ECOGRAPHY

Software Note

chelsa-cmip6 1.0: a python package to create high resolution bioclimatic variables based on CHELSA ver. 2.1 and CMIP6 data

Dirk Nikolaus Karger, Yohann Chauvier and Niklaus E. Zimmermann

Swiss Federal Research Inst. WSL, Switzerland

Correspondence: Dirk Nikolaus Karger (dirk.karger@wsl.ch)

Ecography
2023: e06535
doi: 10.1111/ecog.06535

Subject Editor: Thiago F. Rangel
Editor-in-Chief:
Christine N. Meynard
Accepted 14 February 2023

Data on projected climate conditions for the future is essential for many applications in climate change impact studies. Yet, with the release of the CMIP6, the increasing amount of data poses challenges for users to access and process the data. Here we present the chelsa_cmip6 package to create bioclimatic variables and climatological normals based on CHELSA ver. 2.1 and cloud based CMIP6 data for user defined geographical extents. The package offers simple access to CMIP6 data on cloud servers. It aggregates the CMIP6 data to climatological normals and downscales it from its native resolution to a 30 arcsec grid resolution using the delta change method for user defined future periods. Based on the climatological normals, it creates a set of bioclimatic variables that represent long term climatic means or variability in air temperatures and precipitation. The package offers users a simple way of creating climate change projections for user defined geographical extents and time periods.

Keywords: bioclim, bioclimatic variables, climate projections, climatologies, global climate model, socio-economic pathway

Background

With the release of the Coupled Model Intercomparison Project Phase 6 (CMIP6) (Eyring et al. 2016) a large amount of new data became available for studying the effects of projected climate change on future ecosystems. Compared to CMIP5, CMIP6 involves a larger number of participating research institutions, climate models, socio-economic scenarios, and experiments (O'Neill et al. 2016). While CMIP5 included the output of 20 different climate models, CMIP6 will include the output of approximately 100 climate models. The number of different institutions involved increased from 20 to 49 (wcrp-cmip.github.io/CMIP6_CVs/docs/CMIP6_institution_id.html), and CMIP6 now also includes eight shared socioeconomic pathways (SSPs) (Gidden et al. 2019) instead of the four representative concentration pathways (RCPs) that were used in CMIP5 (van Vuuren et al. 2011).

The output from the different climate models is generally at too coarse spatial resolution, compared to the spatial details needed by most climate impact studies, such as

NORDIC SOCIETY OIKOS

www.ecography.org

ecological studies on organisms that are often implemented at a 1 km resolution or finer (Hurlbert and Jetz 2007). To produce high resolution climate data for impact studies, various approaches exist (Maraun et al. 2010) that mainly fall into two categories: 1) dynamical downscaling of global climate models using regional climate models (RCMs), and 2) statistical downscaling.

Dynamically downscaling of global climate model output using RCMs (Sørland et al. 2021) is usually computationally demanding and requires in-depth knowledge of RCMs and their application. Applying a dynamical downscaling is therefore often impractical if a large ensemble of models is required at high spatial resolution over large geographical extents. In addition, many applications investigating the impact of climate e.g. on biological systems do not require the full range of variables that RCMs can provide, but often only rely on a limited set of variables (Frieler et al. 2017). Biological studies often only use so-called bioclimatic variables (Hijmans et al. 2005, Karger et al. 2017), which describe means or variabilities in climate over long periods of time (usually 30 year climatological normals).

These variables can be created computationally more efficiently by a statistical downscaling, using the so-called delta change method, or climatologically informed interpolation of anomalies (Hay et al. 2000). This method calculates the anomaly for a climatic variable between a specific reference period (e.g. the climatological normals for 1981–2010) in a CMIP6 model and a future period (e.g. the climatological normals for 2041–2070) of the same climate model (Fig. 1). This anomaly is then added to (or multiplied with) a high resolution dataset of climatological normals (e.g. Karger et al. 2017) that has the same contemporary reference period. This method is computationally very efficient compared to dynamical downscaling and additionally relatively robust towards the inherent bias of the models. Delta change corrected CMIP models have long become a standard in macroecological research (Hijmans et al. 2005, Karger et al. 2017). The CHELSA (Karger et al. 2017, 2020, 2021b) and WorldClim datasets (Hijmans et al. 2005) have indeed already been providing delta change downscaled CMIP output, contributing largely to the field of macroecological research over the last decade. However, it is very demanding to provide downscaled CMIP6 data preprocessed for all climate models and SSPs due to the large number of possible combinations. Furthermore, both CHELSA and WorldClim only provide data for a limited amount of periods (e.g. 2011–2040, 2041–2070, 2071–2100 in CHELSA ver. 2.1), which might not always overlap with the time periods needed in specific studies. Finally, these currently available high resolution delta change-based CMIP6 datasets solely provide a full global extent of climate scenarios, which many studies often do not need, as these often only require data for a limited geographical extent. Therefore, a more flexible and automated approach to acquire delta change-downscaled CMIP6 data is needed. Such an approach should allow for specifically selecting the

