PALEO-PGEM-Series: A spatial time series of the global climate over the last 5 million years (Plio-Pleistocene)

How to cite:


For guidance on citations see FAQs.

© 2023 The Authors

https://creativecommons.org/licenses/by/4.0/

Version: Version of Record

Link(s) to article on publisher's website:
http://dx.doi.org/doi:10.1111/geb.13683

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data policy on reuse of materials please consult the policies page.
INTRODUCTION

Climate has an important role in the history of life on Earth, being a major predictor of variation in multiple dimensions of biodiversity over space and time (Svenning et al., 2015). For example, diversity tends to be greater in warmer, wetter and more productive environments (Barreto et al., 2021; Currie et al., 2004; Field et al., 2009; Sandel et al., 2011). These patterns result at least in part from...
extending back hundreds of years or further, with millennial mechanisms driving biodiversity (Rahbek et al., 2019) and to advance the direct or indirect effects of climate on diversification rates (Condamin et al., 2019), range shifts (Jansson & Dynnes, 2002), and other biological processes. Indirect effects of climate on biodiversity are mostly focused on the Holocene (6 kyr), the Last Glacial Maximum (LGM, 21 kyr), and how they differ from the present. These studies consistently show that climate change predicts spatial patterns of species richness (Agárd et al., 2008; Hortal et al., 2017), endemism (Rangel et al., 2018), genetic diversity (Carnaval, 2016), phylogenetic and ecosystem functioning (Odonne & Jetz, 2015), species and lineage turnover (Dobrovolskij et al., 2012), species interaction (Dalsgaard et al., 2012), and community structure (Rowan et al., 2016). However, such a focus on the last 21 kyr restricts attention to the influence of glacial and Northern Hemisphere changes induced by spatial configuration and increases in greenhouse gas concentration (Clark et al., 2012).

Biogeographical and macroecological studies investigating how current biodiversity patterns relate to palaeoclimate are mostly focused on the LGM (21 kyr) and community structure (Rowan et al., 2016). However, a more complete understanding of how historical climatic conditions influenced the origin and maintenance of biodiversity is constrained by the scarcity of high-resolution palaeoclimatic estimates dating back beyond the most recent glacial period. Only recently have palaeoclimate reconstructions at fine spatial and temporal resolution become available (Beyer et al., 2020). These studies consistently show that the strong climate–diversity relationship, one thing is certain: past climatic regimes played an important role in the generation and maintenance of biodiversity (Flantua et al., 2019; Rangel et al., 2018).

TABLE 1
Summary of the freely available global palaeoclimatic estimates at the spatial scales of interest for macroecological studies.

<table>
<thead>
<tr>
<th>Database</th>
<th>Nature of the data</th>
<th>Oldest period covered</th>
<th>Temporal resolution</th>
<th>Smallest spatial scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHELSA-TracE21k</td>
<td>Transient CCSM3-TracE21k simulations coupled with ICE6G dynamic ice sheet model with downscaling</td>
<td>Pleistocene (LGM)</td>
<td>Last 21 kyr every 100 years</td>
<td>0.008°×0.008°</td>
</tr>
<tr>
<td>PaleoView</td>
<td>Downscaling transient CCSM3 simulation with downscaling</td>
<td>Pleistocene (LGM)</td>
<td>Last 21 kyr every 10 years</td>
<td>2.5°×2.5°</td>
</tr>
<tr>
<td>Beyer et al. (2020)</td>
<td>Simulations of the medium-resolution HadCM3 for the last 120,000 years combined with the high-resolution HadCM3 for the last 21,000 years</td>
<td>Pleistocene</td>
<td>Last 120 kyr every 1 or 2 kyr</td>
<td>0.5°×0.5°</td>
</tr>
<tr>
<td>Krapp et al. (2021)</td>
<td>Linear regressions between HadCM3 simulations coupled with external forcings</td>
<td>Pleistocene</td>
<td>Last 800 kyr every 1 kyr</td>
<td>0.5°×0.5°</td>
</tr>
<tr>
<td>Ecolimate</td>
<td>Snapshot PMIP3 simulations with downscaling</td>
<td>Pliocene</td>
<td>6 ka, 21 ka, -3.3–3.0 Ma</td>
<td>0.5°×0.5°</td>
</tr>
<tr>
<td>PaleoClim</td>
<td>Snapshot HadCM3 simulations with downscaling</td>
<td>Pliocene</td>
<td>130 ka, 787 ka, 3.264–3.025 Ma, 3.3 Ma</td>
<td>0.04°×0.04°</td>
</tr>
<tr>
<td>Oscillayers</td>
<td>Pattern-scaled WorldClim LGM anomalies, temporally scaled by oxygen isotope data</td>
<td>Pliocene</td>
<td>Last 5 Myr every 10 kyr</td>
<td>0.04°×0.04°</td>
</tr>
<tr>
<td>PALEO-PGEM-Series</td>
<td>Time series with downscaling</td>
<td>Pliocene</td>
<td>Last 5 Myr every 1 kyr</td>
<td>1°×1°</td>
</tr>
</tbody>
</table>

