



OPEN

DATA DESCRIPTOR

# Replicable Fine-Spatio-Temporal Climate Data for Long-Term Ecology in the Western Mediterranean

Daniel Romera-Romera<sup>1,3</sup> , Francisca Alba-Sánchez<sup>2</sup> , Daniel Abel-Schaad<sup>2</sup> & Diego Nieto-Lugilde<sup>1,3</sup>

Despite the increasing availability of climate data through various databases, obtaining fine temporal and spatial resolution data for past periods remains challenging. Here, we present (i) a toolkit for applying advanced downscaling techniques to coarse-resolution climate datasets within the widely used R programming framework, and (ii) downscaled data for a region recognized as a global biodiversity hotspot. Specifically, this toolkit consists of two R-packages (dsclim and dsclimtools) that were used to downscale seven climate variables for the Western Mediterranean, providing monthly climate data from 22 ka BP to the year 2100 at a spatial resolution of  $11 \times 11$  km. Our aim is to offer open access to a cutting-edge climate dataset for researchers interested in this region and to encourage the reuse of both the dataset and the toolkit, facilitating the creation of similar high-resolution climate products for other regions. Given the ecological importance of this region, we also provide examples of scientific applications, such as spatio-temporal pattern analysis and ecological niche modeling, demonstrating its scientific value.

## Background & Summary

Weather and climate influence a wide range of natural processes, with significant implications for ecological phenomena. For instance, they shape geomorphological processes, such as erosion and soil formation<sup>1</sup> and influence biodiversity by affecting population dynamics, species distribution, community assembly, and ecosystem productivity<sup>2</sup>. Therefore, further study of climate and climate change is crucial to improving our understanding of both the climate itself and the natural processes it affects<sup>3</sup>. This is also of practical importance due to the connections between climate and key areas such as water resources, food supply, health, and the economy<sup>4</sup>. In recent decades, the establishment and coordination of extensive meteorological networks have enabled the generation of global meteorological and climate data. When combined with atmospheric models, interpolation algorithms, and data assimilation methods, these data have facilitated the creation of comprehensive global meteorological and climatological datasets<sup>5</sup>. These datasets are invaluable in many fields, including climatology, geography, agronomy, architecture, urban planning, and economics<sup>6,7</sup>. In ecology and biogeography, they enhance our understanding of macroecological patterns<sup>8</sup>, enable the study of how biological entities (individuals, populations, communities, or biomes) interact with climate<sup>9</sup>, and help predict the impacts of climate change on biodiversity<sup>10</sup>.

Understanding how past environmental changes affected biodiversity, and how it responded, is essential for comprehending these dynamics<sup>11</sup>. This is especially relevant when studying periods of change that parallel the current rate of climate change, as it allows us to anticipate future effects and design appropriate management and conservation strategies<sup>12</sup>. Most studies in this field are based on spatial paleoclimate reconstructions derived from General Circulation Models (GCMs)<sup>13</sup>. These reconstructions are typically validated with independent paleoclimate data or proxies, such as fossil records or stable isotopes. These reconstructions have been

<sup>1</sup>Departamento de Botánica, Ecología y Fisiología Vegetal. Universidad de Córdoba, Córdoba, Spain. <sup>2</sup>Departamento de Botánica. Universidad de Granada, Granada, Spain. <sup>3</sup>These authors contributed equally: Daniel Romera-Romera, Diego Nieto-Lugilde. ✉e-mail: [b62rorod@uco.es](mailto:b62rorod@uco.es)

Database	Spatial resolution	Temporal coverage	Data source
PMIP (different versions)	$>320 \times 160$	6ka and 21 ka Last millennium Early Holocene and 8.2ky BP event 130–125–115 ky BP Mid-Pliocene	GCM Simulation
CMIP (Different versions)	$>320 \times 160$	Present–2100	GCM Simulation
TraCE-21ka	$T31_g \times 3$ ( $\sim 3.75^\circ \times 3.75^\circ$ )	22ka –1990 ce	GCM Simulation
CRU-TS*	$0.5^\circ \times 0.5^\circ$	1901–2018	Weather stations interpolation
ERA5	31 km grid	1940 to present	Data reanalysis from weather stations and simulations
ERA5-Land	$0.1^\circ \times 0.1^\circ$	1950 to present	Data reanalysis from weather stations and simulations
WorldClim	30 seconds or 2.5 minutes**	1970–2000 1960–2018 2021–2100	Spatial interpolation from weather stations and downscaling from GCM simulations
Ecoclimate	$0.5^\circ \times 0.5^\circ$	1950–1999 1900–1949 ~1760 6ka–21ka 3 Ma 2080–2100	Downscaled (Interpolation and Delta method and debias): PMIP3 – CMIP5
CHELSA-TraCE21k	30 arc sec	1901–2016 21ka	Downscaled (Delta method and debias): TraCE-21ka and CHELSA V1.2
PaleoClim	2.5 min 30 sec (CHELSA)	4.2–0.3 ka 8.326–4.2 ka 11.7–8.326 ka 12.9–11.7 ka 14.7–12.9 ka 17.0–14.7 ka ca. 130 ka *ca. 787 ka *3.205 Ma *ca. 3.3 Ma	Downscaled (Delta method and debias): CHELSA and Paleoview
PaleoView	$2.5^\circ \times 2.5^\circ$	21ka	Downscaled (bilinear interpolation): TraCE-21ka
E-OBS	$0.1^\circ \times 0.1^\circ$	1950–2024	Weather stations interpolation.
TerraClimate	$\sim 4$ km	1958–2020	Downscaled (Delta method and debias): CRU Ts4.0/JRA55 and WorldClim

**Table 1.** Description of several climate products in raster format at large/global scale publicly available, including their spatial extent, their finest spatial resolution, their temporal coverage and resolution, their data origin, and data source, and whether they are derived products. \*Global extent except the Antarctica; \*\* depending on time period.

instrumental in advancing our understanding of the climate system<sup>14</sup>, past environmental changes<sup>15</sup>, and the interactions between these processes<sup>16</sup>, as well as in improving future projections<sup>17</sup>.