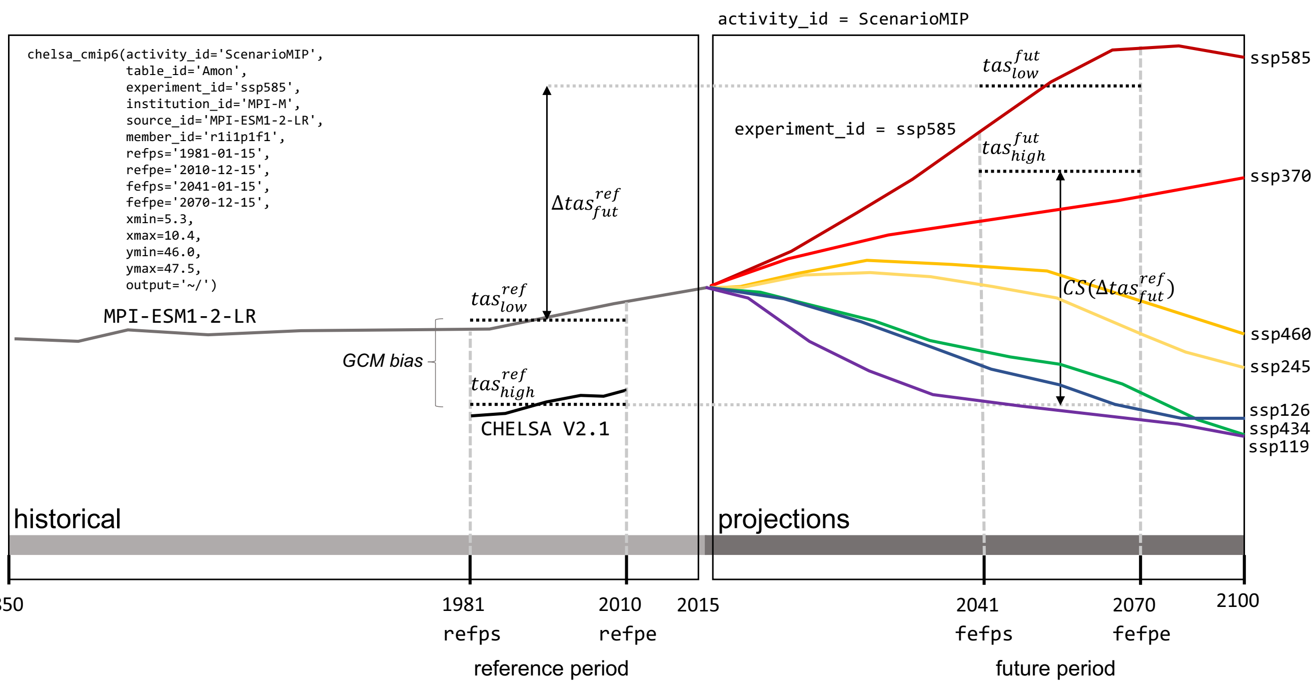

Figure 1. Example of the delta change method applied on the model MPI-ESM1-2-LR for ssp585, the reference period 1981–2010, and the future period 2041–2070 (Example 1). The MPI-ESM1-2-LP gives the low resolution reference period temperature tas^{ref}_{low} and the low resolution future period temperature tas^{fut}_{low}. The difference between these two temperatures is interpolated to 30 arcsec resolution using a qubic-spline $CS(\Delta tas^{ref}_{fut})$ and then added to the high resolution reference temperature tas^{ref}_{high} to get tas^{fut}_{high}. The parameters that need to be set to achieve the shown delta change downscaling are shown in the upper left corner. Both tas^{ref}_{high} and tas^{fut}_{high} are given as output for a specific region that is specified using the arguments: `--xmin 5.3 --xmax 10.4 --ymin 46.0 --ymax 47.5` (see Fig. 2).

desired geographical extent, climate model, SSPs, and time period in order to facilitate the easy provisioning of state of the art climate data for ecological climate impact research.