towards mechanistic approaches to model biodiversity (Cabral et al., 2017; Hagen, 2022; Rangel et al., 2018). For instance, several questions remain unanswered owing to the lack of such data. How did palaeoclimatic fluctuations affect diversification and population dynamics? How did stability and spatial prevalence of environmental conditions over time influence biodiversity? What is the maximum rate of climate change that still allows for some evolutionary rescue? Such knowledge is paramount to predicting how biodiversity might respond to current anthropogenic climatic change (Fordham et al., 2016).

To help fill the gap, we present the PALEO-PGEM-Series, a general circulation model (GCM)-based spatially explicit time series of the Pliocene–Pleistocene (i.e., 5 Ma) downscaled from the outputs of the PALEO-PGEM emulator (Holden et al., 2019). PALEO-PGEM emulates reasonable and extensively validated global estimates of monthly temperature and precipitation for the Plio-Pleistocene every 1 kyr at a spatial resolution of −5°×5° (Holden et al., 2016, 2019). To facilitate and promote the use of these spatio-temporal datasets in biogeographical and macroecological studies, we process the rich and dense set of emulation outputs by downscaling the estimates to a 1°×1° grid and deriving 17 bioclimatic variables that summarize central trends, variability and extreme conditions of known relevance to ecological and evolutionary dynamics. In addition, unlike alternative models (Table 1), we provide uncertainty estimates in the form of SE of the emulated climate over multiple runs of the emulator (see Section 2.4). This allows users to estimate the robustness of their results in the face of the stochastic aspects of the emulations. We encourage users to perform their analysis considering both the mean and the SD of the estimates by, for example, replicating the analyses using the range of possible palaeoclimatic estimates determined by the SE.

## 2 | PALEO-PGEM-SERIES

### 2.1 | Emulation

Most global palaeoclimatic reconstructions at fine spatial resolution come from the downscaling of simulations from GCMs, such as the Palaeoclimate Modelling Intercomparison Project (PMIP3) and the Coupled Model Inter-comparison Project (CMIP5) (Taylor et al., 2012). However, GCM simulations become computationally intractable when reconstructing climate on a million-year scale, causing the general lack of spatio-temporal climatic series for such deep time at fine spatial scales. A recent approach to address this constraint generated a 2 Myr quasi-transient simulation with the Community Earth System Model (Timmermann et al., 2022), which spliced together 21 quasi-100 kyr simulations with orbital variations accelerated by a factor of five, such that each segment consisted of 20 kyr of modelled time but was assumed to represent 100 kyr of real time.

Our PALEO-PGEM-Series overcomes computational limitations by downscaling climatic reconstructions from a GCM-based palaeoemulator able to retain the effects of climate dynamics with sufficient spatial detail for macroecological studies, namely the PALEO-PGEM (Holden et al., 2019; Rangel et al., 2018). The approach applies Gaussian process emulation of the singular value decomposition of ensembles of runs from the intermediate complexity atmosphere–ocean GCM PLASIM-GENIE (planet simulator-grid enabled integrated Earth system model; Holden et al., 2016) with varied boundary-condition forcing. Spatial fields of monthly temperature and precipitation were emulated at 1000-year intervals, driven by time series of scalar boundary-condition forcing of CO₂ (Lüthi et al., 2008), orbit (Berger & Loutre, 1999) and ice volume (Stap et al., 2017) and assuming that the climate is in quasi-equilibrium. This approach neglects “memory” effects that retain an imprint of previous forcing conditions, which are small for the millennial scale of these emulations (i.e., 1000-year time interval). Effectively, this means that PALEO-PGEM does not capture the effects of dynamic processes operating at sub-millennial time-scales, which contrasts with alternative palaeo-reconstructions that use accelerated simulation and overestimate such transient effects (Timmermann et al., 2022).

