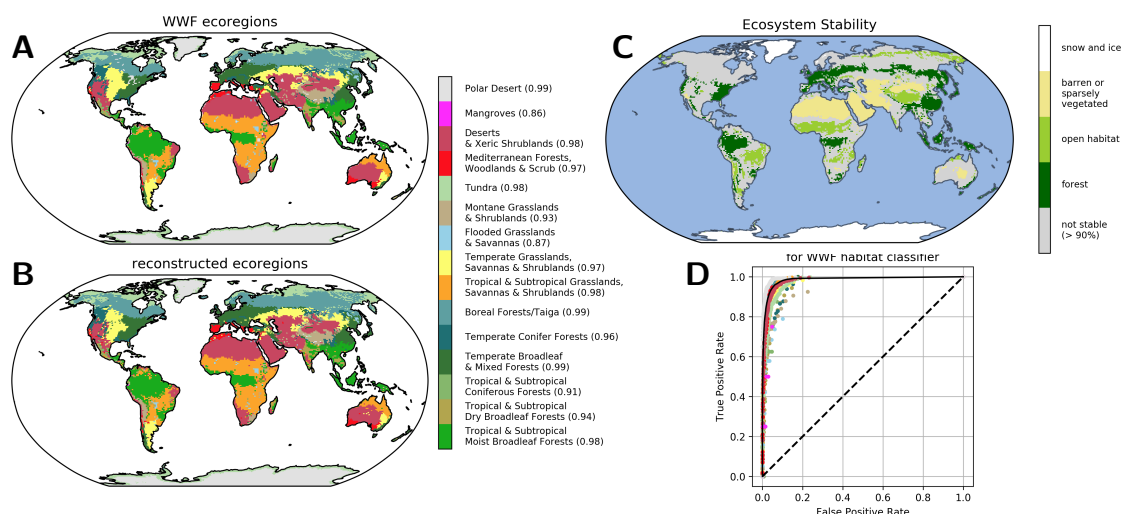
Figure 9. (*A*) Map of 14 major terrestrial habitats as defined by the WWF (Olson et al., 2001) for present-day and (*B*) as reconstructed with GCMET inputs of minimum and annual temperature and minimum and annual precipitation. (*C*) Stability of open habitats, such as grasslands and savannahs, and forest habitats, and sparsely vegetated regions across the world through the last 800,000 years. Regions in which the habitats have been unstable, i.e., of different type, for more than 90% are coloured in grey. (D) Receiver operating characteristic curve for the random forest classifier of the WWF 14 major habitats. The upper left corner represents a perfect prediction of an ecosystem, while the diagonal line represents a prediction made by random guessing. The closer the ROC curve is to the perfection point (0,1) the better the random forest classification is.