



High-resolution downscaling of CMIP6 Earth system and global climate models using deep learning for Iberia

Pedro M. M. Soares, Frederico Johannsen, Daniela C. A. Lima, Gil Lemos, Virgílio A. Bento, and Angelina Bushenkova

Instituto Dom Luiz, IDL, Faculty of Sciences, University of Lisbon, 1749-016 Lisbon, Portugal

Correspondence: Frederico Johannsen (jfjohannsen@ciencias.ulisboa.pt)

Received: 27 June 2023 – Discussion started: 28 June 2023

Revised: 24 October 2023 – Accepted: 19 November 2023 – Published: 12 January 2024

Abstract. Deep learning (DL) methods have recently garnered attention from the climate change community for being an innovative approach to downscaling climate variables from Earth system and global climate models (ESGCMs) with horizontal resolutions still too coarse to represent regional- to local-scale phenomena. In the context of the Coupled Model Intercomparison Project phase 6 (CMIP6), ESGCM simulations were conducted for the Sixth Assessment Report (AR6) of the Intergovernmental Panel on Climate Change (IPCC) at resolutions ranging from 0.70 to 3.75°. Here, four convolutional neural network (CNN) architectures were evaluated for their ability to downscale, to a resolution of 0.1°, seven CMIP6 ESGCMs over the Iberian Peninsula – a known climate change hotspot, due to its increased vulnerability to projected future warming and drying conditions. The study is divided into three stages: (1) evaluating the performance of the four CNN architectures in predicting mean, minimum, and maximum temperatures, as well as daily precipitation, trained using ERA5 data and compared with the Iberia01 observational dataset; (2) downscaling the CMIP6 ESGCMs using the trained CNN architectures and further evaluating the ensemble against Iberia01; and (3) constructing a multi-model ensemble of CNN-based downscaled projections for temperature and precipitation over the Iberian Peninsula at 0.1° resolution throughout the 21st century under four Shared Socioeconomic Pathway (SSP) scenarios. Upon validation and satisfactory performance evaluation, the DL downscaled projections demonstrate overall agreement with the CMIP6 ESGCM ensemble in magnitude for temperature projections and sign for the projected temperature and precipitation changes. Moreover, the advantages of using a high-resolution DL downscaled ensemble of ESGCM cli-

mate projections are evident, offering substantial added value in representing regional climate change over Iberia. Notably, a clear warming trend is observed in Iberia, consistent with previous studies in this area, with projected temperature increases ranging from 2 to 6 °C, depending on the climate scenario. Regarding precipitation, robust projected decreases are observed in western and southwestern Iberia, particularly after 2040. These results may offer a new tool for providing regional climate change information for adaptation strategies based on CMIP6 ESGCMs prior to the next phase of the European branch of the Coordinated Regional Climate Downscaling Experiment (EURO-CORDEX) experiments.

1 Introduction

The Sixth Assessment Report (AR6) of the Intergovernmental Panel on Climate Change (IPCC) was released in August 2021, dramatically calling for urgent action to reduce global greenhouse gas emissions (GGEs) due to the scale of the projected changes for the climate system from the mean state to extremes (IPCC, 2021). The results in the report are based on the Coupled Model Intercomparison Project phase 6 (CMIP6) simulations, which were performed using Earth system and global climate models (ESGCMs), and included runs with a spatial resolution in the range of 0.70 to 3.75°. The IPCC report projects worrying changes in global-scale extreme events, such as significant increases in the frequency and intensity of heat waves, droughts, and extreme precipitation. Although based on global simulations, the AR6 showed particularly pronounced changes on a regional level in some climate change hotspots, like the Mediterranean

region (Turco et al., 2015; Cos et al., 2022; Lionello and Scarascia, 2018).

It is widely accepted that most resolutions used by ESGCMs are still too coarse to represent many regional- to local-scale processes that define the local climate (Randall et al., 2007; Soares et al., 2012; Rummukainen, 2016). This disadvantage highlights the necessity for downscaling methods, at a higher resolution, which often provide regional to local fine-scale information, which is crucial for impact and adaptation studies. There is a plethora of downscaling methods, including dynamical ones using regional climate models (RCMs), statistical ones using statistical downscaling methods (SDMs), and, recently, an umbrella group of the latter, designated as artificial intelligence (AI) approaches, which include machine-learning (ML) and deep-learning (DL) methods.

RCMs are forced at the boundaries by ESGCMs (Dickinson et al., 1989; Giorgi and Bates, 1989; Giorgi and Mearns, 1991; McGregor, 1997; Christensen et al., 2007), using higher resolutions (~ 10 km) in limited area domains, which significantly improve the description of regional to local climates (Giorgi and Mearns, 1999; Laprise, 2008; Rummukainen, 2010, 2016; Feser et al., 2011; Soares et al., 2012, 2017a, b; Rios-Entenza et al., 2014; Giorgi et al., 2016; Lucas-Picher et al., 2017; Cardoso et al., 2019). Nevertheless, considering local and, especially, sub-daily climate features, RCMs still present limitations in capturing sub-grid processes such as convection (Prein et al., 2013). In order to bridge this gap, RCMs are running at very high resolutions, usually described as convective-permitting resolutions (approximately 1 km), where deep convection is explicitly resolved by the grid mesh at grid spacing below 3 km (Prein et al., 2015; Coppola et al., 2020; Pichelli et al., 2021; Soares et al., 2022).

SDMs are based on the establishment of empirical relationships between large-scale atmospheric predictors and local observed predictands describing local climate (Wilby and Wigley, 1997; Fowler et al., 2007; Nikulin et al., 2018; Hertig et al., 2019; Maraun et al., 2019; Gutiérrez et al., 2019a; Rössler et al., 2019; Soares et al., 2019; Widmann et al., 2019). Subsequently, projections of future regional to local climate variables are determined from future large-scale atmospheric conditions. SDMs include model output statistics and perfect-prognosis approaches (Maraun et al., 2010, 2017). However, when compared to dynamical downscaling, the model formulation of SDMs lacks physical constraints and, in general, does not ensure a full multivariate consistency (Le Roux et al., 2018). Since SDMs use observations for training, they are able to overcome the systematic biases often displayed by RCMs. Additionally, since SDMs are not computationally demanding, the need for large computational infrastructures is avoided.

There is a continuous improvement in SDMs, and new AI approaches are being proposed for climate applications, with deep learning (DL) being one of the most promising. DL is

a subdomain of machine learning (ML). In ML, the models learn the optimal value of their parameters automatically. Since parameter tuning is based on the input data fed to the model, the model is able to make predictions when forced by new data (see Alzubi et al., 2018, for an overview of ML). Unlike “shallow” learning models (e.g., random forests and support vector machines), DL models learn non-linear relationships between data due to their “deep” layered structure. DL has become a common approach in research over the past decade (Schmidhuber, 2015), including in Earth sciences in the last few years (Reichstein et al., 2019), thanks to advances in computational power and data availability. For example, the European Centre for Medium-Range Forecasts (ECMWF) features DL as the main showcase in its Destination Earth project (Bauer et al., 2021), which will attempt to create digital twins of the Earth system in the next decade.

The most common DL model type is the artificial neural network (ANN), an attempt to design an artificial analog to the biological neural networks that exist in the human brain. One of the most used types of ANNs is convolutional neural networks (CNNs). These models are widely used in the field of computer vision, as they extract information and identify objects in images (LeCun and Bengio, 1995). However, the value of CNNs is not restricted to computer vision, as CNNs have been used in other research areas, including in Earth sciences, for example, in model parameterization (Chantry et al., 2021a) and ensemble postprocessing (Rasp and Lerch, 2018), showing promising results. Climate downscaling is another promising area benefitting from the implementation of CNNs. Early attempts at downscaling using simple ANN structures were not compelling due to limited input data, computational resources, and scarcer observations (e.g., Wilby et al., 1998; Trigo and Palutikof, 1999). Recent studies have shown more favorable results, equalling and even surpassing classic SDMs (e.g., Baño-Medina et al., 2020; Hernanz et al., 2022; Baño-Medina et al., 2022). Recently, and for the first time, Baño-Medina et al. (2022) were able to downscale climate projections with the aid of DL for precipitation and temperature, based on a set of global climate models (GCMs) from the Coupled Model Intercomparison Project phase 5 (CMIP5). These authors showed that DL reduced the biases in the historical period when compared to an ensemble of RCMs with 0.44° resolution from EURO-CORDEX (European branch of the Coordinated Regional Climate Downscaling Experiment). In addition, the resulting climate change signals have similar spatial patterns to those obtained from the RCMs, and when looking at the uncertainty, the DL preserves the uncertainty of the climate change signal for temperature and reduces it for precipitation.

Despite their promising results, DL methods are viewed with caution in the scientific community due to their black-box nature. DL models usually have thousands (if not millions) of trainable parameters that hinder a physically based explanation for the quality of their results. There have been

attempts to improve the understanding of models' reasoning (e.g., Carter et al., 2018), building an overall framework for DL studies in Earth sciences, including weather/climate modeling and postprocessing, and generating consistent intercomparable studies (Reichstein et al., 2019; Chantry et al., 2021b; Haupt et al., 2021). As a result, the first benchmark dataset for data-driven weather forecasting has been created (Rasp et al., 2020). DL presents other general limitations, including the need for hardware (graphics processing units or GPUs accelerate the model training, while the more common central processing units or CPUs can be computationally costly; Chantry et al., 2021a). Other DL limitations concern the climate research field. Lack of explicit physics in the DL models and the need to split the data in a way that includes long-term patterns and trends (e.g., El Niño–Southern Oscillation or ENSO and global warming) in both training and test phases for long-term datasets (Schultz et al., 2021) are examples of said limitations.

The Iberian Peninsula, within the Mediterranean basin, is a known climate change “hotspot” (Planton et al., 2012; Difenbaugh and Giorgi, 2012; Turco et al., 2015; Russo et al., 2019; Cos et al., 2022) due to its high vulnerability to warming and drying conditions (Argüeso et al., 2012; Cardoso et al., 2019; Soares et al., 2017a, b; Lima et al., 2023a, b; Soares and Lima, 2022), leading to high occurrence of extreme events, such as droughts, heat waves, and wildfires (Hoerling et al., 2012; Bento et al., 2022, 2023; Soares et al., 2023a). Future projections point to a warming trend that is stronger for daytime values during summer and autumn than in other seasons, resulting in an amplification of the daily and annual temperatures (Cardoso et al., 2019; Lima et al., 2023a). Also, a significant reduction in the mean precipitation is projected throughout the entire year (Argüeso et al., 2012; Lima et al., 2023a; Soares et al., 2017a). Concomitant with the projected warming and drying trends, the occurrence of hot and dry extreme events is expected to become more frequent, intense, and longer (Hoerling et al., 2012; Lima et al., 2023b), which may have significant impacts on human and natural sectors, such as agriculture (Bento et al., 2021), forests (Palma et al., 2015, 2018), coastal areas (Pereira and Coelho, 2013), and water resources (Soares and Lima, 2022). The Iberia01 regular gridded product (hereafter Iberia01) is the highest-resolution observational daily dataset including mean, maximum, and minimum temperatures and precipitation, covering the full domain of continental Iberia (Herrera et al., 2019). Iberia01 is commonly used for assessing the performance of ESGCMs (Soares et al., 2022), RCM results (Herrera et al., 2020; Careto et al., 2022a, b), building of multi-model ensembles for climate change assessments (Soares et al., 2023a; Lima et al., 2023a, b), and other studies, such as those related to water availability (Soares and Lima, 2022) and droughts (Páscoa et al., 2021; Soares et al., 2023a).