Several projects have aimed to estimate climate data for the past, present, and future to improve predictions. For instance, the Paleo Model Intercomparison Project (PMIP) seeks to enhance climate models and their projections by analyzing and validating GCMs for different periods, such as the Last Glacial Maximum (LGM; ~21 ka BP) and the mid-Holocene (~6 ka BP). Although these projections were primarily designed for model evaluation, they have also been used in paleoecological studies through downscaling techniques to provide fine-scale paleoclimate reconstructions<sup>18</sup>. Similarly, the TraCE-21ka project used a coupled atmosphere-ocean GCM (Community Climate System Model version 3) to estimate global monthly climate data since the LGM, providing a foundation for other projects aiming to increase spatial resolution (e.g., PaleoView, CHELSA-TraCE21k) or the number of variables available (e.g., CHELSA-TraCE21k). Both PMIP and TraCE-21ka have been combined with modern climate datasets (e.g., WorldClim, CHELSA-TraCE21k) and future projections (e.g., CMIP) to create standardized climate datasets spanning the past, present, and future.

However, all these datasets have limitations. For example, most of them (e.g., PMIP3, TraCE-21ka, PaleoView) are available only at coarse spatial resolutions (Table 1), which can be problematic for regional or local studies where meso- and microclimates are critical. Other projects, such as WorldClim and EcoClimate, have high spatial resolution (~1 km) but limited temporal coverage, focusing on specific periods (e.g., LGM, mid-Holocene, 1970–2000, and 2100). This restricts their usefulness for studying continuous historical biogeographic patterns. CHELSA-TraCE21k addresses this limitation by providing data at 100-year intervals, although it does not retain the monthly temporal resolution of the original TraCE-21ka dataset. Regional analyses have produced finer resolution versions of these datasets, but they are often region-specific, require advanced programming skills, and are not always open access (but see<sup>19</sup>). The complexity of downscaling techniques can also deter non-experts from creating datasets at the desired resolution for specific regions.

Most downscaling efforts in paleoecology use the delta approach, also known as the standard change-factor method, due to its simplicity and speed<sup>20</sup>. However, this method relies on assumptions, such as orography not affecting changes in climate variables equally. More advanced methods, such as mechanistic models, simulate

physical atmospheric dynamics but are computationally intensive and require extensive data<sup>21</sup>. A compromise is the perfect prognosis approach, which uses observational data for both predictors (independent variables) and predictands (dependent variables) during calibration<sup>22</sup>. In this approach, observational data for predictands often comes from reanalysis projects that readjust model predictions through the integration of daily observations. Large-scale circulation variables, well represented by global models, are typically chosen as predictors. This approach assumes that the relationship between micro- and macroclimate is constant over time, making it more robust than the delta approach<sup>20</sup>. While perfect prognosis demands greater computational resources, it is more efficient than mechanistic downscaling in terms of cost and complexity.

The Mediterranean region, located between Africa, Asia, and Europe, is of great significance in biogeography due to its unique combination of diverse climatic conditions, varied topography, and long evolutionary history, which have contributed to its recognition as a biodiversity hotspot<sup>23</sup>. All these factors, and others like its role as a glacial refugia or long impact by human activities<sup>24</sup>, contribute to high levels of species richness, including a large number of endemic species<sup>25</sup>. The Western Mediterranean is particularly important, with a high rate of endemism<sup>26</sup>, and expected to be highly vulnerable to the impacts of climate change, such as increasing warming and reduced precipitation<sup>27</sup>. However, there is a lack of standardized, high-quality climate reconstructions for this region, especially at fine spatial and temporal scales, which are needed for studying past environmental changes and predicting future impacts.

This manuscript aims to develop a toolkit for applying advanced downscaling techniques to coarse-resolution climate datasets within the widely used R programming framework, specifically in ecology and paleoecology. This toolkit, in the form of two R-packages, will then be used to downscale climate data for the Western Mediterranean at the finest possible spatial and temporal resolution. While initially focused on the Western Mediterranean, the toolkit is designed to be adaptable to other regions globally. We aim to provide open access to a cutting-edge climate dataset for those interested in this region, and to promote the reuse of both the dataset and the toolkit to facilitate the creation of similar high-resolution climate products for other regions and scientific applications.

## Methods

To achieve the proposed objectives, two R packages were developed: *dsclim* (<https://dnietolugilde.com/dsclim/index.html>) and *dsclimtools* (<https://dnietolugilde.com/dsclimtools/index.html>). The former carries out the downscaling process, while the latter manages the resulting dataset and, hence, will be most useful for users interested only in working with our dataset. The *dsclim* package was used to select the optimal configuration of the downscaling algorithm by performing a cross-validation across multiple parameter spaces. It was subsequently used to downscale monthly climate reconstructions from the LGM (22 ka BP; 1000 years before the end of the LGM) to future projections (up to 2100 cal years), using data from the TraCE-21ka project, the Coupled Model Intercomparison Project 5 (CMIP5), and the UERRA reanalysis project. This process yielded a high-resolution dataset for the Western Mediterranean region, both spatially and temporally (DS-TraCE-21ka)<sup>28</sup>. The *dsclimtools* package was subsequently used to analyze the downscaled dataset, assessing various aspects of climate change in the region since the LGM until the end of the 21st century, including spatio-temporal patterns. Additionally, our downscaled data (DS-TraCE-21ka data)<sup>28</sup> were compared to the external CHELSA-TraCE21k dataset. Both packages are openly available in their respective GitHub repositories and documented through websites. These sites offer instructional guides on downscaling climate data and loading the dataset into R, with potential applications.

**Package development.** The developed packages integrate existing R packages (e.g., *ncdf4*, *loader*, *downscaleR*, *transformer*, etc.), rather than creating an entirely new framework. The *dsclim* package supports loading data from diverse sources, standardizing these data, and executing the downscaling process. On the other hand, the *dsclimtools* package simplifies the management and usage of downscaled data (DS-TraCE-21ka)<sup>28</sup>, allowing for filtering and combining datasets. By adhering to existing standard formats in R for spatio-temporal data (e.g., *stars* or *terra* packages), *dsclimtools* also facilitates statistical calculations (e.g., mean absolute deviation) and the derivation of bioclimatic variables using other R packages (e.g., *dismo*).