### 2.2 | Downscaling and deriving bioclimatic variables

Here, we downscaled the monthly temperature and precipitation emulations using anomaly adjustments (Osborn et al., 2016) and a current climate baseline from CHELSA (climatologies at high resolution for the Earth’s land surface areas) for the years 1981–2010 (Karger et al., 2017). CHELSA adequately captures fine-scale variations in climate, especially in areas of great topographical variation and with sparse and unreliable climate stations (Karger et al., 2017), which makes it a suitable baseline for downscaling. We downscaled the emulations for land areas that are available in the supplementary material of Holden et al. (2019) from the native resolution of −5°×5° onto a higher-resolution global grid of 1°×1° every 1000-year interval (5000 time slices; Figure 1). For each time step, we interpolated the climate anomalies linearly between the emulated climate for time t (E_t) and for pre-industrial (E_0), a steady-state pre-industrial baseline with modern orbital forcing and CO₂ concentration of 280 ppm) onto the high-resolution spatial grid. The interpolation was performed using a distance-based weighted mean from the centroid of the downscaled cell to the centres of the cells at the original resolution. To preserve a realistic spatial heterogeneity in estimated palaeoclimate, we applied additive anomaly adjustment (Equation 1) using the current climate, C_{t}, from CHELSA (Karger et al., 2017):

$$C_{t} = C_{0} + (E_{t} - E_{0}). \tag{1}$$

However, in very dry regions the additive anomaly adjustment applied to precipitation might generate negative estimates if the underlying climate model fails to represent observed extreme dryness.
FIGURE 1 Representation of PALEO-PGEM-Series time series. Monthly temperature and precipitation were averaged across the global land surface for each time step. Black lines indicate mean values, and the grey shading represents the model uncertainty, measured as the SE of the climate estimated across multiple GCM emulations. Time is in years before present.
Thus, we applied multiplicative anomaly adjustment (Equation 2) for precipitation only, in regions where $E_0 > C_0$.

$$C_t = C_0 \times (E_t / E_0).$$

(2)

The multiplicative anomaly adjustment represents better small changes in precipitation in permanently arid regions that are not well captured by the emulator, preventing the occurrence of unrealistic negative precipitation estimates in hyper-arid deserts and overestimates in wetter areas. For instance, the multiplicative term enabled us to reconstruct the hyper-aridity of the Atacama Desert and the seasonally hyper-wet monsoons. We use the additive adjustment for precipitation when $E_0 < C_0$ because this prevents the possibility of unrealistically high precipitation when $E_t / E_0 \gg 1$. Note that multiplicative and additive adjustments yield similar estimates in regions where $E_0 = C_0$.

From the downscaled monthly temperature and precipitation estimates, we derived 17 bioclimatic variables that summarize central tendencies (e.g., annual means), seasonality (e.g., range of monthly means) and extreme conditions (e.g., maximum and minimum monthly means) of relevance for ecological and evolutionary studies (Table 2). Mean diurnal range (BIO 2) and isothermality (BIO 3) cannot be derived from PALEO-PGEM-Series because the diurnal cycle was not resolved. Maximum temperature (BIO 5) and minimum temperature (BIO 6) are derived from monthly means and thus refer to the maximum and minimum monthly means of temperatures in the year. Consequently, temperature annual range (BIO 7) is also with reference to monthly means.

### 2.3 Validation

The modern climate of PLASIM-GENIE was validated by Holden et al. (2016), supplemented with validation of the simulated Asian monsoon (Thomson et al., 2021). Holden et al. (2019) provide an extensive validation of the palaeomulator (i.e., PALEO-PGEM) against model inter-comparisons of the mid-Holocene, the LGM, the Last Interglacial (LIG) and the mid-Pliocene warm period, with glacial–interglacial variability over the last 800,000 years validated against observationally based global temperature reconstructions. Emulated palaeoclimatic validations were supplemented in the study by Wang et al. (2021) with comparisons of mean annual temperature and precipitation against the LGM PMIP4 ensemble (Kageyama et al., 2021) that showed how the emulated mean LGM annual temperature was approximately centred within the highly uncertain PMIP4 envelope. The emulator has also been shown to reproduce the large-scale regionally heterogeneous hydroclimatic behaviour observed in proxy records spanning 130–70 ka (Nilsson-Kerr et al., 2021).