A problem with such a flexible and automated approach is in the size and quantity of CMIP6 data needed as input. End users often neither have the storage capacity nor the computational power to store and process all the input data from CMIP6 locally. Several approaches have been developed recently to allow access of native CMIP6 data through cloud servers, and to perform operations on them using lazy loading methods. For instance, the Pangeo initiative (https://pangeo-data.github.io/pangeo-cmip6-cloud/accessing_data.html) provides cloud access to CMIP6 model output in the Zarr format. This format stores data and metadata as a collection of files in compressed chunks that can be read individually or in parallel from the cloud without the need of downloading it. The metadata of these files can then be processed with python tools such as *xarrays* (Hoyer et al. 2022), which may remotely connect to a file and perform operations via lazy loading. Zarr files, similar to the original netCDF files of CMIP6, are chunked geographically, meaning that only parts of the data that encompass a user-defined geographical extent need to be downloaded.

Here, we present the 'chelsa-cmip6' python package that builds up on these new developments of cloud access to CMIP6 data and provides a convenient way to create bioclimatic data for any available model, SSP, experiment, or time period covered in CMIP6. It uses *xarrays* to remotely access CMIP6 output from the Pangeo file cloud server and only provides climate data for the geographical extent a user is interested in, minimizing memory demands, data transfer, and storage. The package uses the delta change method with the CHELSA ver. 2.1 1981–2010 climatological normals of precipitation and mean, minimum and maximum temperatures (Karger et al. 2021a) as high resolution reference dataset to create future climatological normals for these variables for any given time period. It then uses these climatological normals to calculate a set of bioclimatic variables for both the reference period and the future time period.

Methods and features

Delta change method – climatologically aided interpolation of anomalies

The delta change method applied here (Fig. 1) uses the long-term climatological normals of a reference period (*ref*) and calculates the anomaly (Δ) towards a chosen future long-term climatological normal period (*fut*) for a given variable (Arnell et al. 2001). For mean daily 2 m air temperatures (*tas*), as well as daily maximum (*tasmax*) and minimum (*tasmin*) 2 m air temperatures, an additive delta change method is applied so that the anomaly is given by (Eq. 1):

$$\Delta tas_{fut}^{ref} = tas_{low}^{fut} - tas_{low}^{ref} \qquad (1)$$

This calculation is performed at the native resolution of the CMIP6 model grid (*low*). The anomaly Δtas_{fut}^{ref} is then interpolated to the high 30 arcsec. resolution of a high-resolution (*high*) climatology using a cubic-spline interpolation (*CS*) (Hall and Meyer 1976) and added to the high resolution reference climatology (tas_{high}^{ref}), so that (Eq. 2):

$$tas_{high}^{fut} = tas_{high}^{ref} + CS(\Delta tas_{fut}^{ref}) \qquad (2)$$

The same applies for *tasmax* and *tasmin*. This calculation is done separately for each month from January to December.

For precipitation (*pr*), an additive delta change method would generate values that could be potentially negative. The package therefore uses a multiplicative method on the derived anomalies. By adding a constant of 0.01 kg m^{-2} day^{-1} to both the reference and the future data to avoid division by zero (Eq. 3).

$$\Delta pr_{fut}^{ref} = \frac{pr_{low}^{fut} + 0.01\, kg\, m^{-2} day^{-1}}{pr_{low}^{ref} + 0.01\, kg\, m^{-2} day^{-1}} \qquad (3)$$

Similar to temperatures, the anomalies are then interpolated using a cubic-spline interpolation to a 30 arcsec. resolution and multiplied to a high-resolution reference climatology using (Eq. 4):

$$pr_{high}^{fut} = pr_{high}^{ref} * CS(\Delta pr_{fut}^{ref}) \qquad (4)$$

The delta change method as applied here is relatively insensitive regarding individual model bias of the GCM used, as it only uses the difference (ratio) for a given variable between a reference period and a future period.

Structure of the chelsa_cmip6 package

The 'chelsa-cmip6' package consists of three distinct parts. The main function *chelsa_cmip6.py* that calls both the *GetClim.py* module, and the *BioClim.py* module (Fig.2).
The GetClim.py module connects to remote CMIP6 data and calculates climatological normals
The *GetClim.py* module provides functions to create monthly climatologies from climate simulation data from CMIP6 using climate observation data from CHELSA ver. 2.1 at a 30 arcsec. grid resolution for a given area of choice. The *GetClim.py* module contains functions to connect to CMIP6 data via the Google cloud storage provided by the Pangeo initiative and to read the data into *xarrays*. It also creates monthly climatological normals in CMIP6 time series using the delta change method for a given time period that can be chosen by the user.