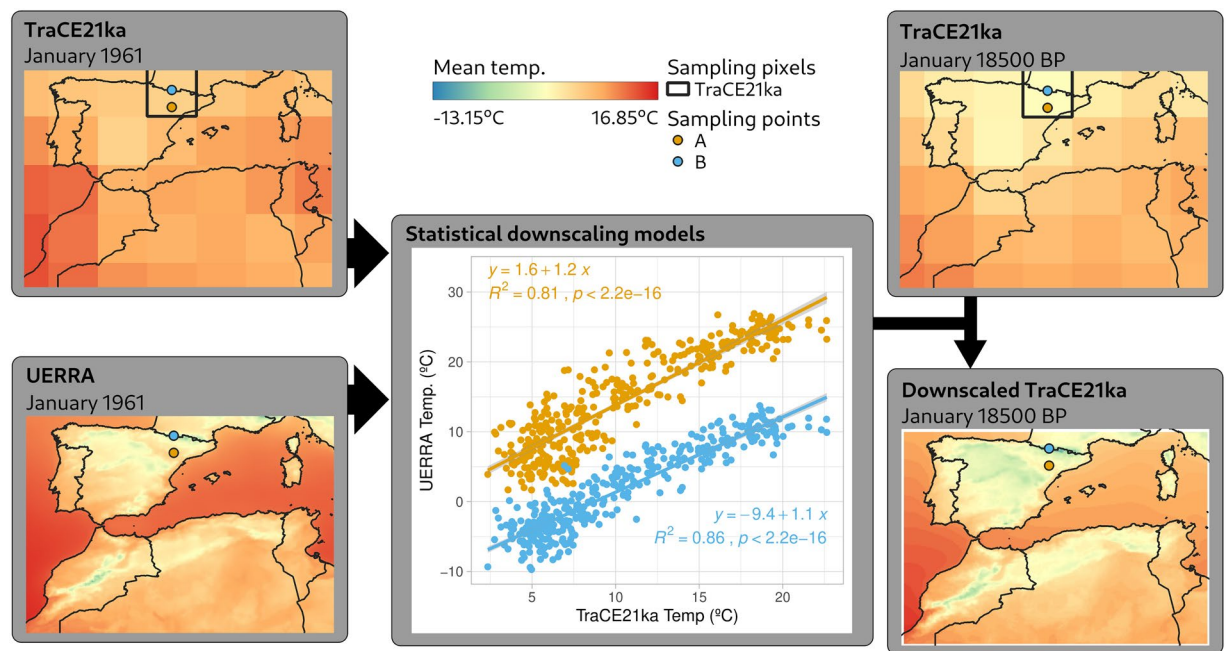
**Source climate data.** Climate data were retrieved from three sources to allow for an overlap period between past, present, and future periods at coarse resolution, as well as present data (i.e., historical period) at fine resolution. The database for the historical period was chosen for its superior spatial resolution. In all cases, identical variables were obtained across the three periods: maximum temperature, minimum temperature, average temperature, precipitation, wind speed, atmospheric pressure, relative air humidity, and cloud cover.

Paleoclimate data at coarse resolution were obtained from the TraCE-21ka paleoclimatic simulation<sup>29</sup>. This experiment provides a monthly reconstruction of climate since 22 ka BP until 1990, utilising the standard T31\_gx3 grid with a spatial resolution of approximately  $3.75^\circ \times 3.75^\circ$ . Further detailed information regarding the dataset, including the meaning of file names, variable names, and units, can be found in the following websites: <https://www.cgd.ucar.edu/projects/trace> and <https://www2.cgd.ucar.edu/ccr/strandwg/TraCE>. Further information can also be found on the documentation website for the model and the developing laboratory: <http://www.cesm.ucar.edu/projects/community-projects/LENS/data-sets.html> and <http://www.cesm.ucar.edu/models/atm-cam/docs/cam2.0/UsersGuide/UG-45.html>.

The reanalysis climate data for the historical period (1961–1990) were obtained from the UERRA project<sup>30</sup>. The project provides reanalysis data for Europe and North Africa from two different systems (UERRA-HARMONIE and MESCAN-SURFEX) that produce data in the Lambert Conformal Conic Grid at a temporal resolution of six hours and two spatial resolutions: The UERRA-HARMONIE system provides data at a resolution of  $11 \text{ km} \times 11 \text{ km}$ , while the MESCAN-SURFEX system offers data at a resolution of  $5.5 \text{ km} \times 5.5 \text{ km}$ .

Models	Institutions
Community Earth System Model, version 1 - Community Atmosphere Model, version 5 (CESM1-CAM5)	National Science Foundation, Department of Energy, National Center for Atmospheric Research (NSF-DOE-NCAR)
The Commonwealth Scientific and Industrial Research Organisation Mark version 3.6.0 (CSIRO-Mk3-6-0)	Commonwealth Scientific and Industrial Research Organisation - Queensland Climate Change Centre of Excellence (CSIRO-QCCCE)
Institut Pierre Simon Laplace Climate Model, version 5 A, Medium Resolution (IPSL-CM5A-MR)	Institut Pierre-Simon Laplace (IPSL)

**Table 2.** CMIP5 models and the institutions that provided model output used in this study.



**Fig. 1** Diagrammatic workflow for the perfect prognosis downscaling approach. The simple regressions displayed in the scatter plot represent an example of those used in the actual downscaling approach, since they only include a predictor variable while the actual downscaling was performed with multiple regressions including eight predictor variables combined in different ways (spatial, local and global predictors). Number of points in scatterplot is 360 (30 years  $\times$  12 months) for each spatial point (blue or orange).