We complement these validations by comparing the down-scaled climate data of PALEO-PGEM-Series with proxy data from various times and regions (i.e., model–proxy comparison). Initially, we correlated the time series of 17 bioclimatic variables from PALEO-PGEM-Series with those of 19 terrestrial proxies over the past 800 ka that were compiled by Krapp et al. (2021), including proxies of different types, such as benthic oxygen isotopes ($\delta^{18}$O), surface pollen and magnetic susceptibility. We found good agreement between proxy and model, with only 5 of the 19 proxy time series not showing correlations $>|0.5|$ (Supporting Information Table S1). Each proxy was best approximated by different bioclimatic variables, suggesting that proxies of different types might capture different aspects of climate (e.g., mean temperatures, summer or winter temperatures, seasonality). Here, we illustrate the time series of each proxy paired with annual mean temperature (BIO 1) from PALEO-PGEM-Series and Krapp et al. (2021), for comparison with the latter model (Figure 2). Overall, the BIO 1 estimates from PALEO-PGEM-Series are coherent with the temporal patterns from proxy at various locations and of different natures, such as those based on speleothem oxygen isotope $\delta^{18}$O (e.g., Figure 2a), the ratio between free Fe$_2$O$_3$ (Fed) and total Fe$_2$O$_3$ (Fet) (Fed/Fet ratio; Figure 2k), magnetic susceptibility (e.g., Figure 2n)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>Monthly mean</td>
</tr>
<tr>
<td>Precipitation</td>
<td>Monthly mean</td>
</tr>
<tr>
<td>BIO 1</td>
<td>Annual mean temperature</td>
</tr>
<tr>
<td>BIO 4</td>
<td>Temperature seasonality (SD x 100)</td>
</tr>
<tr>
<td>BIO 5</td>
<td>Maximum monthly temperature</td>
</tr>
<tr>
<td>BIO 6</td>
<td>Minimum monthly temperature</td>
</tr>
<tr>
<td>BIO 7</td>
<td>Temperature annual range (BIO 5–BIO 6)</td>
</tr>
<tr>
<td>BIO 8</td>
<td>Mean temperature of the wettest quarter</td>
</tr>
<tr>
<td>BIO 9</td>
<td>Mean temperature of the driest quarter</td>
</tr>
<tr>
<td>BIO 10</td>
<td>Mean temperature of the warmest quarter</td>
</tr>
<tr>
<td>BIO 11</td>
<td>Mean temperature of the coldest quarter</td>
</tr>
<tr>
<td>BIO 12</td>
<td>Annual precipitation</td>
</tr>
<tr>
<td>BIO 13</td>
<td>Precipitation of the wettest month</td>
</tr>
<tr>
<td>BIO 14</td>
<td>Precipitation of the driest month</td>
</tr>
<tr>
<td>BIO 15</td>
<td>Precipitation seasonality (coefficient of variation)</td>
</tr>
<tr>
<td>BIO 16</td>
<td>Precipitation of the wettest quarter</td>
</tr>
<tr>
<td>BIO 17</td>
<td>Precipitation of the driest quarter</td>
</tr>
<tr>
<td>BIO 18</td>
<td>Precipitation of the warmest quarter</td>
</tr>
<tr>
<td>BIO 19</td>
<td>Precipitation of the coldest quarter</td>
</tr>
</tbody>
</table>

Note: Tab-delimited text files containing the mean and SD of the estimates are available in FigShare (https://figshare.com/s/d45714f7212de7225fe2). Temperature-related variables are in degrees Celsius and precipitation-related variables in millimetres. All bioclimatic variables are derived from monthly means.
and rainfall (Figure 2q). We also found a strong correspondence between annual mean temperature from PALEO-PGEM-Series and Krapp et al. (2021) (Figure 2).

For the LIG (130–116 ka), we compared the proxy anomalies of mean annual surface temperature relative to 1961–1990 that were compiled by Turney and Jones (2010) with the anomalies in PALEO-PGEM-Series mean annual and summer temperatures at 127 ka, the thermal maximum, relative to a steady-state pre-industrial baseline (Figure 3). PALEO-PGEM-Series reconstructs both positive and negative anomalies in mean annual temperature at 127 ka, similar to what has been found for the proxies, but with a different spatial pattern and magnitude (Figure 3). The proxies point to a latitudinal gradient of warmer anomalies in northern regions, especially in Siberia (Figure 3a), whereas PALEO-PGEM-Series under-predicts such high-latitude warming and points to warmer conditions in central Europe (Figure 3b,c). Given that the proxy anomalies are with reference to modern climate and the emulated anomalies with reference to a steady-state pre-industrial, we expected that the former could have smaller amplitudes of warm anomalies. However, we found that the proxies have a greater amplitude of warm anomalies (~5 to 15°C; Figure 3) than the emulated and downscaled climate from PALEO-PGEM-Series (~0.2 to 6°C and ~1 to 2°C of summer and mean temperature anomalies, respectively; Figure 3b,c). That is in line with the overall inability of climate models to reproduce the global mean LIG annual warming of ~1 to 2°C inferred from proxy-based reconstructions (McKay et al., 2011; Otto-Biesner et al., 2013; Turney & Jones, 2010).