The BioClim.py module calculates bioclimatic variables based on climatological normals

The *BioClim.py* module calculates various bioclimatic parameters from climatological data. 19 of these bioclimatic

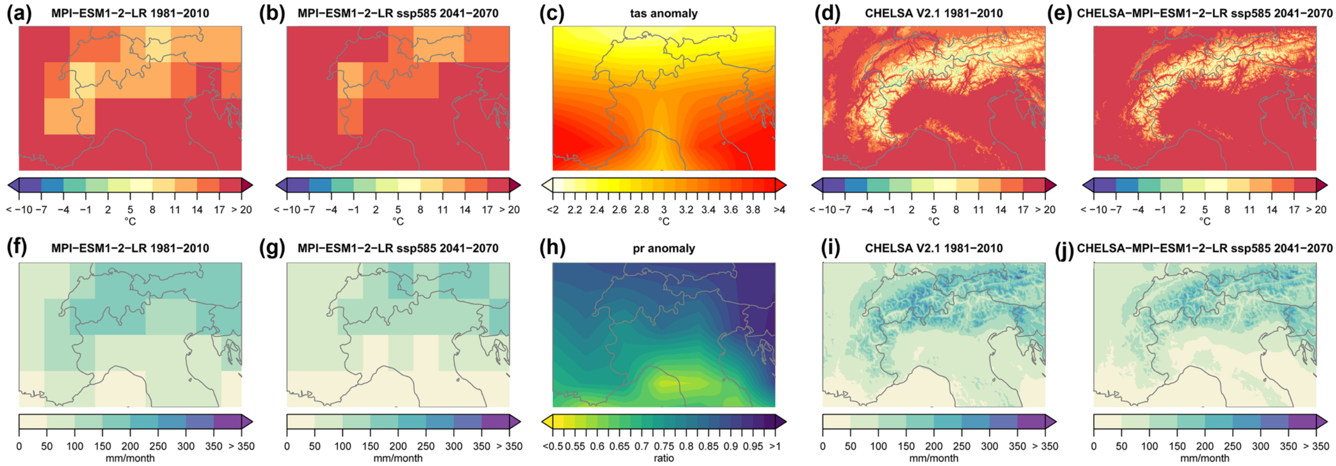

Fig.2 Spatial example of the delta change method using the MIP-ESM1-2-LR model for ssp585 in central Switzerland for the future time period 2041–2070 (Example 1). (a) Shows the CMIP6 model mean monthly 2 m air temperature data for June on its native resolution for the reference period 1981–2010, (b) shows the same model for the future period 2041–2070 and (c) the air temperature anomaly. This anomaly is added to the high resolution temperature climatology of the reference period from CHELSA ver. 2.1 (d) resulting in the high resolution future data for the period 2041–2070 (e). Panels (f)–(j) show the same process for July precipitation, only that the ratio instead of the absolute difference is used in this case.

variables originate from ANUCLIM (Xu and Hutchinson 2013) and WorldClim (Hijmans et al. 2005) and describe a range of bioclimatic conditions that are commonly used in biological and ecological research. Additionally, the package provides a calculation of growing degree days.

Bioclimatic variables based on annual quarters (Q) use a definition of a quarter as a running three-month interval. For each month, the pro-, and preceding month is used together with the focal month. At the end of the year (December) and the beginning of the year (January) a quarter is defined using the months in the beginning of the year (e.g. when the focal month is December, it uses (November, December, January).

The bioclimatic variables included in *chelsa_cmip6* are defined by:

bio1 – Mean annual temperature. The average daily 2 m air temperature over all 12 months. It is given by (Eq. 5):

$$bio1 = \frac{\sum_{i=1}^{12} tas_i}{12} \tag{5}$$

bio2 – Annual mean diurnal range. The mean of the monthly temperature ranges between all months. It is given by (Eq. 6):

$$bio2 = \frac{\sum_{i=1}^{12} (tasmax_i - tasmin_i)}{12} \tag{6}$$

bio3 – Isothermality gives the ratio between the day-to-night temperatures range relative to the summer-to-winter temperature range. It is given by (Eq. 7):

$$bio3 = \frac{bio2}{bio7} \tag{7}$$

bio4 – Temperature seasonality – Quantifies the variation of temperature over the year and is given by the standard deviation (SD) of *tas* over all 12 months (Eq. 8):

$$bio4 = SD\{tas_1, \ldots, tas_{12}\} \tag{8}$$

bio5 – Maximum temperature of the warmest month, given by (Eq. 9):

$$bio5 = max(\{tasmax_1, \ldots, tasmax_{12}\}) \tag{9}$$

bio6 – Minimum temperature of the coldest month, given by (Eq. 10):

$$bio6 = min(\{tasmin_1, \ldots, tasmin_{12}\}) \tag{10}$$

bio7 – Annual temperature range. The total range of temperatures over a year, given by (Eq. 11):