The data can be obtained from the Climate Data Store on the Copernicus website (<https://cds.climate.copernicus.eu/datasets/reanalysis-uerra-europe-single-levels>). Further technical details may be found in the aforementioned Climate Data Store. The resolution of the precipitation variable (in the MESCAN-SURFEX) was aligned with that of UERRA-HARMONIE by calculating the mean value of the four pixels at 5.5 km, which were equivalent to each of the pixels at 11 km in the rest of variables. We opted in favor of UERRA against other data sources because of its superior spatial resolution (e.g. CRU-TS and ERA5), being a reanalysis rather than an interpolation (e.g. E-OBS), and/or being specially designed for the European and Mediterranean context rather than at a global scale (e.g. ERA5-Land).

The future climate data were obtained from the CMIP5 multi-model ensemble<sup>31</sup>, which are accessible from a variety of sources, including <https://pcmdi.llnl.gov/mips/cmip5/> and the Climate Data Store. The CMIP5 data were obtained for two distinct time periods. In terms of the historical period, this corresponds to the interval between 1850 and 2005 as defined by the CMIP5 project. In contrast, projections for future scenarios were made for the 2006–2100 interval. Despite the availability of data from multiple GCMs within the CMIP5 datasets, only three models (i.e., CESM1-CAM5, CSIRO-Mk3-6-0 and IPSL-CM5A-MR; see Table 2) possessed all the selected variables for all four representative pathways (RCPs; RCP2.6, RCP4.5, RCP6.0 and RCP8.5). Consequently, the variables were downloaded from these three GCMs and four RCPs.

**Downscaling.** A perfect prognosis downscaling process was carried out using the climate4R framework<sup>32</sup> for R<sup>33</sup>. The core of this statistical downscaling process is the construction of a transfer function between the predictand variables (i.e., fine resolution data) and a linear combination of the predictor variables (i.e., data at coarse resolution) for each pixel at fine resolution<sup>22</sup>. Accordingly, the model is calibrated using historical data at two distinct spatial resolutions, with a GCM estimated for each pixel at fine resolution (see Fig. 1 for a graphical illustration). Once the coefficients of the regression models have been calculated, the software uses past or future data at coarse resolution as input and applies the regression models to them, thereby obtaining past or future data at fine resolution. This process is repeated for each variable that is to be downscaled.



Model configuration	Spatial predictors	Global variables	Local variables	Number of pixels in local variables
M1.sp	True	All	None	
M21	False	All but focal variable	Focal variable	1
M21.sp	True	All but focal variable	Focal variable	1
M24	False	All but focal variable	Focal variable	4
M24.sp	True	All but focal variable	Focal variable	4
M31	False	None	All	1
M34	False	None	All	4

**Table 3.** Focal variable is the variable that is being downscaled.

The transfer function was built using Generalized Linear Models (GLMs) with a Gaussian link family for all variables except precipitation and wind speed, which were fitted according to a Gamma link family. Prior to the formal downscaling, a number of configurations of the downscaling algorithm were tested and evaluated. To this end, a cross-validation was conducted, whereby the historical period (1961–1990) was split into five-year intervals. In each cross-validation exercise, the models were fitted using 25 years of data, with the remaining five years used for model evaluation. Consequently, six models were fitted for each configuration. In particular, different configurations were tested in order to define the predictor variables (see Table 3). These included the use of ‘spatial predictors’ (i.e. the predictor variables were transformed and reduced using the main axis of a Principal Component Analysis that explain more than 70% of the original variation); the use of ‘global predictors’ (i.e. the application of values from all pixels in the study area at coarse resolution as predictors); and the use of ‘local predictors’ (i.e. the application of only values from a specific number of closest pixels as predictors). All tests were conducted using the full set of eight variables at the original, coarse resolution. The downsampled data from each configuration were evaluated by calculating the bias for each dataset, comparing the downsampled values (i.e., the estimated fine-resolution values for each pixel at a monthly time resolution) against the observed data from the UERRA dataset. In particular, three summary metrics of error were calculated for each pixel, representing the mean, standard deviation, and skewness of bias, respectively. The optimal configuration for downscaling the entire period was identified as M1.sp, in which spatial predictors were employed, and PCA axes explaining up to 70% of the variance were retained.

The resulting downsampled dataset was then compared with an external source of data, namely the CHELSA-TraCE21k dataset. As illustrated in Table 1, the CHELSA-TraCE21k project provides climate data at a spatial resolution of 30 arc sec from 21 ka to the present at one-hundred-year intervals. For purposes of comparison, these data were spatially aggregated to align with the resolution of our own dataset. A comparison between CHELSA-TraCE21k and our own dataset is of interest in order to identify any procedural bias or error present in any of the datasets. The information was extracted from the two datasets at two different pixels of the downsampled data (high spatial resolution), which were part of the same pixel of the TraCE-21ka dataset (coarse spatial resolution) and corresponded with two different orographic conditions (i.e. valley bottom and mountain range).