The most consistent and widely used high-resolution climate change dataset for Iberia remains the EURO-CORDEX

and CORDEX-Core runs (Jacob et al., 2014, 2020). These regional climate simulations were forced by the previous CMIP5 global climate simulations and are becoming less useful after the recent release of the CMIP6 results forced by the Shared Socioeconomic Pathway (SSP) and Representative Concentration Pathway (RCP) greenhouse gas emissions scenarios. At present, the new EURO-CORDEX simulations protocol, forced by CMIP6 runs, is being finished, and widespread availability of new simulations and results for the scientific community and society is not expected before 1–2 years' time. Additionally, the building of new multi-model and multi-approach ensembles is highly beneficial to assess the robustness and uncertainty of future climate projections (Lima et al., 2023a). The increasing need for exploring and updating regional climate information for Iberia requires and benefits from the use of other approaches to downscale the current CMIP6 runs.

In the present study, a DL methodology based on the work of Baño-Medina et al. (2022) is used to downscale, in a consistent manner, the CMIP6 runs at high resolution for Iberia. A matrix of plausible futures is used to select the CMIP6 models considered to be in agreement with the EURO-CORDEX evaluation study (Sobolowski et al., 2023). The DL algorithm is trained using ERA5 and compared to the high-resolution regular gridded dataset Iberia01 (Herrera et al., 2019) for the current climate, covering the period 1979–2014, and is then used to downscale future projections in agreement with four SSP–RCP scenarios, namely SSP1–RCP2.6, SSP2–RCP4.5, SSP3–RCP7.0, and SSP5–RCP8.5 (O'Neill et al., 2016), for three future periods throughout the 21st century, namely the beginning of the century (2015–2040), middle of the century (2041–2070), and end of the century (2071–2100). First, different architectures of DL are trained and evaluated for the present climate and then multi-model projections are performed, based on a simple-averaged multi-model ensemble approach. This study is focused on four of the main climate variables and their extremes: minimum, mean, and maximum temperatures and precipitation. The main goals of this study are to understand the accuracy of downscaling CMIP6 GCMs to a much finer spatial resolution using DL and to take advantage of the advancing AI methods to compile information that may be crucial to timely assist mitigation and adaptation plans being developed at the national, regional, and local levels within Iberia.

2 Data and methods

2.1 Study area

The Iberian Peninsula (IP) is located in the southwestern tip of Europe (Fig. 1), bordered by the Atlantic Ocean and the Mediterranean Sea. The IP sits in a climate transition zone between the arid and semiarid climates of subtropical re-

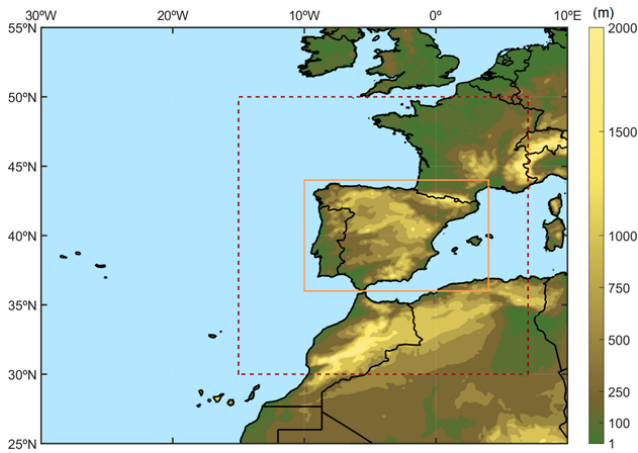


Figure 1. Consideration of western Europe and northwestern Africa topography (m) and of the domains of the Earth system and global climate model predictors (dashed red line) and predictands (full orange line).

regions and the humid temperate climates of northern Europe. Despite having a surface area of smaller than 600 000 km², it shows a diverse climate, with significant regional variations. In fact, while the northern and northwestern regions are marked by long rainy seasons and temperate summers, the south and southeast are characterized by long and hot summers and a clear dry season. The interior regions are defined by a continental climate, with hot summers and cold winters. Additionally, local and regional topographic features play a significant role in modulating climate features throughout the IP. Here, the IP domain is considered to be the land area between 36 and 44° N and ° W and 4° E (Fig. 1; inside the orange line). The predictor domain (Fig. 1; dashed red line) is a larger region than the IP domain to ensure that large-scale phenomena are included in the information provided by the predictors to train the DL models.

2.2 ERA5 reanalysis

ERA5 is the latest European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis (Hersbach et al., 2020) produced within the Copernicus Climate Change Service (C3S). ERA5 provides a comprehensive, high-resolution record of the global atmosphere, land surface, and ocean from 1950 onwards. Benefitting from advanced research and model physics development, outputs are archived at 0.25° × 0.25° horizontal and 1 h time resolutions, considering 137 atmospheric levels up to 0.01 hPa. The ECMWF Integrated Forecasting System (IFS) Cy41r2, used operationally for forecasting from March to November 2016, is used to produce ERA5. Additional details are available in Hersbach et al. (2020). Here, the period from 1 January 1979 to 31 December 2014 is considered. The original ERA5 reanalysis data were interpolated to a 1° × 1° horizontal reso-

lution, using a bilinear interpolation method, to build a common grid to the CMIP6 ESGCMs (Sect. 2.3).

2.3 CMIP6 Earth system global climate models

The ESGCMs selected for the current study closely follow the model array built in Sobolowski et al. (2023) for the ongoing CMIP6 dynamical downscaling that is being performed (i.e., the regional climate model simulations of EURO-CORDEX phase II). The authors thoroughly analyzed the ability of the CMIP6 ESGCMs to describe the most important large-scale features that define the European climate, such as the storm track position, and that span the AR6 IPCC climate sensitivity range. The ESGCMs considered are listed in Table 1; understandably, the list is additionally constrained by the availability of the predictors data. The predictor data were extracted for the domain in Fig. 1 (inside the dashed red line), limited by 30° N–50° N, 15° W–7° E, and then interpolated to a common grid at a 1° × 1° resolution, using the bilinear interpolation method.

2.4 Iberia01 observational regular gridded dataset

The Iberia01 regular gridded product is the highest-resolution observational daily dataset including mean, maximum, and minimum temperatures and precipitation that covers the full domain of continental Iberia (Herrera et al., 2019). This observational dataset was built using an unprecedented number of ground station observations, 275 for temperatures and 3486 for daily accumulated precipitation, resulting in a high-quality regular gridded dataset at 0.1° × 0.1° horizontal resolution. Here, the Iberia01 product is used both to calibrate and evaluate the deep learning approach for the period 1979–2014 (same as ERA5).

2.5 Deep learning methodology

Convolutional neural networks (CNNs) are deep learning (DL) model structures specializing in extracting features automatically from geospatial data. The architecture of a CNN model includes convolutional layers that perform feature identification and extraction using filters that apply the mathematical operation of cross-correlation to the data (LeCun and Bengio, 1995; see Fig. 3 of Baño-Medina et al., 2020). The general outline of one epoch, i.e., a full cycle of the training phase, is as follows:

- The 2D filters in a convolutional layer “scan” the set of predictor variables, computing a set of filter maps based on each filter, highlighting different features/patterns of the original data. These filter maps are then used as input for the following convolution layer.
- The output of the final convolutional layer is flattened (reshaped to 1D) before being fed to the fully connected (dense) layer that follows.

Table 1. CMIP6 Earth system global climate models. Note that “Hor. and vert. res.” stands for horizontal and vertical resolution.

ESGCM (CMIP6)	Institute	Reference	Hor. and vert. res. atmosphere	Hor. and vert. res. ocean
ACCESS-CM2	CSIRO/BOM	Bi et al. (2013)	$1.25^\circ \times 1.875^\circ$, L85	$1.00^\circ \times 1.00^\circ$, L50
MPI-ESM1-2-HR	MPI	Müller et al. (2018); Gutjahr et al. (2019)	$0.90^\circ \times 0.90^\circ$, L95	$0.40^\circ \times 0.40^\circ$, L40
IPSL-CM6A-LR	IPSL	Boucher et al. (2020)	$1.25^\circ \times 2.50^\circ$, L79	$0.5\text{--}1.00^\circ \times 1.00^\circ$, L75
MIROC6	AORI/NIES/JAMSTEC	Tatebe et al. (2019)	$1.40^\circ \times 1.40^\circ$, L81	$1.00^\circ \times 1.00^\circ$, L62
NorESM2-MM	NCC	Seland et al. (2020)	$0.90^\circ \times 1.25^\circ$, L32	$1.00^\circ \times 1.00^\circ$, L53
UKESM1-0-LL	UKMO	Sellar et al. (2019)	$1.25^\circ \times 1.875^\circ$, L85	$1.00^\circ \times 1.00^\circ$, L75
CNRM-ESM2-1	CNRM	S��ferian et al. (2019)	$1.40^\circ \times 1.40^\circ$, L91	$1.00^\circ \times 1.00^\circ$, L75

- The units in a dense layer are connected to every unit in the previous and following layers, allowing the network to learn potential relationships between all units in successive layers. The final dense layer must have the size of the target data in order to generate the predictions.
- The predictions are compared with the observations by calculating the loss, according to the loss function defined by the user.
- Finally, the model attempts to lower the loss by the use of the stochastic gradient descent optimization algorithm, tuning the parameters of each model layer according to the direction of the gradient that minimizes the loss the fastest. This process begins in the output layer, computing the gradients on that layer, and backtracks all the way to the first convolutional layer in what is known as the back-propagation algorithm.

The model then repeats the training until it reaches a convergence mark defined by the user (usually a set number of epochs after the loss stops decreasing). While the learnable parameters are optimized automatically by the model, there is a set of hyperparameters that is defined by the user, including the following:

- the maximum number of epochs that the model can run;
- the batch size of observations used to tune the model in each training cycle; and
- the learning rate at which the model incorporates new information after each epoch.