$$bio7 = bio5 - bio6 \tag{11}$$

bio8 – Mean temperature of the wettest quarter. This variable gives the mean temperature during the wettest three-month period of the year. First the precipitation of the wettest quarter is determined using Eq. 20. We name i^* the first month of the wettest quarter identified in Eq. 20. Then the mean temperature for the wettest quarter is calculated by (Eq. 12):

$$bio8 = \frac{\sum_{i^*}^{i^*+3} tas_{i^*}}{3} \tag{12}$$

bio9 – Mean temperature of the driest quarter. This variable gives the mean temperature during the driest three-month period of the year. First the driest quarter is determined using the Eq. 21. Here i^* gives the first month of the driest quarter identified in Eq. 21. Then the mean temperature for the driest quarter is calculated similar to bio8 by (Eq. 13):

$$bio8 = \frac{\sum_{i^*}^{i^*+3} tas_{i^*}}{3}$$

(13)

bio10 – Mean temperature of the warmest quarter. This variable gives the mean temperature during the warmest three-month period of the year by (Eq. 14):

$$bio10 = max \left(\left| \frac{\sum_{i=1}^{3} tas_i}{3}, \frac{\sum_{i=2}^{4} tas_i}{3}, \ldots, \frac{\sum_{i=10}^{12} tas_i}{3}, \frac{\sum_{i=11}^{1} tas_i}{3}, \frac{\sum_{i=12}^{2} tas_i}{3} \right| \right) \quad (14)$$

bio11 – Mean temperature of the coldest quarter. This variable gives the mean temperature during the coldest three-month period of the year by (Eq. 15):

$$bio11 = min \left(\left| \frac{\sum_{i=1}^{3} tas_i}{3}, \frac{\sum_{i=2}^{4} tas_i}{3}, \ldots, \frac{\sum_{i=10}^{12} tas_i}{3}, \frac{\sum_{i=11}^{1} tas_i}{3}, \frac{\sum_{i=12}^{2} tas_i}{3} \right| \right) \quad (15)$$

bio12 – Annual precipitation – The sum of all precipitation (liquid and solid) that fell within one year, given by (Eq. 16):

$$bio12 = \sum_{i=1}^{12} pr_i \quad (16)$$

bio13 – Precipitation of the wettest month, given by (Eq. 17):

$$bio13 = max\left(\left\{ pr_1, \ldots, pr_{12} \right\}\right) \quad (17)$$

bio14 – Precipitation of the driest month, given by (Eq. 18):

$$bio14 = min\left(\left\{ pr_1, \ldots, pr_{12} \right\}\right) \quad (18)$$

bio15 – Precipitation seasonality. variation in monthly precipitation over the year given by the coefficient of variation CV (Eq. 19):

$$bio15 = CV\left(\left\{ pr_1, \ldots, pr_{12} \right\}\right) \quad (19)$$

bio 16 – Precipitation of the wettest quarter given by (Eq. 20):

$$bio16 = max \left(\left| \sum_{i=1}^{3} pr_i, \sum_{i=2}^{4} pr_i, \ldots, \sum_{i=10}^{12} pr_i, \sum_{i=11}^{1} pr_i, \sum_{i=12}^{2} pr_i \right| \right) \quad (20)$$

bio 17 – Precipitation of the driest quarter given by (Eq. 21):

$$bio17 = min \left(\left| \sum_{i=1}^{3} pr_i, \sum_{i=2}^{4} pr_i, \ldots, \sum_{i=10}^{12} pr_i, \sum_{i=11}^{1} pr_i, \sum_{i=12}^{2} pr_i \right| \right) \quad (21)$$

bio 18 – Precipitation of the warmest quarter. This variable gives the precipitation sum during the warmest three-month period of the year. First the warmest quarter is determined by Eq. 14. Here i^* gives the first month of the warmest quarter identified in Eq. 14. Then the precipitation for the warmest quarter is calculated by (Eq. 22):

$$bio18 = \sum_{i^*}^{i^*+3} pr_{i^*} \quad (22)$$

bio19 – Precipitation of the coldest quarter. This variable gives the precipitation sum during the coldest three-month period of the year. First the coldest quarter is determined by Eq. 15. Here i^* gives the first month of the coldest quarter identified in Eq. 15. Then the precipitation for the coldest quarter is calculated by (Eq. 23):

$$bio19 = \sum_{i^*}^{i^*+3} pr_{i^*} \quad (23)$$

gdd – Growing degree days. The growing degree days give the amount of temperature that accumulates above a given threshold temperature over the course of the year . It has been used to understand e.g. the phenology of plants and animals (Anandhi 2016, Cayton et al. 2015). The function for *gdd* used a B-spline interpolation S(tasi, ti) to approximate daily estimates from monthly tas, with t_i being a vector of mean Julian days of the year marking the middle yearday (Julian day) of the month (t ∈ [349, 15, 45, 74, 105, 135, 166, 196, 227, 258, 288, 319, 349, 15]), tas_i being the mean daily 2 m air temperature for the respective month, and i ∈ [0, 1, …, 13]. The growing degree sum is then given as the integral (Eq. 24):

$$gdd = \int_0^t \left(max(tas_t - tas_b, 0) \right) dt \quad (24)$$

where tas_b is the baseline (default is 5°C) temperature and t represents Julian day.