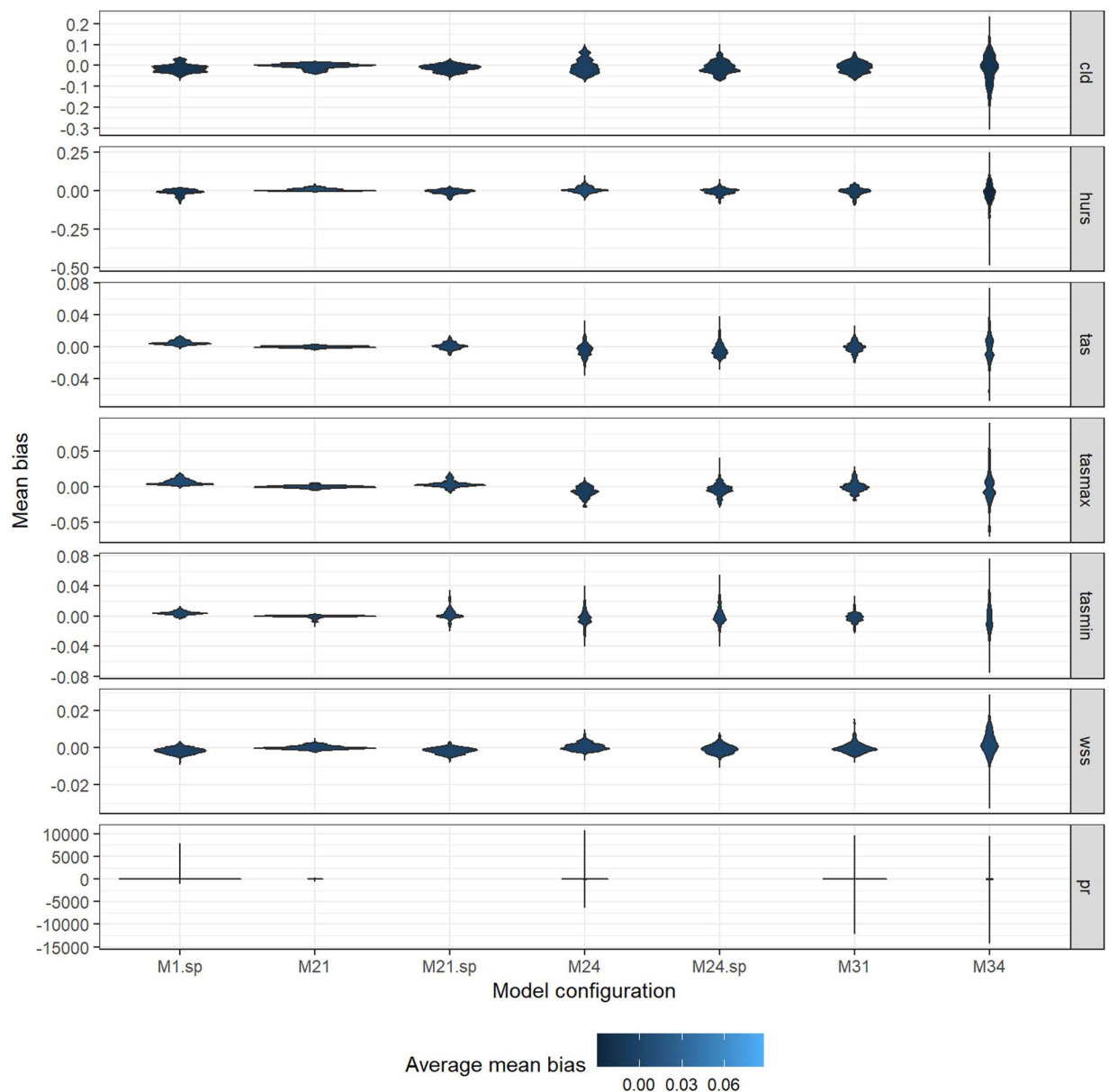
### Data Records

Our dataset is stored on Zenodo<sup>28</sup>, organized within the following folder structure. There is one folder for each climate change scenario (i.e., RCP 2.6, RCP 4.5, RCP 6.0, and RCP 8.5). Within each RCP folder, there are three subfolders corresponding to the GCMs used (i.e., CESM1-CAM5, CSIRO-Mk3-6-0, and IPSL-CM5A-MR), and each GCM folder has one subfolder per variable containing a NetCDF file for each year. On the other hand, the past data is fragmented per time and variable since it is heavier than the previously mentioned. All the folders have the same naming structure, the years range and the variable itself (e.g., –22000\_–16000\_BP\_pr, precipitation data from –22000 to –16000 BP).

The downsampled dataset spans the Western Mediterranean, including Southern Europe and North Africa, covering the geographical area defined by the coordinates 11.04° W to 12.04° E and 27.96° S to 44.04° N. It provides monthly climate data since 22 ka BP up to the year 2100. The variables included in the dataset are: maximum temperature (°C), minimum temperature (°C), average temperature (°C), precipitation (mm), wind speed (m/s), relative air humidity (percent), and cloud cover (fraction). All variables are presented at a spatial resolution of 11 × 11 km, ensuring high-resolution data for regional and local-scale analyses.