The main goal of DL is to achieve generalization; i.e., the ability to make quality predictions when given new, never-before-seen data (extrapolation). Such a feature is particularly important when training DL models for climate studies, due to global warming and other long-term trends. The model structures considered in this study were retrieved from the Ba  o-Medina et al. (2020) and are described in Table 2.

Although all models have similar structures, differing only in the small details, they are designed in such a way that every model is slightly more complex than the previous one. All models comprise the following:

- three convolutional layers (the first two layers have 50 and 25 filters each);
- a final dense layer that outputs the predictions; and
- the same hyperparameters (batch size = 100 and learning rate = 0.0001).

The differences among the models are as follows:

- The third convolutional layer has 1 filter in BMlinear and BM1 architectures and 10 filters in BM10 and BMdense architectures.
- BMdense presents two additional dense layers, both with 50 units, prior to the output layer.
- The activation function in every layer of every model is the rectified linear unit (ReLU), a non-linear function, except in BMlinear, in which the function is linear.

The loss function used for the temperature predictions is based on the mean squared error. For precipitation, however, the DL models feature a multi-output structure (see Fig. 3 in Ba  o-Medina et al., 2020). Instead of predicting precipitation directly, the model attempts to obtain three parameters: the shape (alpha) and scale (beta) of the gamma distribution and the probability of precipitation (p). This is achieved by applying a custom loss function that computes the negative log likelihood of the Bernoulli–gamma distribution (Canon, 2008), following the methodology presented in Ba  o-Medina et al. (2020). The precipitation value is obtained by multiplying the alpha and beta parameters.

2.6 Selection of predictors, training, and evaluating

The predictors selected follow the Ba  o-Medina et al. (2022) study and are included in Table 3. The data were pre-

Table 2. CNN architectures used in this study (adapted from Baño-Medina et al., 2022). The architecture is divided into one input and one output layer and several hidden layers in between. Numbers represent the units in each hidden layer, with convolutional layers in bold; otherwise, dense layers are given. The input format is lat × long × 15 (five predictors times three pressure levels), and the output is a 6523 × 1 vector (the number of 0.1° land grid points over Iberia).

Model	Architecture	Activation function	Rationale
BMlinear	Input – 50 – 25 – 1 – output	Linear	Using convolutions to perform the downscaling
BM1	Input – 50 – 25 – 1 – output	ReLU	Add non-linearity to the model structure
BM10	Input – 50 – 25 – 10 – output	ReLU	Increase the number of filters in the last convolution layer
BMdense	Input – 50 – 25 – 10 – 50 – 50 – output	ReLU	Deepen the model structure

processed before being used to train and evaluate the DL models. The ERA5 variables, used as predictors, were standardized to facilitate the training of the DL models. Grid points with missing data in the CMIP6 ESGCMs were filled with an average of the surrounding grid points. If the surrounding grid points had missing data as well, then a domain average was applied. Afterwards, the dataset was standardized (with the same parameters used for ERA5). The ESGCMs were bias-corrected in relation to ERA5 through a simple mean-variance scaling method. The climate change trend was removed in the future scenarios before the bias correction and reintroduced afterwards (Vrac and Ayar, 2017).

Two stages were pursued with the aim of training and evaluating the four architectures (Fig. 2). The first stage was to train them using ERA5 predictors (Table 3), considering the 1979–2004 period, validating their performance between 2005 and 2009, and finally testing the architectures for the period 2010–2014. This process was performed to obtain each of the four predictands, namely daily mean temperature (T), daily minimum temperature (T_{\min}), daily maximum temperature (T_{\max}), and daily accumulated precipitation (Pr) (Table 3). The results of the DL downscaled predictands from ERA5 were then compared with the Iberia01 reference data. In this case, since the DL used ERA5 reanalysis predictors, the evaluation was performed with daily synchronized climate data. This evaluation, conducted between 2010 and 2014, was based on error metrics such as the bias, the root mean squared error (RMSE), the standard deviation ratio (SDR), the Perkins' skill score (PSS), and the relative operating characteristic skill score (ROCSS).

The mean bias, used for temperature and precipitation is defined as

$$\text{Bias} = \frac{1}{N} \sum_{k=1}^N (m_k - o_k), \quad (1)$$

where o_k and m_k are, respectively, the observed and modeled time series, and N is the total number of grid points.

The root mean squared error (RMSE), used for temperature and precipitation, is defined as

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{k=1}^N (m_k - o_k)^2}. \quad (2)$$

The standard deviation ratio, used only for temperature, is expressed as

$$\sigma_n = \frac{\sigma_m}{\sigma_o} = \frac{\sqrt{\frac{1}{N} \sum_{k=1}^N (m_k - \underline{m})^2}}{\sqrt{\frac{1}{N} \sum_{k=1}^N (o_k - \underline{o})^2}}, \quad (3)$$

where σ_o and σ_m are standard deviations of the observed and modeled time series, respectively, while \underline{o} and \underline{m} represent the respective mean values.

The Perkins' skill score (PSS; Perkins et al., 2007) quantifies the model's ability to reproduce the observed probability distribution functions (PDFs) as follows:

$$S = 100 \times \sum_{i=1}^B \min[E_{m,i}, E_{o,i}] \quad , \quad (4)$$

where E_m and E_o are, respectively, the modeled and observed empirical PDFs, and $\min[E_{m,i}, E_{o,i}]$ is the minimum between the two values. B is the total number of bins used to compute the PDF.

Finally, the relative operating characteristic skill score (ROCSS) is given by

$$\text{ROCSS} = 2 \times \text{Area under the ROC curve} - 1. \quad (5)$$

For the extreme values, the 2nd and 98th percentiles of T , the 10th (90th) percentile of T_{\min} (T_{\max}), and the 98th percentile of Pr were computed and compared with those from Iberia01 (bias).

The second stage consisted in training the DL architectures with ERA5; this time, the complete 1979–2014 period was used. These architectures were then used to downscale the individual CMIP6 ESGCMs for the same period for each of the four predictands. The resulting DL downscaled ESGCMs, at 0.1° horizontal resolution (ESGCM-DL), are non-synchronized with the Iberia01, and consequently, only a statistical comparison was performed. Therefore, Julian years with 365 multi-year daily means were computed for each ESGCM-DL and for the Iberia01, and a performance evaluation based on the same error metrics as in the first stage was conducted. Finally, a simple average ensemble was built for each architecture, containing seven ESGCM-DL models, and compared to the 1° ESGCM ensemble, the 1° ERA5 reanalysis, and the interpolated 0.1° ERA5 reanalysis.

Table 3. ERA5 and CMIP6 predictors and predictands and the respective atmospheric levels.

Levels		850–700 hPa–500 hPa			
Predictors (daily)	Temperature (ta)	Humidity (hus)	Geopotential height (zg)	Zonal wind speed (ua)	Meridional wind speed (va)
Levels		Surface			
Predictands (daily)	Mean temperature (tas)	Minimum temperature (tasmin)	Maximum temperature (tasmax)	Precipitation (pr)	

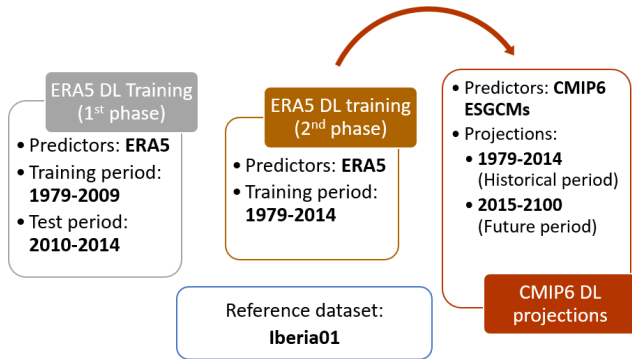


Figure 2. Summary of the two phases of the methodology (detailed in Sect. 2.6) describing the predictors and training and projections periods considered in each phase.

2.7 Future climate projections

The present climate historical period considered here corresponds to 1981–2010. The future climate projections are focused on three periods: 2015–2040 (beginning of the 21st century), 2041–2070 (middle of the 21st century), and 2071–2100 (end of the 21st century), encompassing four CMIP6 SSPs (Rozenberg et al., 2014; O’Neill et al., 2016), namely SSP1–2.6, SSP2–4.5, SSP3–7.0, and SSP5–8.5. These scenarios range from a strong mitigation level, resulting in low greenhouse gas emissions (GGEs), with CO₂ emissions cut to net zero around 2075 (SSP1–2.6); an intermediate trajectory of future GGEs, with CO₂ emissions maintaining current levels until 2050 and then reducing, but not achieving, net zero by 2100 (SSP2–4.5); and, finally, two scenarios with increasing GGEs with SSP3–7.0 and SSP5–8.5, where the former considers that CO₂ emissions double by 2100, and the latter considers a 3-fold increase by 2075. In this study, results of future climate projections correspond to anomalies (differences) between the future and the historical climatological values, as given by a simple averaged multi-model ensemble consisting of all ESGCM-DL outputs. It should be noted that the projected temperature increase depends on the chosen historical period. Downscaling using the four DL algorithms is performed for each ESGCM considered in this study (Table 1) for the disclosed future periods. As reference, the change signal linked to all ESGCMs is also computed.

The future ESGCM-DL-projected climate of Iberia is analyzed in terms of mean climate and extreme values. Anomaly maps for the annual projected changes for Iberia are presented for all variables, where the differences between the 1° ESGCMs and 0.1° ESGCM-DL projections are highlighted. Box plots summarizing the projected changes (median, interquartile range, and variability) are also presented for the four predictands.

3 Results

3.1 Evaluation of DL forced by ERA5

The four DL architectures are trained and validated with ERA5 for the 1979–2004 and 2005–2009 periods, respectively, and finally tested during 2010–2014 against Iberia01 considering minimum, mean, and maximum temperatures and precipitation. The performance evaluation metrics are shown in Fig. 3 (*T*), Fig. 4 (*T*_{min}), Fig. 5 (*T*_{max}), and Fig. 6 (Pr). The comparison between the ERA5 (interpolated to 0.1° horizontal resolution; iERA5 from here on) and Iberia01 is also shown as a reference for all fields (dark gray box plot).