Example

We provide examples for two common applications of high resolution climate data in macroecology. First, we show how long term normals, i.e. means over periods of e.g. 30 years, can be produced using chelsa-cmip6. Such data are often used in applications such as species distribution modelling for example (Guisan and Zimmermann 2000, Guisan and Thuiller 2005). In addition we show how monthly timeseries can be derived using chelsa-cmip6. Such monthly timeseries are often used as input for dynamic vegetation models such as Fate-HD (Boulangeat et al. 2014), LPJ-Guess (Smith et al. 2014), or MigCLIM (Engler and Guisan 2009).

Step 1: Checking if all needed input data is already available from CMIP6

Before using the chelsa-cmip6 function it is important to check that all input data is actually provided by cmip6, as all climate models from CMIP6 provide all the necessary input data needed for *chelsa_cmip6.py*. *chelsa_cmip6.py* will only work for CMIP6 models that are both available for the historical experiment (following the definition of CMIP as 1850–2015) and the respective activity (activity_id) of interest (e.g. ScenarioMIP that covers the different SSP experiments from 2016 to 2100). This may be verified by using the CMIP6 data search interface (https://esgf-node.llnl.gov/search/cmip6/), with which you can filter the different parameters needed by the chelsa_cmip6.py function (e.g. experiment_id, activity_id, etc.) and inspect if the dataset exists (e.g. by using the parameters given in the example you will see that all necessary input data is available). To check if the historical data exists for the model, just change the activity_id to 'CMIP' and the experiment_id to 'historical'. Make sure the four variables needed do exist for both the historical experiment and the respective experiment from the activity chosen:

1) pr (precipitation rate)
2) tas (mean daily 2 m air temperature)
3) tasmax (maximum daily 2 m air temperature)
4) tasmin (minimum daily 2 m air temperature)

The *chelsa_cmip6* function only runs for CMIP6 models for which all needed variables *tas*, *tasmax*, *tasmin*, *pr*, are available for both the user selected reference and the future period. The standard reference period is 1 January 1981 to 31 December 2010. If another reference period is chosen, the code conducts a delta change for this period as well. Best practice would be to choose the standard reference period. If the combination of model, experiment, member, etc. is not available in the CMIP6 data store, the error 'IndexError: index -1 is out of bounds for axis 0 with size 0' will be displayed.

Example 1: Providing the input parameters resulting in long term climatological normals 2041–2070

Creating long term climatological normals and the related bioclimatic variables that are commonly used in species distribution modeling is controlled via the fefps and fefpe parameters of the chelsa_cmip6 function. You can use function by running the following command in python a python prompt (for an tutorial how to use chelsa-cmip6 in R [www.r-project.org], see Supporting information). Open a python prompt by either typing python in your terminal in Linux, or a command prompt in Windows. First import the chelsa_cmip6 function into your environment using:

```
from chelsa_cmip6.GetClim import chelsa_cmip6
```

Then the chelsa_cmip6 function can be used by providing a set of custom parameters. The example parameters here are given for future climate data from the CMIP6 model MPI-ESM1-2-LR for ssp585 (Wieners et al. 2019), and the future period of 2041–2070, with the reference period 1981–2010 and the region between 5.3°–10.4° longitude and 46.0°–47.5° latitude. The output will be save in your home directory (~/, on a linux system, or : 'C:/Users/your_user_name/' on a Windows machine). The parameters of the chelsa_cmip6 function as follows:

```
chelsa_cmip6(activity_id='ScenarioMIP',
             table_id='Amon',
             experiment_id='ssp585',
             institution_id='MPI-M',
             source_id='MPI-ESM1-2-LR',
             member_id='r1i1p1f1',
             refps='1981-01-15',
             refpe='2010-12-15',
             fefps='2041-01-15',
             fefpe='2070-12-15',
             xmin=5.3,
             xmax=10.4,
             ymin=46.0,
             ymax=47.5,
             output='~/')
```

The output generated by the code consists of netCDF4 files. There will be separate files for each variable and two files each for the reference (refps–refpe) and the future period (fefps–fefpe). Additionally, there will be netCDF4 files for the different bioclimatic variables, one each for the reference (refps–refpe) and the future period (fefps–fefpe).