Lastly, we compared the temperature anomaly from the Plio-Pleistocene with respect to a pre-industrial baseline using proxies compiled by Salzmann et al. (2013) at various time intervals between 3.8 and 1.8 Ma. Since these proxies are expected to indicate mean annual temperatures (bio 1), we derived the maximum and minimum bio 1 anomalies within the time intervals relevant to each proxy (Figure 4). Maximum temperature anomalies from PALEO-PGEM-Series (Figure 4b) show a better correspondence with the temperature estimated from proxies (Figure 4a) than the minimum anomalies (Figure 4c), by pointing to overall warmer conditions in the Plio-Pleistocene than in the pre-industrial. Yet, such maximum temperature anomalies from PALEO-PGEM-Series have a smaller amplitude (1 to 7°C, Figure 4b) than the ones calculated from proxies (~4 to 20°C, Figure 4a). Particularly, anomalies in northern Russia calculated from PALEO-PGEM-Series are milder than those from calculated from proxy, similar to model-proxy comparisons of an alternative climatic model (Salzmann et al., 2013) and to what we observe in the model-proxy comparison for the LIG (Figure 3). Extreme anomalies, in both PALEO-PGEM-Series and proxy data, are mostly in areas where our model presents greater variation between estimates from different emulation runs (smallest circles in Figure 4b,c), reinforcing the importance of considering not only the mean but also the SE of estimates from PALEO-PGEM-Series (see Section 2.4 on model limitations and uncertainties).

Generally, some deviations between emulated estimates and proxy data are expected, as with other palaeoclimatic models (e.g., Otto-Biesner et al., 2013; Salzmann et al., 2013). These mismatches could stem from the model (see Section 2.4 for a comment on the limitations of our model), but also at least in part from the interpretation of proxy data (Hossain et al., 2018).

### 2.4 Limitations and uncertainties

PALEO-PGEM-Series is subject to limitations on accuracy and applicability, as is the case for all palaeoclimatic estimates. The paper by Holden et al. (2019) includes a discussion section on the main issues affecting PALEO-PGEM. This includes discussions on the limitations of low-resolution and intermediate-complexity climate models, the use of limited climate forcing (i.e., concentration of atmospheric CO₂, orbital configuration and ice sheet state), the assumption of a fixed relationship between global sea-level reconstructions from benthic oxygen isotopes and the spatial configuration of ice sheets, and limitations of spatial downscaling. In addition, the land-sea mask is not altered in the simulations, which could underestimate climate change in certain regions (e.g., Haywood et al., 2020), such as in the Fennoscandian ice sheet.

We accommodate aspects of model uncertainty in PALEO-PGEM-Series by reporting the mean and SE of the emulated climate over 10 stochastic GCM emulations (Figure 1). This constitutes an estimate of the error incurred in statistically emulating the underlying simulations, and we suggest that users take advantage of such uncertainty estimates by conducting sensitivity tests to assess the robustness of results derived using PALEO-PGEM-Series. Although these uncertainty estimates are technically distinct from errors in the simulator itself, the magnitudes and spatial patterns of these two sources of uncertainty have been shown to be generally similar (Holden et al., 2019). An important exception is that while the emulator uncertainty (which we report here) captures much of the uncertainty seen in multi-model inter-comparisons, PALEO-PGEM-Series cannot fully represent all model uncertainty because it is derived from a single configuration of a single simulation model. Most particularly, the 90% uncertainty range of climate sensitivity (3.8 ± 0.6°C)
FIGURE 3  Visual comparison of temperature anomalies of the Last Interglacial Period (LIG) from (a) proxy data encompassing the period of 130–116 ka (Turney & Jones, 2010), and PALEO-PGEM-Series (b) mean and (c) summer temperatures at the LIG maximum (127 ka).