Considering *T* (Fig. 3), the following three main outcomes emerge: (1) all of the DL approaches display rather small errors and even slight improvements in comparison with iERA5, such as concerning RMSE and the PSS; (2) the DL architectures present less variability in accuracy metrics (bias) than iERA5, but in some cases, the error distribution of the latter is more closely centered around zero than for the DL outcomes; and (3) the four architectures present small biases for extreme values. When considering the total bias, the four architectures show somewhat interchangeable results, with median values slightly below zero for BMlinear, virtually zero for BM1, and slightly over zero for BM10 and BMdense. The small warm bias found for the 2nd percentile of *T* is observed both in the DL outcomes and in the iERA5. However, the cold bias found in the 98th percentile of *T* is only found in the DL outcomes, with values close to zero for the iERA5. Regarding the RMSEs, DL results show lower variability ranges than iERA5 and overall lower median values with increasing DL complexity. The interquartile range for the RMSEs of the DL results encompasses values from 1.25 to 1.5 °C. In terms of SDR, in relation to Iberia01,

the iERA5 shows the median value closest to 1; however, it also shows the largest interquartile distance and largest variability range (from ~ 0.90 to ~ 1.08 when compared with ~ 0.93 to ~ 1.04 for the DL outcomes). Finally, regarding the PSS, the four architectures show more similarity between distributions of T with Iberia01 than the iERA5. A distinction between the DL outcomes for this error metric is rather unnoticeable.

Regarding T_{\min} (Fig. 4), the overall results show the following: (1) the DL architectures present better results in comparison to Iberia01 than the simple interpolation of ERA5, showing lower RMSEs, biases closer to zero, and larger PSS values; and (2) the four architectures show, to some extent, similar results between them. Following a more detailed analysis, the median biases presented by the four architectures are all near-zero, while the iERA5 shows a median bias larger than 1°C . Furthermore, when comparing the bias of the architectures to the interpolated ERA5, a lower interquartile range (circa 1°C) is observable in the first range compared to the latter ($\sim 2^\circ\text{C}$). Additionally, a narrower extreme bias variability range (about 2.5°C versus about 4°C , respectively) is seen. Results for extreme low temperatures (bias p10) are in line with the total bias; nevertheless, these results show a slight tendency to lower median biases as the complexity of the architecture increases. This is also noticeable in the precision metric, with reduced RMSEs for increasing DL architecture complexity. However, here, BM10 and BMdense show very similar results. All architectures present median RMSEs below 2°C , which is the third quartile of the three more complex ones below this value as well. The maximum RMSE does not surpass 3°C . On the other hand, the iERA5 shows a median RMSE slightly above 2°C , the third quartile close to 3°C , and a maximum value above 4.5°C . Similar to the T results in Fig. 3, the median SDR is closer to 1 for the iERA5; nevertheless, the DL architectures show greater variability ranges for T_{\min} in comparison to T . Among the architectures, BMdense is the one with a standard deviation ratio median closer to 1. Finally, considering the PSS values, it is once again noticeable that the distributions of the down-scaled ERA5 using DL and Iberia01 tend to match better with the increase in complexity of the architectures.

When assessing T_{\max} (Fig. 5), the following three main results may be highlighted: (1) all the error metrics are improved by the DL methods when compared with iERA5; (2) the DL architectures show much less variability in the biases and RMSEs in comparison to iERA5 (having Iberia01 as reference); and (3) the four architectures show, to some extent, similar results between them. In terms of bias, and considering the four architectures, neither T_{\max} nor the extreme T_{\max} show cold or warm biases, as both are centered around zero. Conversely, iERA5 shows a cold bias in both cases. Once again, the precision tends to be larger with more complex DL architectures, with BMdense showing lower RMSEs. The T_{\max} SDR between BM10 and Iberia01 seems to indicate a better agreement than Bilinear, BM1, and BM-

dense. Nevertheless, the four architectures present SDR values closer to 1 when compared with iERA5. Finally, the matching between the four DL architectures outcomes distributions and Iberia01 is greater than for iERA5. The high-quality DL results for temperatures with regard to iERA5 are rather promising, since those variables are assimilated by ERA5.

For Pr, the error metrics from the comparison between the DL downscaled ERA5, iERA5, and Iberia01 are shown in Fig. 6. In this case, between the four architectures, BMdense is the most accurate and precise one, also surpassing iERA5. While the iERA5 shows an overall slightly negative bias (median of -8 mm), the DL outcomes show generally positive values (between 1 and 10 mm). Considering the extreme Pr (98th percentile), all approaches show an underestimation, performing slightly worse than the iERA5, despite the lower error variability ranges. The distributions of the RMSEs of iERA5 and Bilinear, BM1, and BM10 show somewhat similar results for the median and overall variability. In this case, BMdense presents the best results. Finally, the ROCSS shows that all four DL architectures have better skill at representing Pr over the Iberian Peninsula, in comparison with iERA5, with median values ranging between 0.82 and 0.86, which contrasts with 0.67 for iERA5.

The results from Figs. 3 to 6 show that the four DL architectures are successful in downscaling temperature and precipitation from ERA5 at high resolution, presenting, in the vast majority of instances, a better performance than iERA5. Given the similar behavior of the four architectures, choosing the “best” one is not straightforward. BM10 and BMdense show the best precision (RMSEs) for the four variables (with BMdense being the most precise). However, considering the biases, BMdense produces the best results for 10th percentile T_{\min} , Bilinear comes first for the 2nd percentile of T , and BM10 produces more accurate results for T_{\max} . Regarding Pr, BMdense and BM1 retain the best performance for the mean and 98th percentile. Therefore, a clear distinction between architectures for all variables is not meaningful. Therefore, and assuming that all DL architectures are able to partially contribute to the overall performance of the down-scaled datasets, the ensemble-building process considers all DL downscaling equally for each ESGCM.

3.2 Evaluation of DL forced by the ESGCMs

In this section, the error metrics comparing the DL downscaling of the CMIP6 ESGCMs and Iberia01 are displayed for the four analyzed variables (T , T_{\min} , T_{\max} and Pr) and evaluated in the context of the baseline dataset errors, like the CMIP6 ESGCMs at 1° and ERA5 (interpolated at both 1 and 0.1° , henceforth iERA5-1 and iERA5-0.1 or simply iERA5). Note that, for each DL downscaled ESGCM, a four-member ensemble, comprising the results from the four DL architectures, is considered.

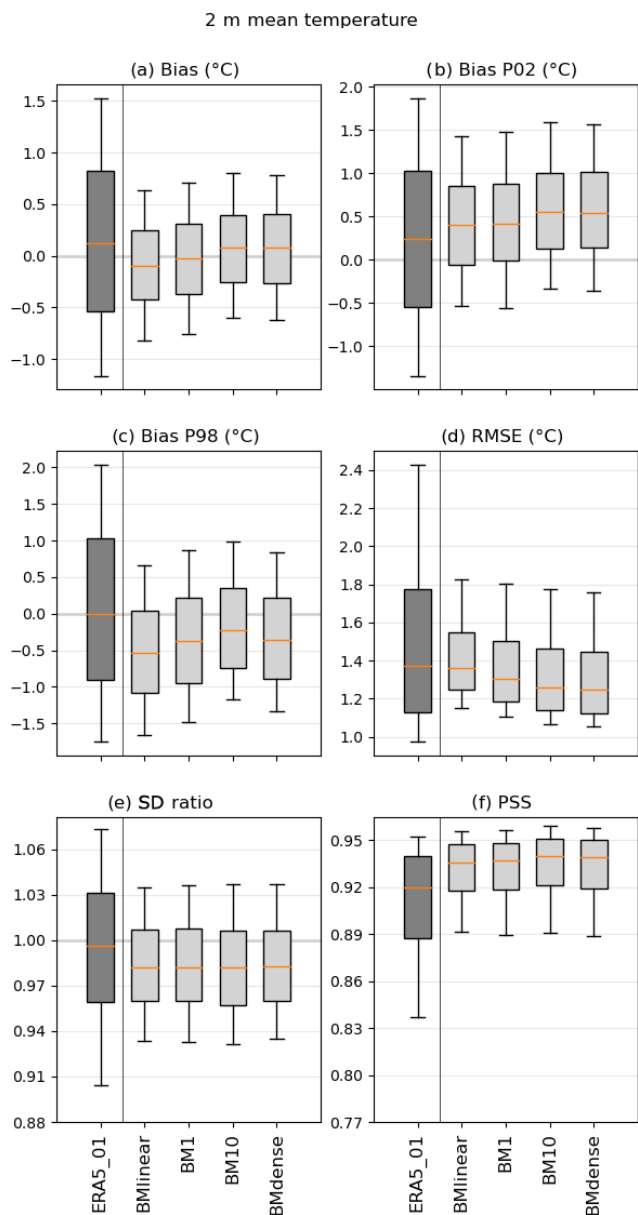


Figure 3. Error measures of the DL downscaling of ERA5 for the daily mean temperature (2010–2014) in relation to the Iberia01 observations. The errors considered are bias, bias of the 2nd and 98th percentiles, root mean square error (RMSE), standard deviation ratio (SDR), and Perkins’ skill score (PSS). As reference, the errors in ERA5 interpolated to 0.1° are also shown. Each box plot represents the value of all grid points of the output of each CNN model forced with ERA5. The box represents the interval between the 25th and 75th percentiles. The orange line is the median value, and the lower (upper) whisker represents the 10th (90th) percentile.

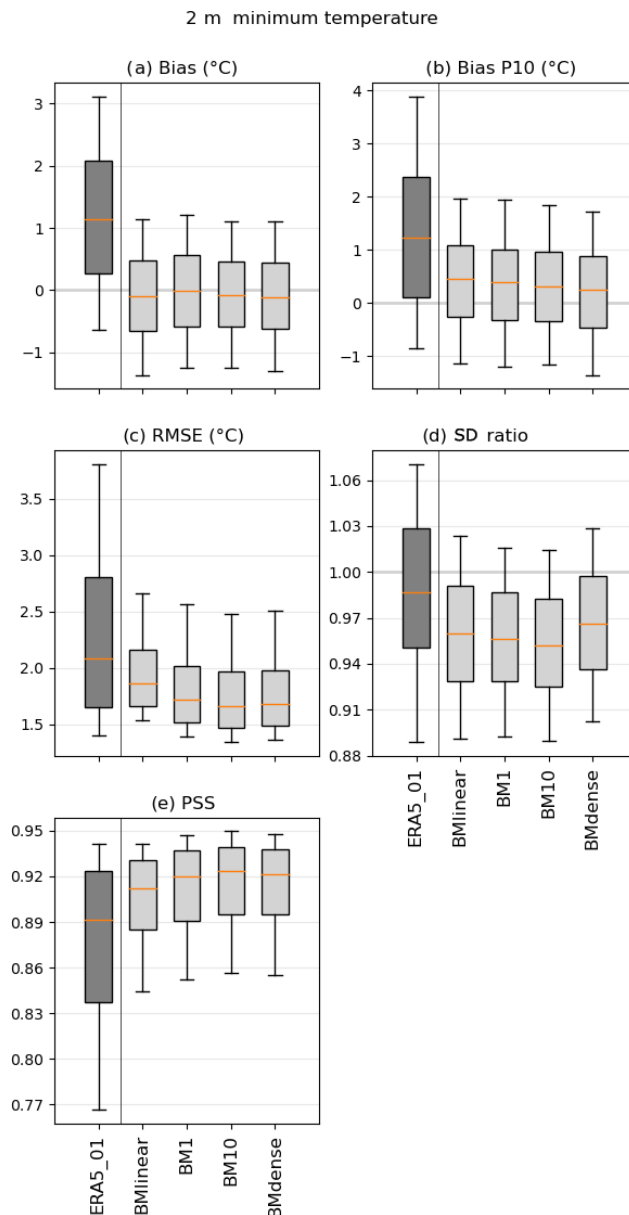


Figure 4. Error measures of the DL downscaling of ERA5 for the daily minimum temperature (2010–2014) in relation to the Iberia01 observations. The errors considered are bias, bias of the 10th percentile, root mean square error (RMSE), standard deviation ratio (SDR), and Perkins’ skill score (PSS). As reference, the errors in ERA5 interpolated to 0.1° are also shown. Each box plot represents the value of all grid points of the output of each CNN model forced with ERA5. The box represents the interval between the 25th and 75th percentiles. The orange line is the median value, and the lower (upper) whisker represents the 10th (90th) percentile.