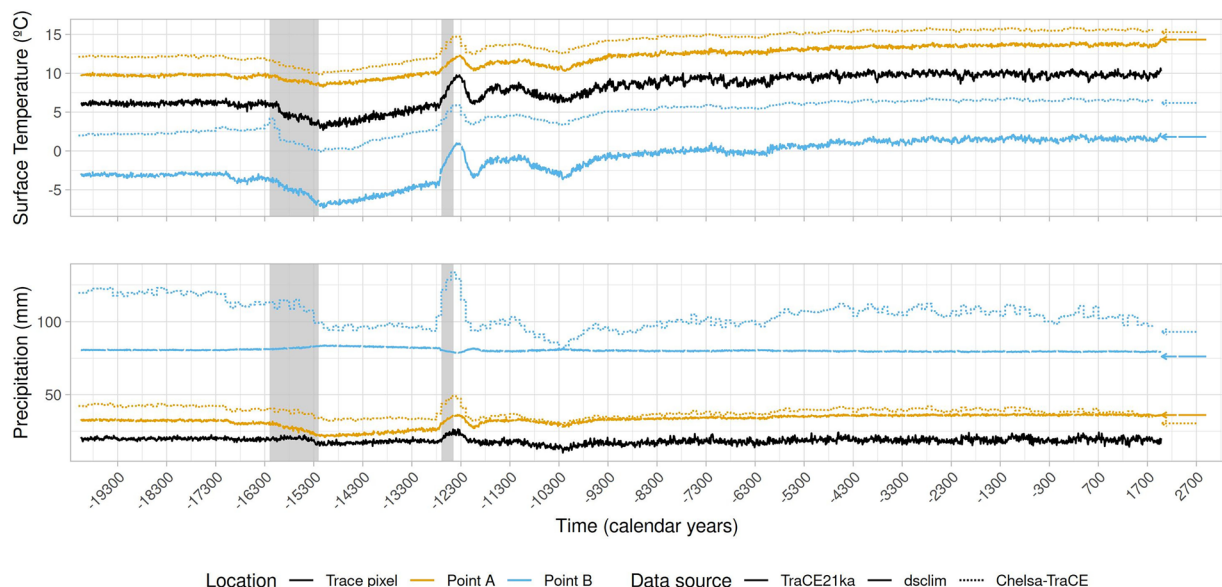
### Technical Validation

The cross-validation conducted for the downscaling of the historical period produced good results for all model configurations, showing consistently low mean bias across variables (Fig. 2). However, configurations without spatial predictors (M21, M24, M31, and M34) showed increased variability in both the standard deviation and skewness of the bias (Figures S1, S2 respectively). As a result, the simplest model (M1.sp) was selected to downscale all variables since the LGM until the year 2100. Furthermore, the spatial pattern of the bias from the simplest model (M1.sp) showed no artifacts at the borders of the TraCE-21ka pixels, suggesting that the bias from this downscaling configuration has a smoother spatial pattern. Although these evaluation metrics were favorable, it is important to acknowledge that the cross-validation relied on a relatively short historical period of only thirty years, while our downscaling approach aims to downscale thousands of years. Hence, this method’s main limitation is that uncertainties increase as the examined period deviates from the cross-validation period, requiring caution when interpreting the results and underscoring the need for external validation of the dataset.



**Fig. 2** Violin plots for mean bias in cross-validation of the downscaling of the historical period (1961–1990) along model configurations: M1.sp (uses spatial predictors but do not use global or local predictors), M21 (uses the focal variable as local predictors with 1 local pixel, but no spatial or global predictors), M21.sp (uses the focal variable as local predictor with 1 local pixel and spatial predictor, but no global predictors), M24 (uses the focal variable as local predictors with 4 local pixel, but no spatial or global predictors), M24.sp (uses the focal variable as local predictors with 4 local pixel and spatial predictors, but no global predictors), M31 (uses all variables as local predictors with 1 local pixel, but no spatial or global predictors), M34 (uses all variables as local variables with 4 local pixel, but no spatial or global predictors).

To circumvent the limitation of our cross-validation, we compared our dataset with the original TraCE-21ka and the CHLSA-TraCE21k dataset, the last one relying on the same low resolution climate data (i.e. TraCE-21ka) but with different high resolution contemporary climate data and downscaling approach. When comparing the average temperature values from the original TraCE-21ka pixel with two downscaled pixels (both within the same TraCE-21ka original pixel) and the same two pixels from the CHLSA-TraCE21k data, two main conclusions emerged (Fig. 3). First, the higher resolution of both downscaled datasets enabled the differentiation between points with varying topographies; both downscaled datasets successfully captured the higher temperatures in the valley and the lower temperatures in the mountainous region. Second, both approaches agreed on the distinct dynamics observed between the two points, where mountain areas showed less variation over time compared to the valley. This alignment between the two datasets confirms that fine-resolution data can capture topographic effects on average temperature, providing valuable insights for biogeographical studies, especially at micro- or meso-scales. As the time series fluctuations align well (e.g., Bølling–Allerød, Younger Dryas), we can

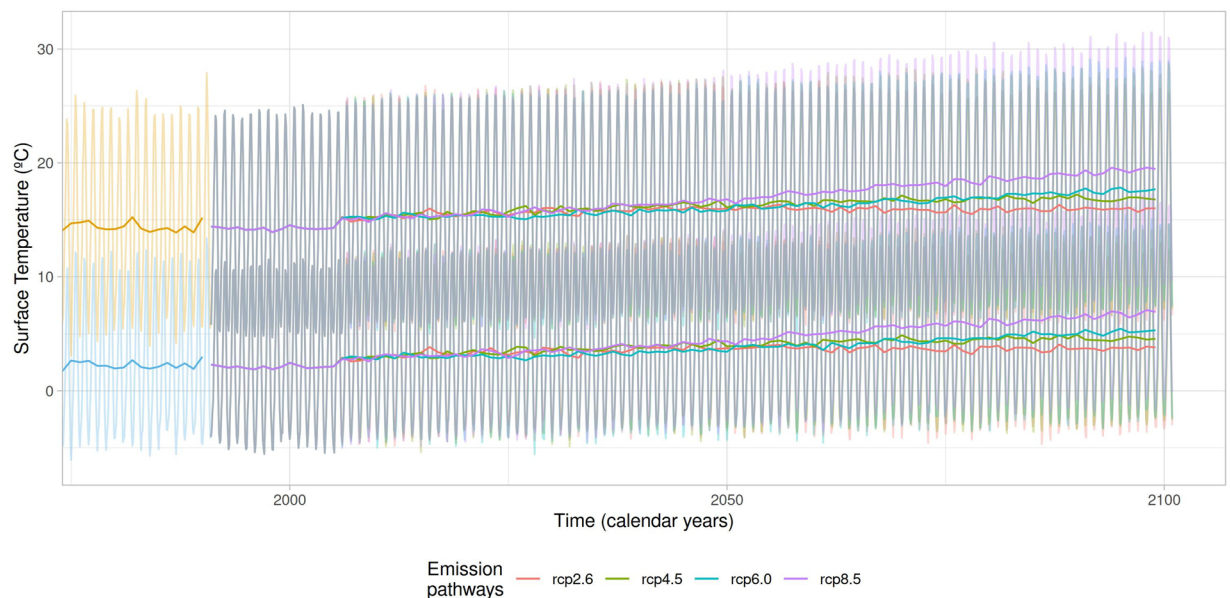


**Fig. 3** Comparison of temperature and precipitation as reconstructed by -CHLSA-TraCE21K and DS-TraCE-21ka (our dataset of the downscaled TraCE-21ka). Black line is the temperature/precipitation as in the original TraCE-21ka for the coarse resolution pixel highlighted in Fig. 1. Orange and blue lines are the temperatures/precipitations in the orange and blue points, respectively, in Fig. 1 according to CHLSA-TraCE21k (dashed line) and DS-TraCE-21ka (continuous line). Arrows indicates the mean historical temperature/precipitation in the original data sources at fine resolution (UERRA for DS-TraCE-21ka and Present CHLSA for CHLSA-TraCE21k). Grey is representing two periods of rapid climate change in the series, cooling and warming, that are used to illustrate uses of the dataset.