Anomalies from proxy data are relative to 1961–1990 and those from PALEO-PGEM-Series are relative to a steady-state pre-industrial baseline with modern orbital forcing and CO₂ concentration of 280 ppm. Red indicates higher temperatures during the LIG than the pre-industrial period. Dots are scaled by the inverse of the SE in the climate estimated across multiple GCM emulations, such that the bigger the circle, the smaller the variation in PALEO-PGEM-Series.
Figure 4: Visual comparison of temperature anomalies (i.e., past minus pre-industrial temperature) from (a) proxy data compiled by Salzmann et al. (2013) and (b) maximum and (c) minimum annual temperatures from PALEO-PGEM-Series encompassing the Plio-Pleistocene (1.8–3.8 Ma). Red indicates higher temperatures during the Plio-Pleistocene than in the pre-industrial period. Dots represent the confidence in proxy (following Salzmann et al., 2013) and PALEO-PGEM-Series, the latter being scaled by the inverse of the SE in the climate estimated across multiple GCM emulations. The bigger the circle, the smaller the variation in estimates from PALEO-PGEM-Series.
is understated relative to multi-model estimates of $3.2 \pm 1.3$°C (Flato et al., 2014). Where appropriate, climate sensitivity uncertainty could be inflated through a global scaling of emulated temperature anomalies.

### 2.5 Data availability

We made available tab-delimited text files in FigShare (https://figshare.com/s/d45714f7212de7225fe2) containing monthly temperature (in degrees Celsius), monthly precipitation (in millimetres) and 17 bioclimatic variables estimated over the last 5 Myr every 1000 years (Table 2). The data refer to the mean and the SD of the estimates across the 10 stochastic GCM emulation runs. Each file contains 19,233 rows representing geographical gridded cells and 5003 columns, the first two with coordinates (longitude and latitude) and the others with the estimates for each time slice, starting from 5 Ma and ending with the pre-industrial period. Each column is named with “T” followed by a number representing the time (in years) divided by 1000 (i.e., T5000 corresponds to time 5 Ma). These tables can easily be converted into spatial files (e.g., raster) and visualized and explored using GIS libraries. Additionally, we provide an R script to derive the 17 bioclimatic variables from the monthly estimates (code adapted from the function `biovars` in R package `dismo`; Hijmans et al. 2021).

### 3 | APPLICABILITY

PALEO-PGEM-Series can be used to explore a series of long-standing ecological, evolutionary, biogeographical and macroecological hypotheses of great relevance to understanding the drivers of diversity and forecasting organisational response to future climatic change. For example, our spatio-temporal series proved highly relevant to reconstruct long-term patterns and dynamics of biotas (e.g., populations, species, biomes and ecosystems; Mondanaro et al., 2020; Raia et al., 2020; Wang et al., 2021) and to simulate the origin and maintenance of biodiversity (Diniz-Filho et al., 2019; Rangel et al., 2018). The time frame of PALEO-PGEM-Series, which ranges from thousands to millions of years, offers the possibility to study the evolutionary responses of species to past climate change, moving from the current focus on the LGM and the inference of climatic stability as a driver of species persistence (i.e., extinction rates and range shifts) to a driver of species generation (i.e., speciation rates) at evolutionary scales (Svenning et al., 2015).

### ACKNOWLEDGEMENTS

P.B.H. and N.R.E. were funded by NERC (grant no. NE/P015093/1). E.B. was supported by a doctorate fellowship from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)—Finance Code 001, and a postdoctoral fellowship from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (no. 787638) granted to Catherine H. Graham. This study was partially financed by the MCTIC/CNPq (grant no. 465610/2014-5) and Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG grant no. 201810267000023) in the context of the National Institute of Science and Technology (INCT) in Ecology, Evolution and Biodiversity Conservation.

### CONFLICT OF INTEREST

None declared.

### DATA AVAILABILITY STATEMENT

Tab-delimited text files with the monthly temperature, monthly precipitation and 17 bioclimatic variables estimated for the last 5 Myr every 1000 years together with the R script to derive the bioclimatic variables are available in FigShare (https://figshare.com/s/d45714f7212de7225fe2). The palaeoclimatic emulator used to derive such estimates is available in Holden et al. (2019).

### ORCID

Elisa Barreto https://orcid.org/0000-0002-3372-7295

### REFERENCES


Berger, A., & Loutre, M.-F. (1999). *Parameters of the Earth’s orbit for the last 5 Million years in 1 kyr resolution* [Data set]. PANGAEA. https://doi.org/10.1594/PANGAEA.56040


Communications Earth & Environment, 2(1). https://doi.org/10.1038/s43247-021-00133-7


BIOSKETCH

We are a team of evolutionary ecologists and palaeoclimatologists interested in the role of climate on eco-evolutionary dynamics.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.