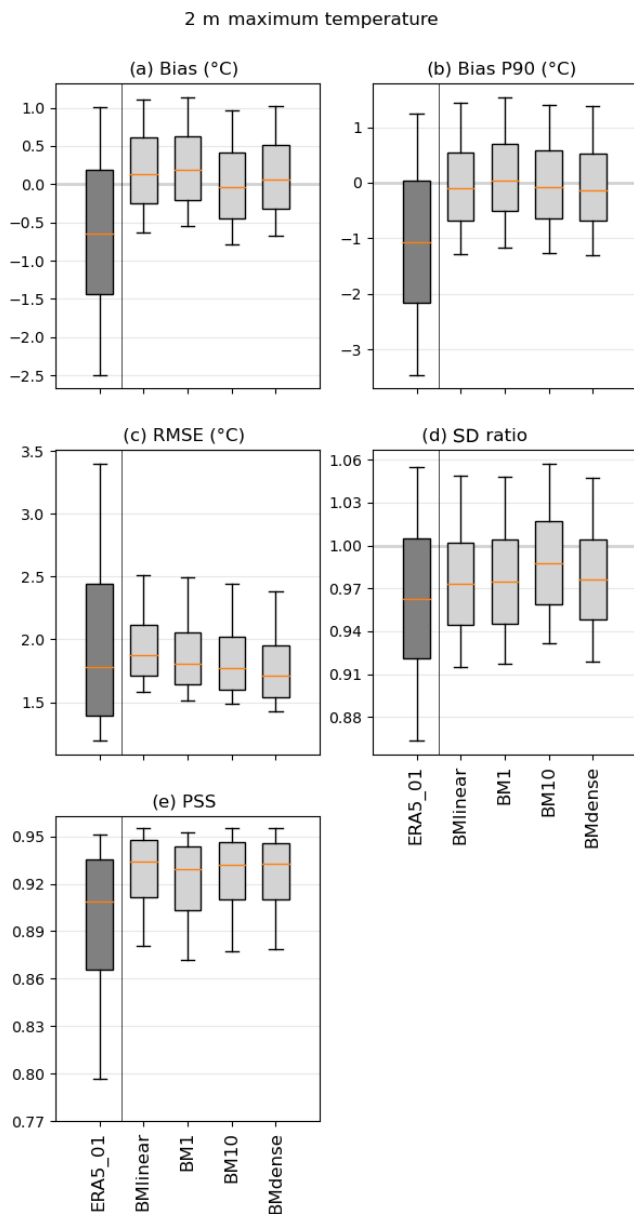


Figure 5. Error measures of the DL downscaling of ERA5 for daily maximum temperature (2010–2014) in relation to the Iberia01 observations. The errors considered are bias, bias of the 90th percentile, root mean square error (RMSE), standard deviation ratio (SDR), and Perkins' skill score (PSS). As reference, the errors in ERA5 interpolated to 0.1° are also shown. Each box plot represents the value of all grid points of the output of each CNN model forced with ERA5. The box represents the interval between the 25th and 75th percentiles. The orange line is the median value, and the lower (upper) whisker represents the 10th (90th) percentile.

In general, for T (Fig. 7), the DL ESGCMs show a much better performance in comparison to the ESGCMs, even with regard to the ERA5, at 0.1° . All biases for the DL ESGCM results are around zero and show small variabilities (below 0.3°C). The forcing ESGCMs display both positive and

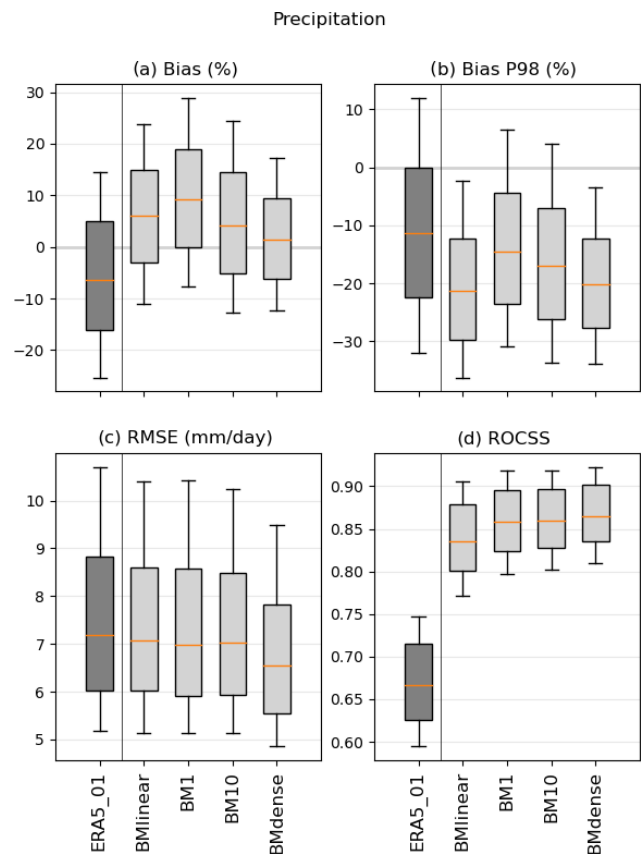


Figure 6. Error measures of the DL downscaling of ERA5 for daily precipitation (2010–2014) in relation to the Iberia01 observations. The errors considered are bias, bias of the 98th percentile, root mean square error (RMSE), and relative operating characteristic skill score (ROCSS). As reference, the errors in ERA5 interpolated to 0.1° are also shown. Each box plot represents the value of all grid points of the output of each CNN model forced with ERA5. The box represents the interval between the 25th and 75th percentiles. The orange line is the median value, and the lower (upper) whisker represents the 10th (90th) percentile.

negative median values for the three biases (total, 2nd, and 98th percentiles), ranging in general between -1 and 2°C , but some rise to 3°C . The medians for iERA5 are generally closer to zero (below 0.3°C). Regarding the RMSE, the DL downscaled ESGCMs show similar values, below 0.5°C , while the iERA5 values are typically below 1°C , and most of the ESGCMs reach almost 4°C , except for MIROC6, which exceeds this threshold. The SDR of all models is around 1; nevertheless, the DL ensembles for each ESGCM present less variability. Finally, the PSS metric shows that the DL ESGCMs are able to represent the Iberia01 PDFs remarkably well, yielding scores above 0.93. The ESGCMs display median PSS values between 0.8 and 0.9 and are characterized by large variability.

Considering T_{\min} (Fig. 8), the biases for the DL ESGCMs results are around zero (ranging no more than 0.5°C), while

the forcing ESGCMs and iERA5 show mainly positive values, with medians reaching 4.5 and 2 °C, respectively. The error variability range for the DL ESGCMs is considerably smaller than for the ESGCMs counterparts. For the extreme T_{\min} values (10th percentile), a similar pattern is visible; however, there are slightly greater biases for the ESGCMs. Regarding RMSEs, the DL ESGCM ensemble shows a great improvement with values around 1 °C, whereas the medians of ESGCMs and ERA5 exceed 4 °C and reach ~2 °C, respectively, accompanied by much larger variability ranges. In terms of SDR, all DL downscaling medians are near 1 and with rather small interquartile ranges when compared with ESGCMs and iERA5, reaching 0.3 units. Finally, the PSS metric consistently reveals the added value of the DL ensemble in representing the PDFs with values ~0.94 that compare with values in the range of 0.70 and 0.87 of the ESGCMs.

Regarding T_{\max} (Fig. 9), the DL ensemble shows a clear improvement with regard to the forcing ESGCMs, with median biases less than -0.2 °C, compared with a general underestimation of median values that reach 4 °C for T_{\max} and its 90th percentile. The MIROC6 is the only model overestimating T_{\max} in ~1 °C. The RMSE values display a striking reduction given by the DL approaches, with median RMSE values ranging from 4 and 1.5 °C to less than 0.5 °C. The SDRs are closer to 1 than the ESGCM counterparts and iERA5. Considering the PSS, similarly to what was previously shown, the DL downscaled inter-member variability ranges between 0.92 and 0.97, contrasting with the forcing ESGCMs and iERA5 (although the median PSS values for iERA5 are also high at above 0.9).

Finally, for precipitation (Fig. 10), the performance of the DL ensembles for each ESGCM is less remarkable than for temperatures. Nevertheless, the DL downscaling outperforms the forcing counterparts in all the error metrics, presenting lower errors and variability ranges. Biases, both for Pr and its 98th percentile, point to a general underestimation, ranging between -25 % and 10 %, yet corresponding to much lower overall differences in comparison to Iberia01 than the ESGCMs and iERA5. For the RMSE, the DL ESGCMs and iERA5 are relatively equivalent, with median values of about 10 mm d⁻¹. However, the ESGCM RMSEs show values above 70 mm d⁻¹ but with medians between 18 and 35 mm d⁻¹. A similar behavior is identifiable for the ROCSS, with good results for both the DL ESGCMs and iERA5, with most median values above 0.95, while the ESGCMs median ROCSS values are in the range of 0.7 and 0.9 and extreme values reach -0.2. In contrast, the PSS values of the DL downscaled ESGCMs show lower values, with medians around 0.5, which are smaller than the ~0.72 of the ESGCMs. In some sense, this is not that surprising since we are comparing the ESGCM and Iberia01 precipitation at 1°, which has a much smoother spatial pattern than at 0.1.