consider the downscaled dataset externally validated, eliminating the possibility of significant procedural errors. However, some differences were noted between our downscaled data and the CHLSA-TraCE21k data. In the case of the warmer region, our dataset shows temperatures 2–3 °C lower than CHLSA-TraCE21k, while in the colder region, our dataset is around 5 °C lower. These differences are already present in the original data used for the historical period in both downscaling procedures (CHLSA data for CHLSA-TraCE21k and UERRA data for our downscaling). Consequently, these discrepancies reflect limitations in the fine-resolution source data, rather than any bias introduced during downscaling. This consistency raises concerns about mixing downscaled datasets with different baselines. It also suggests that similar patterns could arise if we used a different data source for the high spatial resolution data in the historical period (calibrating phase; e.g. E-OBS or ERA5-Land), with uncertainties arising from the differences between data sources. Further venues of research could explore these differences to provide better assessments of uncertainties due to data sources.

The precipitation comparison (Fig. 3), on the other hand, revealed some other discrepancies. Both datasets suggested lower precipitation in the valley than in the mountainous pixel, but the precipitation values in both datasets exceeded those of the original TraCE-21ka dataset at coarse resolution. This discrepancy could arise from fine-resolution pixels with exceptionally low precipitations within the corresponding coarse-resolution pixel. Additionally, CHLSA-TraCE21k exhibited greater precipitation variability than our dataset and TraCE-21ka, particularly in the mountain. However, past studies based on soil development have documented more stable precipitation regimes in floodplains<sup>34</sup>. Given the challenges of accurately estimating precipitation at global scale, this higher variability in the CHLSA-TraCE21k is not unexpected. It is noteworthy that the precipitation trend observed in the mountainous pixel conflicts with the general trend observed in the coarse resolution pixel. This is not entirely surprising since precipitation is often influenced by local processes, such as convection or rain shadow effects, which can cause variations at a finer resolution. In this case, downscaling at a regional scale could be an advantage to reconstruct paleoprecipitation regimes compared to global approaches (i.e., CHLSA-TraCE-21k), although we cannot confirm it. Unfortunately, literature on the topic doesn't help to support the trends observed by the two approaches; discrepancies in precipitation trends at local scale during the last deglaciation have been reported to range from reduced precipitation to no change, or even highly spatially variable anomalies<sup>35</sup>. Given the higher variability in patterns and uncertainties in downscaling of precipitation, the choice of the high spatial resolution data for the historical period could have a greater impact than for temperature, which reinforces the need for further research exploring the uncertainties associated with data sources (e.g. ERA5-Land or E-OBS).

When incorporating future data alongside past data, both datasets demonstrated a coherent progression, as they were based on the same baseline (UERRA reanalysis, see Fig. 4). This smooth transition minimizes biases that might arise from merging different datasets. The consistency between historical reconstructions and future projections is critical for analyzing long-term climate trends. Notably, all future scenarios, regardless of emission



**Fig. 4** Temperatures for the same two points of the study area for the calibration period (1960–1990) and the average future projections among three climate models (e.g. CESM1-CAM5, CSIRO-Mk3-6-0, and IPSL-CM5A-MR) under four different Representative Concentration Pathways (RCPs; rcp2.6, rcp4.5, rcp6.0, rcp8.5).

levels, followed this trajectory, lending further validity to our methodological approach and allowing for a robust comparison of future climate projections with historical data.

### Code availability

Code can be found in both the Github repository <https://github.com/dromera2/DS-TraCE-21ka> and the Zenodo repository<sup>36</sup>.

Received: 22 January 2025; Accepted: 23 April 2025;

Published online: 06 May 2025

### References

- Musso, A., Ketterer, M. E., Greinwald, K., Geitner, C. & Egli, M. Rapid decrease of soil erosion rates with soil formation and vegetation development in periglacial areas. *Earth Surface Processes and Landforms* **45**, 2824–2839 (2020).
- Harrison, S., Spasojevic, M. J. & Li, D. Climate and plant community diversity in space and time. *Proceedings of the National Academy of Sciences* **117**, 4464–4470 (2020).
- Grimm, N. B. *et al.* The impacts of climate change on ecosystem structure and function. *Frontiers in Ecology and the Environment* **11**, 474–482 (2013).
- Din, M. S. U. *et al.* World Nations Priorities on Climate Change and Food Security. in *Building Climate Resilience in Agriculture: Theory, Practice and Future Perspective* (eds. Jatoti, W. N. *et al.*) 365–384. [https://doi.org/10.1007/978-3-030-79408-8\\_22](https://doi.org/10.1007/978-3-030-79408-8_22) (Springer International Publishing, Cham, 2022).
- Overpeck, J. T., Meehl, G. A., Bony, S. & Easterling, D. R. Climate Data Challenges in the 21st Century. *Science* **331**, 700–702 (2011).
- Finn, D. How do we build collaborative science for better urban planning and climate change adaptation? **2020**, SY001-06 (2020).
- Vrontisi, Z. *et al.* Macroeconomic impacts of climate change on the Blue Economy sectors of southern European islands. *Climatic Change* **170**, 27 (2022).
- Alba-Sánchez, F. *et al.* Past and present potential distribution of the Iberian Abies species: a phytogeographic approach using fossil pollen data and species distribution models. *Diversity and Distributions* **16**, 214–228 (2010).
- Hansen, A. J. *et al.* Global Change in Forests: Responses of Species, Communities, and Biomes: Interactions between climate change and land use are projected to cause large shifts in biodiversity. *BioScience* **51**, 765–779 (2001).
- Varol, T., Cetin, M., Ozel, H. B., Sevik, H. & Zeren Cetin, I. The Effects of Climate Change Scenarios on Carpinus betulus and Carpinus orientalis in Europe. *Water Air Soil Pollut* **233**, 45 (2022).
- Abel-Schaad, D. *et al.* Are Cedrus atlantica forests in the Rif Mountains of Morocco heading towards local extinction? *The Holocene* **28**, 1023–1037 (2018).
- Urban, M. C. *et al.* Improving the forecast for biodiversity under climate change. *Science* **353**, aad8466 (2016).
- Grassl, H. Status and Improvements of Coupled General Circulation Models. *Science* **288**, 1991–1997 (2000).
- Janson, L. & Rajaratnam, B. A Methodology for Robust Multiproxy Paleoclimate Reconstructions and Modeling of Temperature Conditional Quantiles. *Journal of the American Statistical Association* **109**, 63–77 (2014).
- Maguire, K. C., Nieto-Lugilde, D., Fitzpatrick, M. C., Williams, J. W. & Blois, J. L. Modeling Species and Community Responses to Past, Present, and Future Episodes of Climatic and Ecological Change. *Annual Review of Ecology, Evolution, and Systematics* **46**, 343–368 (2015).
- Svenning, J.-C., Eiserhardt, W. L., Normand, S., Ordóñez, A. & Sandel, B. The Influence of Paleoclimate on Present-Day Patterns in Biodiversity and Ecosystems. *Annual Review of Ecology, Evolution, and Systematics* **46**, 551–572 (2015).
- Macleán, I. M. D. Predicting future climate at high spatial and temporal resolution. *Global Change Biology* **26**, 1003–1011 (2020).
- Lima-Ribeiro, M. S. *et al.* EcoClimate: a database of climate data from multiple models for past, present, and future for macroecologists and biogeographers. *Biodiversity Informatics* **10** (2015).
- Lorenz, D. J., Nieto-Lugilde, D., Blois, J. L., Fitzpatrick, M. C. & Williams, J. W. Downscaled and debiased climate simulations for North America from 21,000 years ago to 2100AD. *Sci Data* **3**, 160048 (2016).
- Wilby, R. *et al.* Guidelines for Use of Climate Scenarios Developed from Statistical Downscaling Methods. (2004).