Example 2: Providing the input parameters resulting in a monthly time series 2016–2100

Creating a monthly timeseries for the same model requires only an adaptation of the fefps, and fefpe parameter of the function. Here we show an example using a simple loop in python. The output will be a netCDF files for each month from 2016, 2100 for tas, tasmax, tasmin, and pr, and an annual timeseries for the bioclimatic variables.

```
for year in range(2016,2101):
    chelsa_cmip6(activity_id='ScenarioMIP',
                 table_id='Amon',
                 experiment_id='ssp585',
                 institution_id='MPI-M',
                 source_id='MPI-ESM1-2-LR',
                 member_id='r1i1p1f1',
                 refps='1981-01-15',
                 refpe='2010-12-15',
                 fefps=year+'-01-15',
                 fefpe=year+'-12-15',
                 xmin=5.3,
                 xmax=10.4,
                 ymin=46.0,
                 ymax=47.5,
                 output='~/')
```

Discussion

The *chelsa-cmip6* package offers convenient access to the entire CMIP6 data archive and downscales it to user defined time periods at user defined geographical extents. Since the first publication of the CHELSA ver. 1.0 data in 2017 (Karger et al. 2017) and the associated downscaled CMIP5 climatologies, there have been repeated requests to make downscaled data for different models, time periods, or RCPs available. While covering the wide range of CMIP5 models and RCPs was still somewhat feasible by only providing data for fixed time periods, this approach became impractical with the wealth of information provided by CMIP6. With the CMIP6 data being available in the cloud, and accessible via lazy loading procedures, the provisioning of a package that utilizes these functions now allows application of the downscaling for a limited geographical extent on end user's computers without the need of cluster computing or high memory servers. The approach applied here does not require downloading the entire input dataset needed for the downscaling and provides an easy solution for researchers to create bioclimatic variables for the time series they require for any given CMIP6 climate model, time period, or experiment. This comes with a large gain in time to make CMIP6 data available at high spatial resolution and help researchers and practitioners in need of high resolution bioclimatic CMIP6 data.

While the package provides easy access to bioclimatic data for any time period requested, it should be kept in mind that the inherent deficits of the delta change method still need to be considered. The delta change method, although robust towards inherent model bias, assumes that the high resolution patterns that are detected in the high resolution reference climatologies remain similar in the future, which is not always realistic in all cases (Christensen et al. 2008). It therefore does not account for changes in e.g. wind patterns to determine orographic precipitation for example (Karger et al. 2021b). For such applications, more sophisticated dynamical downscaling using numerical regional climate models (e.g. Sørland et al. 2021) or terrain-based mechanistic downscaling (e.g. Karger et al. 2022) would be required. The delta change method used in *chelsa-cmip6* therefore offers a compromise between the more exact representation of climatic processes derived from computationally expensive and complex numerical models (Schär et al. 2019), and the ability to downscale large ensembles for applications in biological or ecological studies.

To cite chelsa-cmip6 1.0 or acknowledge its use, cite this Software note as follows, substituting the version of the application that you used for 'version 1.0':

Karger, D. N., Chauvier, Y., Zimmermann, N. E. 2023. chelsa-cmip6 1.0: A python package to create high resolution bioclimatic variables based on CHELSA ver. 2.1 and CMIP6 data. – Ecography 2023: e06535 (ver. 1.0).

Funding – DNK and NEZ acknowledge funding from: the WSL internal grant exCHELSA, the 2019–2020 BiodivERsA joint call for research proposals, under the BiodivClim ERA-Net COFUND program, with the funding organizations Swiss Natl Science Foundation SNF (project: FeedBaCks, 193907), as well as the Swiss Data Science Project: SPEEDMIND. PB, DNK and NEZ the Swiss Data Science Project: COMECO. DNK acknowledges funding to the ERA-Net BiodivERsA – Belmont Forum, with the national funder Swiss Natl Science Foundation (20BD21_184131), part of the 2018 Joint call BiodivERsA-Belmont Forum call (project 'FutureWeb'), as well as the WSL internal grant ClimEx.

Author contributions

Dirk Nikolaus Karger: Conceptualization (lead), Code curation (equal), Formal Analysis (equal), Methodology (lead), Writing-original draft (lead), Writing-review and editing (equal), Funding acquisition (equal). **Yohann Chauvier**: Code curation (equal), Formal Analysis (equal), Writing-review and editing (equal). **Niklaus E. Zimmermann**: Conceptualization (support), Methodology (equal), Writing-review and editing (equal), Funding acquisition (equal).

Transparent peer review

The peer review history for this article is available at https://publons.com/publon/10.1111/ecog.06535.