3.3 Iberian future mean climate

The evaluation of the ability of the DL architectures to downscale both the ERA5 and the ESGCMs during the historical climate provided the necessary confidence to apply this method to downscale the future ESGCMs climate simulations. Therefore, here, the projected changes from the DL downscaled ESGCM ensemble are shown, as obtained from the comparison of three future time slices (2015–2040, 2041–2070, and 2071–2100) with the 1981–2010 historical period, in terms of anomalies (i.e., future minus historical). The four SSP–RCP pairs are analyzed (SSP1–2.6, SSP2–4.5, SSP3–7.0, and SSP5–8.5) for each of the four variables. The simple-averaged unweighted ensembles were built considering all ESGCMs and DL architectures. Therefore, the DL ESGCM ensembles are composed of 28 members (seven models times four architectures). Figures 11 to 14 refer to the projections for T , T_{\min} , T_{\max} and Pr, respectively. If fewer than two-thirds of the ESGCMs members agree on the change signal, then the grid point is signalized with a gray dot, which reveals the lack of robustness of the projected change. A spatial comparison between the projected changes from the 1° ESGCMs ensemble, the 0.1° DL downscaled ESGCM ensemble, and the interpolated version of the latter, at 1° (to offer a fair comparison with the original datasets), is conducted to highlight the differences and added value brought by the DL downscaled ensembles.

The future projected changes for T are displayed in Fig. 11 for the forcing 1° ESGCMs (labeled “1° GCM” in the panels) and for the DL downscaled ESGCM ensemble (labeled “DL-MM_01” in the panels) for the three future time slices under the four scenarios. Overall, the results show a projected increase in T , starting from the 2015–2040 period and continuing towards the end of the 21st century (Fig. 11a). Naturally, the SSP1–2.6 (SSP5–8.5) scenario depicts the smallest (greatest) changes. Under SSP1–2.6, projected changes of up to 2.5 °C are discernible, and the patterns exhibit analogous characteristics when comparing the ensemble of ESGCM to the downscaled ensembles generated using DL (Fig. 11a). This similarity is also evident in the remaining scenarios, although with the additional advantage of DL downscaled ESGCM ensembles displaying more detailed patterns of warming. Both DL and ESGCM ensembles demonstrate temperature increases of up to 1.5, 3.5, and 6 °C during the periods 2015–2040, 2041–2070, and 2071–2100, respectively, under the SSP5–8.5 scenario. But the corresponding median warming values for Iberia are around 1.23, 2.5, and 5 °C. In the case of SSP2–4.5 and SSP3–7.0, there is less pronounced warming, although it may still reach up to 3.5 and 4.5 °C, respectively. These results are more easily observed by condensing the spatial information into box plots (Fig. 11b). Overall, differences between DL and ESGCM ensemble are more pronounced from the middle of the century onwards, especially for the two worst-case scenarios (SSP3–7.0 and SSP5–8.5).

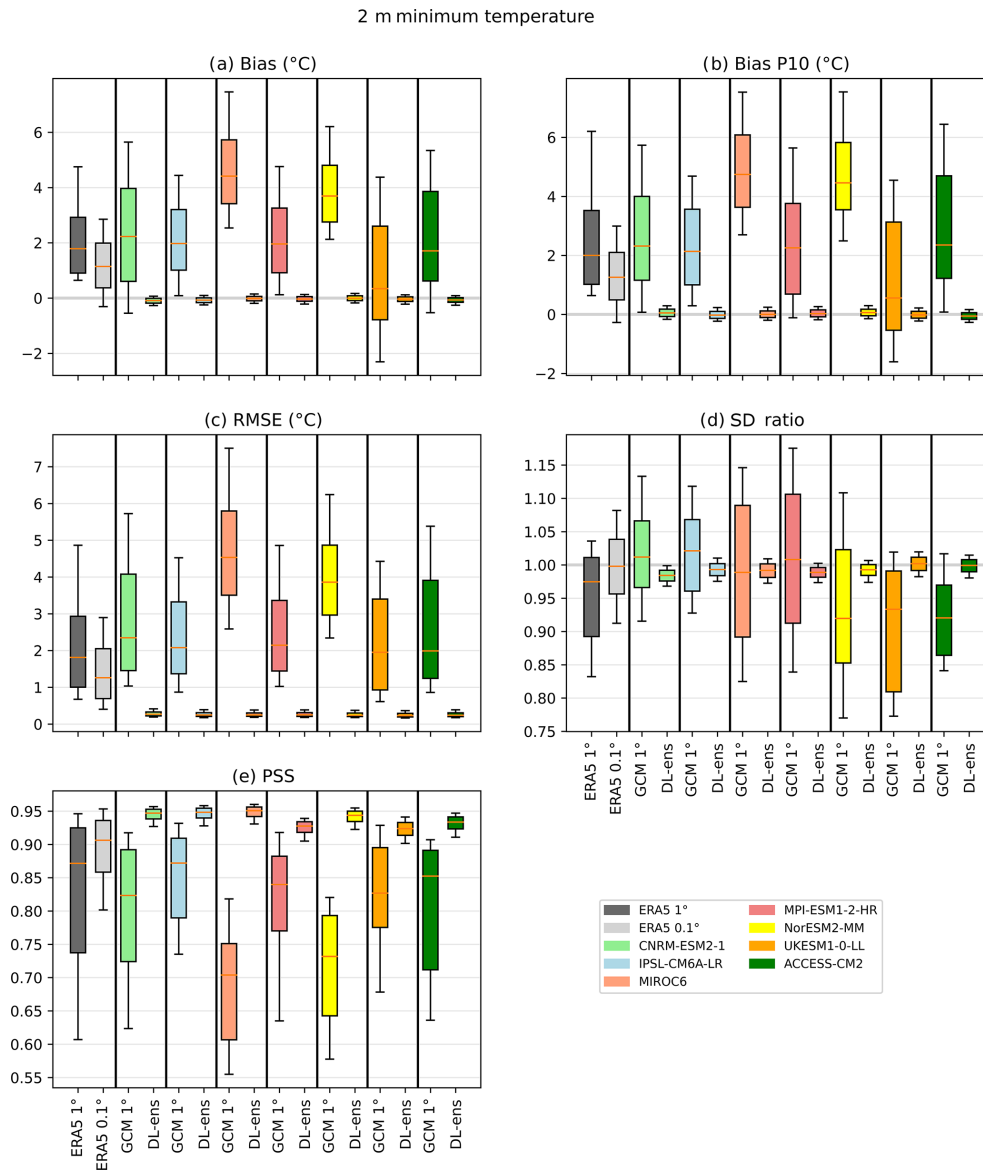


Figure 8. Error measures of the DL downscaling of CMIP6 ESGCMs for daily minimum temperature (1979–2014) in relation to the Iberia01 observations. The errors considered are bias, bias of the 10th percentile, root mean square error (RMSE), Perkins’ skill score (PSS), and standard deviation ratio (SDR). As reference, the errors in ERA5 interpolated to 1 and 0.1° and the errors in CMIP6 ESGCMs at 1° are also shown. Each box plot represents the value of all grid points of the output of all CNN models pooled together forced with each CMIP6 ESGCM. The box represents the interval between the 25th and 75th percentiles. The orange line is the median value, and the lower (upper) whisker represents the 10th (90th) percentile.

present in the ESGCM ensemble projections. This behavior is also depicted in the box plots of Fig. 12b.

T_{\max} (Fig. 13) presents similar characteristics to T and T_{\min} . At the beginning of the 21st century (2015–2040), the magnitude of the projections from both the ESGCM and DL ensemble ranges from 0.5 to 2 °C in most of the Iberian Peninsula (Fig. 13a), independently of the scenario. In the mid-21st century (2041–2070), projections from both the ESGCM and DL downscaled ensembles represent a similar range of projected changes (up to 3.5 °C, depending on the

scenario; Fig. 13b). By 2071–2100, warming values are almost 2-fold those of the middle of the century, surpassing 6.5 °C in the worst-case scenario. It should be highlighted that the DL downscaled ensemble shows different areas of extreme projected increases in T_{\max} (towards south) that are not present in the ESGCM ensemble (where the largest warming is found towards more central and northeastern regions).

Figure 14 shows the Pr projected changes for the future time slices and scenarios, which, in this case, consider