21. Ekström, M., Grose, M. R. & Whetton, P. H. An appraisal of downscaling methods used in climate change research. *WIREs Climate Change* **6**, 301–319 (2015).
22. Manzanas, R. *et al.* Statistical Downscaling in the Tropics Can Be Sensitive to Reanalysis Choice: A Case Study for Precipitation in the Philippines. *Journal of Climate* **28**, 4171–4184 (2015).
23. Comes, H. P. The Mediterranean Region: A Hotspot for Plant Biogeographic Research. *The New Phytologist* **164**, 11–14 (2004).
24. Blondel, J. The ‘Design’ of Mediterranean Landscapes: A Millennial Story of Humans and Ecological Systems during the Historic Period. *Hum Ecol* **34**, 713–729 (2006).
25. Abellán, P. & Svenning, J.-C. Refugia within refugia – patterns in endemism and genetic divergence are linked to Late Quaternary climate stability in the Iberian Peninsula. *Biological Journal of the Linnean Society* **113**, 13–28 (2014).
26. Cheikh Albassatneh, M. *et al.* Spatial patterns of genus-level phylogenetic endemism in the tree flora of Mediterranean Europe. *Diversity and Distributions* **27**, 913–928 (2021).
27. Cos, J. *et al.* The Mediterranean climate change hotspot in the CMIP5 and CMIP6 projections. *Earth System Dynamics* **13**, 321–340 (2022).
28. Romera-Romera, D., Alba-Sánchez, F., Abel-Schaad, D. & Nieto-Lugilde, D. DS-TraCE-21ka. *Zenodo* <https://doi.org/10.5281/zenodo.14067329> (2025).
29. Feng, H. Simulating Transient Climate Evolution of the Last deglaciation with CCSM3, Doctor of Philosophy, Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, WI, USA, 161 pp. (2011).
30. Copernicus Climate Change Service, Climate Data Store. UERRA regional reanalysis for Europe on single levels from 1961 to 2019. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). <https://doi.org/10.24381/cds.32b04ec5> (2019).
31. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society* **93**, 485–498 (2012).
32. Iturbide, M. *et al.* The R-based climate4R open framework for reproducible climate data access and post-processing. *Environmental Modelling & Software* **111**, 42–54 (2019).
33. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing (2022).
34. Casana, J. Mediterranean valleys revisited: Linking soil erosion, land use and climate variability in the Northern Levant. *Geomorphology* **101**, 429–442 (2008).
35. Beghin, P. *et al.* What drives LGM precipitation over the western Mediterranean? A study focused on the Iberian Peninsula and northern Morocco. *Clim Dyn* **46**, 2611–2631 (2016).
36. Romera-Romera, D. dromera2/DS-TraCE-21ka: Fine-Spatio-Temporal and Long-Term Climate Data for the Western Mediterranean in Ecology and Biogeography. *Zenodo* <https://doi.org/10.5281/zenodo.15023379> (2025).

## Acknowledgements

We acknowledge the World Climate Research Programme’s Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table 2 of this paper) for producing and making available their model output. For CMIP the U.S. Department of Energy’s Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. UCO pre-doctoral contracts of the Own Plan of Research of the University of Córdoba. [Romera-Romera]. This publication is part of the project I + D + i TED2021-132631B-I00 funded by MICIU/AEI/ 10.13039/501100011033 and by NextGenerationEU/PRTR. Grant TED2021-130133B-I00 to Diego Nieto Lugilde through Conserva3 project, funded by MCIN/AEI/10.13039/501100011033 and the “European Union NextGenerationEU/PRTR”. Grant to Diego Nieto Lugilde and Daniel Romera Romera through the project PID2022-140794NB-I00 funded by MICIU/AEI/ 10.13039/501100011033 and by FEDER, EU.

## Author contributions

Daniel Romera-Romera and Diego Nieto-Lugilde: Conceptualization (equal); Formal analysis (lead); Software (lead); Writing – original draft (lead). Francisca Alba-Sánchez and Daniel Abel-Schaad: Conceptualization (equal); Writing – review & editing (equal); Funding acquisition (equal).

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41597-025-05067-9>.

**Correspondence** and requests for materials should be addressed to D.R.-R.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025