Data availability statement

This article contains no additional data.

Supporting information

The Supporting information associated with this article is available with the online version.

References

Arnell, N., Liu, C., Conpagnucci, R., Hanaki, K., Howe, C., Mailu, G., Shicklomanov, I. and Stakhiv, E. 2001. Hydrology and water resources. – In: Mccarthy, J. J., Canziani, O. F., Leary, N. A., Dokken, D. J., White, K. S. (eds), IPCC climate change 2001: impacts, adaptation & vulnerability, the third assessment report of working group II of the intergovernmental panel on climate change (IPCC). Cambridge Univ. Press.

Christensen, J. H., Boberg, F., Christensen, O. B. and Lucas-Picher, P. 2008. On the need for bias correction of regional climate change projections of temperature and precipitation. – Geophys. Res. Lett. 35: L20709.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J. and Taylor, K. E. 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. – Geosci. Model Dev. 9: 1937–1958.

Frieler, K. et al. 2017. Assessing the impacts of 1.5°C global warming – Simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b). – Geosci. Model Dev. 10: 4321–4345.

Gidden, M. J. et al. 2019. Global emissions pathways under different socioeconomic scenarios for use in CMIP6: a dataset of harmonized emissions trajectories through the end of the century. – Geosci. Model Dev. 12: 1443–1475.

Hall, C. A. and Meyer, W. W. 1976. Optimal error bounds for cubic spline interpolation. – J. Approx. Theory 16: 105–122.

Harris, C. R. et al. 2020. Array programming with NumPy. – Nature 585: 357–362.

Hay, L. E., Wilby, R. and Leavesley, G. H. 2000. A comparison of delta change and downscaled gcm scenarios for three mountainous basins in the United States. – J. Am. Water Resour. Assoc. 36: 387–397.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. and Jarvis, A. 2005. Very high resolution interpolated climate surfaces for global land areas. – Int. J. Climatol. 25: 1965–1978.

Hoyer, S. et al. 2022. xarray. – https://doi.org/10.5281/zenodo.6628513.

Hurlbert, A. H. and Jetz, W. 2007. Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. – Proc. Natl. Acad. Sci. USA 104: 13384–13389.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Peter Linder, H. and Kessler, M. 2017. Climatologies at high resolution for the earth's land surface areas. – Sci. Data 4: 170122.

Karger, D. N. et al. 2021a. Climatologies at high resolution for the earth's land surface areas. – https://www.envidat.ch/dataset/chelsa-climatologies.

Karger, D. N., Wilson, A. M., Mahony, C., Zimmermann, N. E. and Jetz, W. 2021b. Global daily 1 km land surface precipitation based on cloud cover-informed downscaling. – Sci. Data 8: 307.

Karger, D. N., Lange, S., Hari, C., Reyer, C. P. O. and Zimmermann, N. E. 2022. CHELSA-W5E5 v1.0: W5E5 v1.0 downscaled with CHELSA v2.0. – ISIMIP Repository, https://doi.org/10.48364/ISIMIP.836809.3.

Kurtzer, G. M., Sochat, V. and Bauer, M. W. 2017. Singularity: scientific containers for mobility of compute. – PLoS One 12: e0177459.

Lafon, T., Dadson, S., Buys, G. and Prudhomme, C. 2013. Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods. – Int. J. Climatol. 33: 1367–1381.

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M. and Thiele-Eich, I. 2010. Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. – Rev. Geophys. 48: RG3003.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K. and Sanderson, B. M. 2016. The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. – Geosci. Model Dev. 9: 3461–3482.

Schär, C. Fuhrer, O., Arteaga, A., Ban, N., Charpilloz, C., Di Girolamo, S., Hentgen, L., Hoefler, T., Lapillonne, X., Leutwyler, D., Osterried, K., Panosetti, D., Rüdisühli, S., Schlemmer, L., Schulthess, T., Sprenger, M., Ubbiali, S. and Wernli, H. 2019. Kilometer-scale climate models: prospects and challenges. – Bull. Am. Meteorol. Soc. 101: E567–E587.

Sørland, S. L. et al. 2021. COSMO-CLM regional climate simulations in the Coordinated Regional Climate Downscaling Experiment (CORDEX) framework: a review. – Geosci. Model Dev. 14: 5125–5154.

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J. and Rose, S. K. 2011. The representative concentration pathways: an overview. – Clim. Change 109: 5.

Wieners, K.-H. et al. 2019. MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 ScenarioMIP ssp585. – Earth System Grid Federation.

Xu, T. and Hutchinson, M. F. 2013. New developments and applications in the ANUCLIM spatial climatic and bioclimatic modelling package. – Environ. Model. Softw. 40: 267–279.