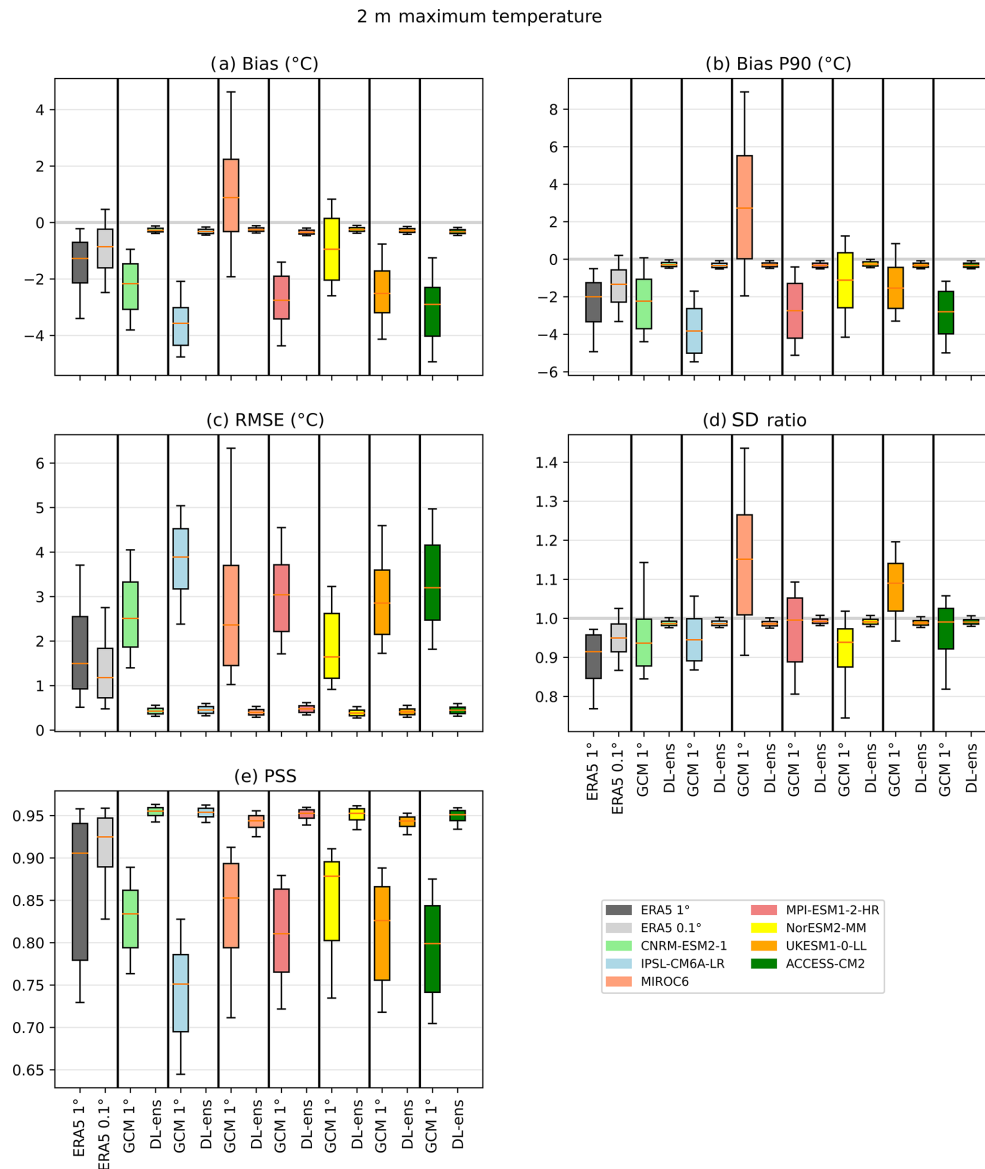


Figure 9. Error measures of the DL downscaling of CMIP6 ESGCMs for daily maximum temperature (1979–2014) in relation to the Iberia01 observations. The errors considered are bias, bias of the 90th percentile, root mean square error (RMSE), standard deviation ratio (SDR), and Perkins' skill score. As reference, the errors in ERA5 interpolated to 1 and 0.1° and the errors in CMIP6 ESGCMs at 1° are also shown. Each box plot represents the value of all grid points of the output of all CNN models pooled together and forced with each CMIP6 ESGCM. The box represents the interval between the 25th and 75th percentiles. The orange line is the median value, and the lower (upper) whisker represents the 10th (90th) percentile.

the mean daily accumulated values and their changes (in mm d^{-1}). In contrast, the changes depicted in Fig. 14 are rather different from the ones in Figs. 11 to 13. While the ESGCM ensemble projects a rather homogeneous decrease in the mean daily precipitation for all future periods and scenarios, the DL downscaled ensemble shows mostly consistent decreases in the western and northern areas of Iberia and non-robust regional increases throughout central and eastern Iberia, independently of the period and scenario. It is important to emphasize that most of these projected increases

are not robust (i.e., fewer than two-thirds of the ensemble members agree on the signal), whereas almost all projected decreases are robust (Fig. 14a). Negative Pr projections are found mainly in the northern, western, and southwestern portions of Iberia, increasing in area and robustness towards 2100 and with the SSP5–8.5 scenario. These features are in overall agreement with the ESGCM ensemble; nevertheless, there is much increased detail due to the enhanced horizontal resolution. In fact, for the 2071–2100 time slice under the SSP5–8.5, the ESGCM (DL downscaled) ensemble shows

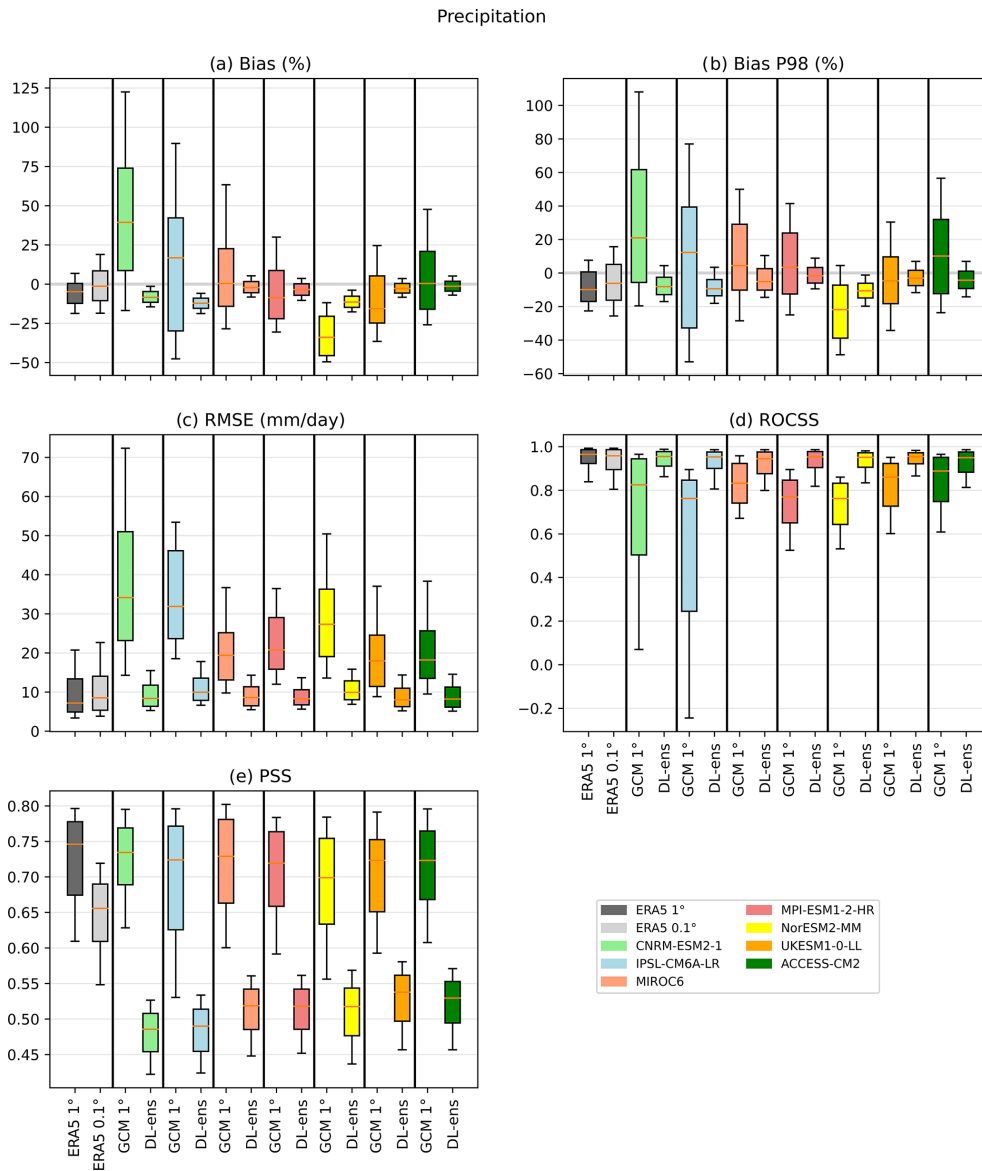


Figure 10. Error measures of the DL downscaling of CMIP6 ESGCMs for daily precipitation (1979–2014) in relation to the Iberia01 observations. The errors considered are bias, bias of the 98th percentile, root mean square error (RMSE), Perkins’ skill score, and relative operating characteristic skill score (ROCSS). As reference, the errors in ERA5 interpolated to 1 and 0.1° and the errors in CMIP6 ESGCMs at 1° are also shown. Each box plot represents the value of all grid points of the output of all CNN models pooled together and forced with each CMIP6 ESGCM. The box represents the interval between the 25th and 75th percentiles. The orange line is the median value, and the lower (upper) whisker represents the 10th (90th) percentile.

projected decreases of down to -0.75 mm d^{-1} (-1 mm d^{-1} in the northern and northwestern Iberia). The box plots in Fig. 14b are largely affected by the compensating effect of different signal projected changes, resulting in overall larger ranges of projected change (even for the interpolated DL ensemble, at 1°) and median values closer to zero, in comparison with the ESGCM ensemble. Nonetheless, an overall decrease in the Iberian precipitation is visible in that the one for the DL ensemble is smaller than the one shown by the forcing ESGCM ensemble.

3.4 Iberian future climate extremes

Considering climate extremes, in this section, the projected changes in the three climate extreme indices are compared for both the ESGCM and the DL downscaled ESGCM ensembles, similar to Sect. 3.3. For T_{\min} and T_{\max} , the 10th and 90th percentiles were considered, respectively, while, for the extreme precipitation, the 95th percentile of the daily mean accumulated values was computed.

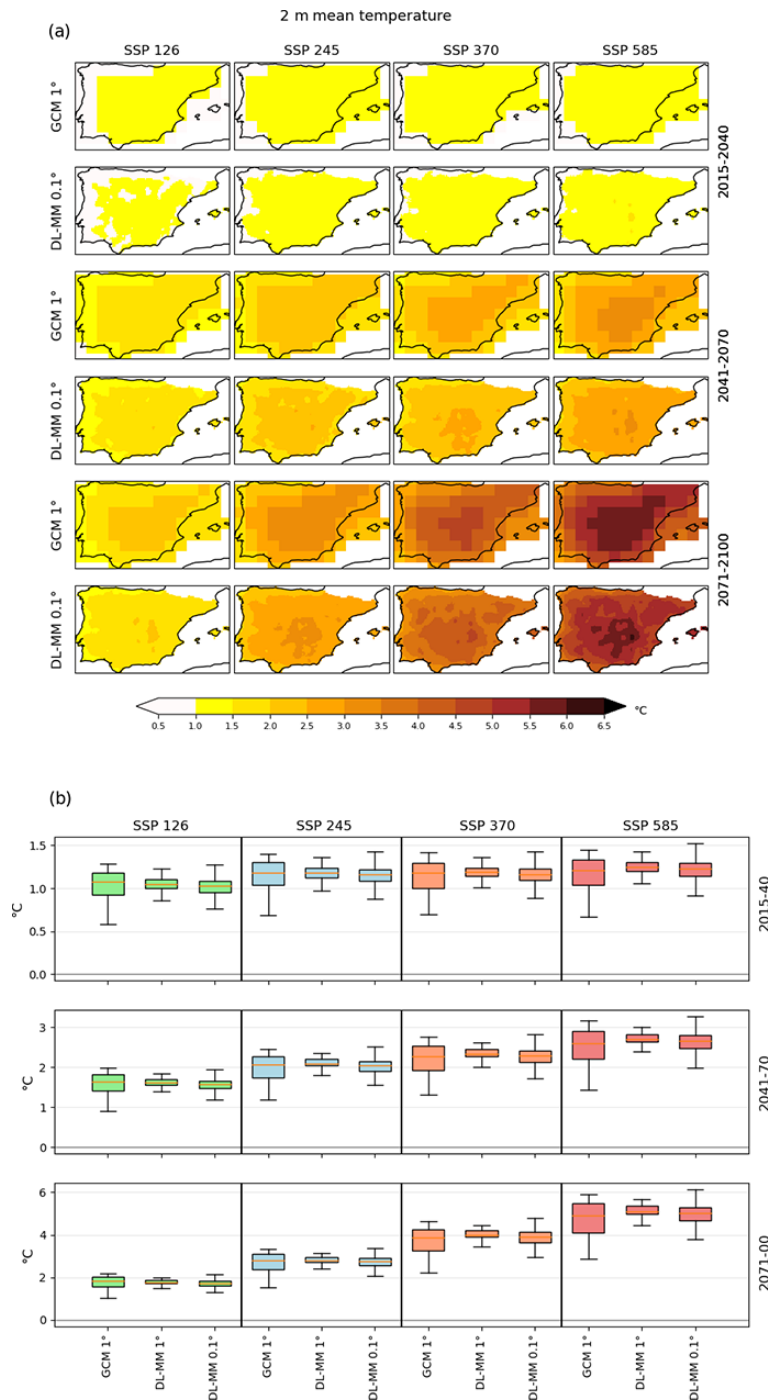


Figure 11. Mean temperature relative changes given by the DL CMIP6 ESGCM multi-model ensemble at 0.1° for SSP1–2.6, SSP2–4.5, SSP3–7.0, and SSP5–8.5 (2015–2040, 2041–2070, 2071–2100 minus 1981–2010)/1981–2100. **(a)** Maps. Gray dots specify grid points where fewer than two-thirds of the DL CMIP6 ESGCM pairs agree on the change signal (no occurrences). **(b)** Box plots. The DL CMIP6 ESGCM multi-model ensembles were interpolated to 1°, and the results are also displayed. As reference, the change signal linked to the ESGCM ensemble at 1° is also shown in panels **(a)** and **(b)**. Each box plot represents the value of all grid points of the output of all CNN models pooled together and forced with all CMIP6 ESGCMs pooled together. The box represents the interval between the 25th and 75th percentiles. The orange line is the median value, and the lower (upper) whisker represents the 10th (90th) percentile.

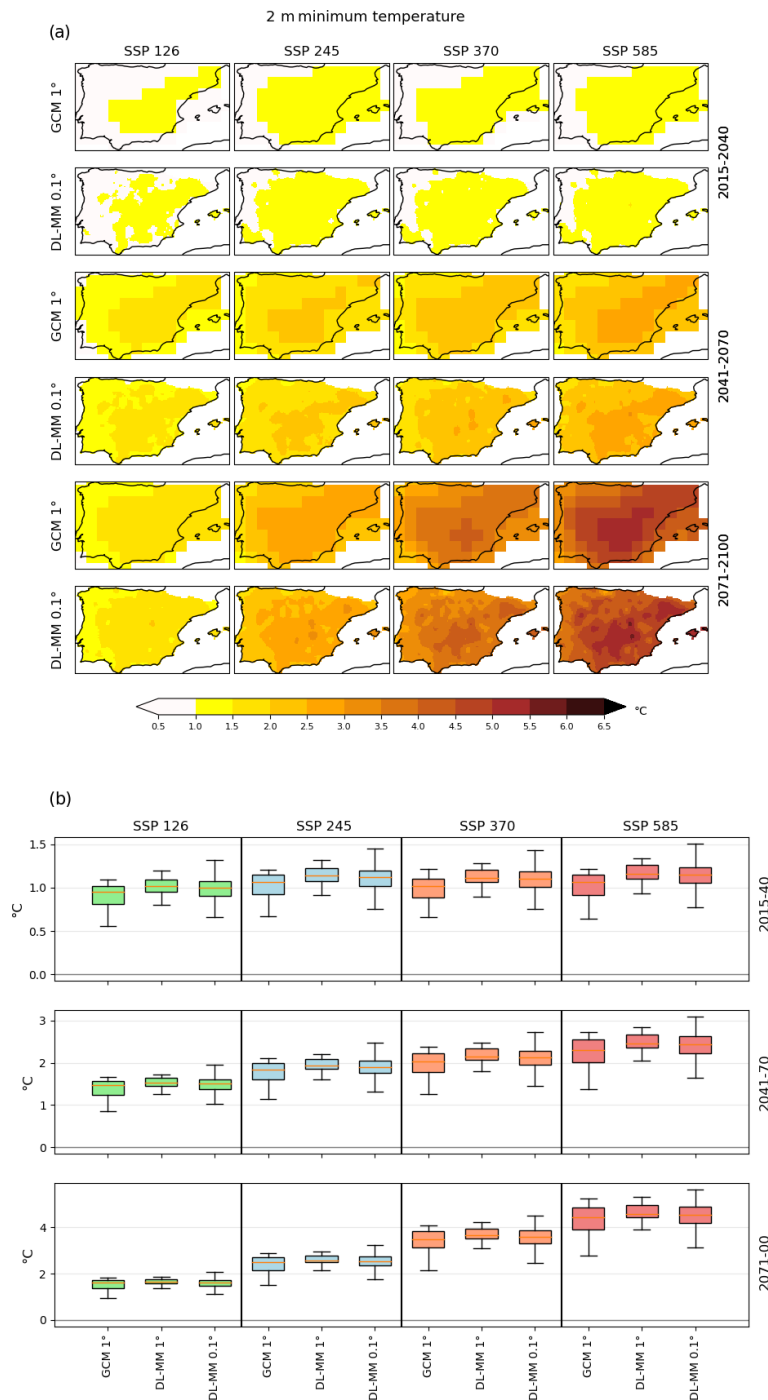


Figure 12. Minimum temperature relative changes given by the DL CMIP6 ESGCM multi-model ensemble at 0.1° for SSP1–2.6, SSP2–4.5, SSP3–7.0, and SSP5–8.5 (2015–2040, 2041–2070, 2071–2100 minus 1981–2010)/1981–2100. **(a)** Maps. Gray dots specify grid points where fewer than two-thirds of the DL CMIP6 ESGCM pairs agree on the change signal (no occurrences). **(b)** Box plots. The DL CMIP6 ESGCM multi-model ensembles were interpolated to 1°, and the results are also displayed. As reference, the change signal linked to the ESGCM ensemble at 1° is also shown in panels **(a)** and **(b)**. Each box plot represents the value of all grid points of the output of all CNN models pooled together and forced with all CMIP6 ESGCMs pooled together. The box represents the interval between the 25th and 75th percentiles. The orange line is the median value, and the lower (upper) whisker represents the 10th (90th) percentile.

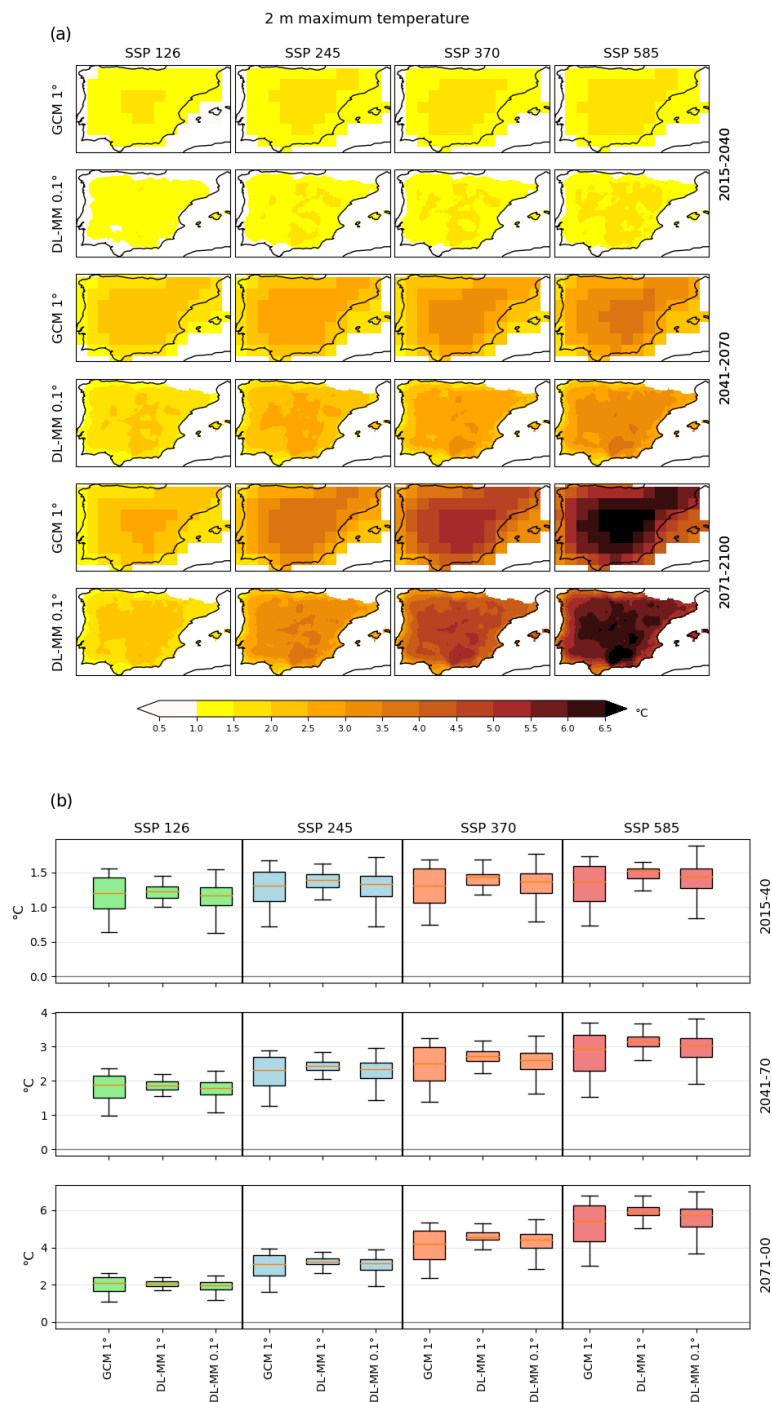


Figure 13. Maximum temperature relative changes given by the DL CMIP6 ESGCM multi-model ensemble at 0.1° for SSP1–2.6, SSP2–4.5, SSP3–7.0, and SSP5–8.5 (2015–2040, 2041–2070, 2071–2100 minus 1981–2010)/1981–2010. **(a)** Maps. Gray dots specify grid points where fewer than two-thirds of the DL CMIP6 ESGCM pairs agree on the change signal (no occurrences). **(b)** Box plots. The DL CMIP6 ESGCM multi-model ensembles were interpolated to 1° , and the results are also displayed. As reference, the change signal linked to the ESGCM ensemble at 1° is also shown in panels **(a)** and **(b)**. Each box plot represents the value of all grid points of the output of all CNN models pooled together and forced with all CMIP6 ESGCMs pooled together. The box represents the interval between the 25th and 75th percentiles. The orange line is the median value, and the lower (upper) whisker represents the 10th (90th) percentile.

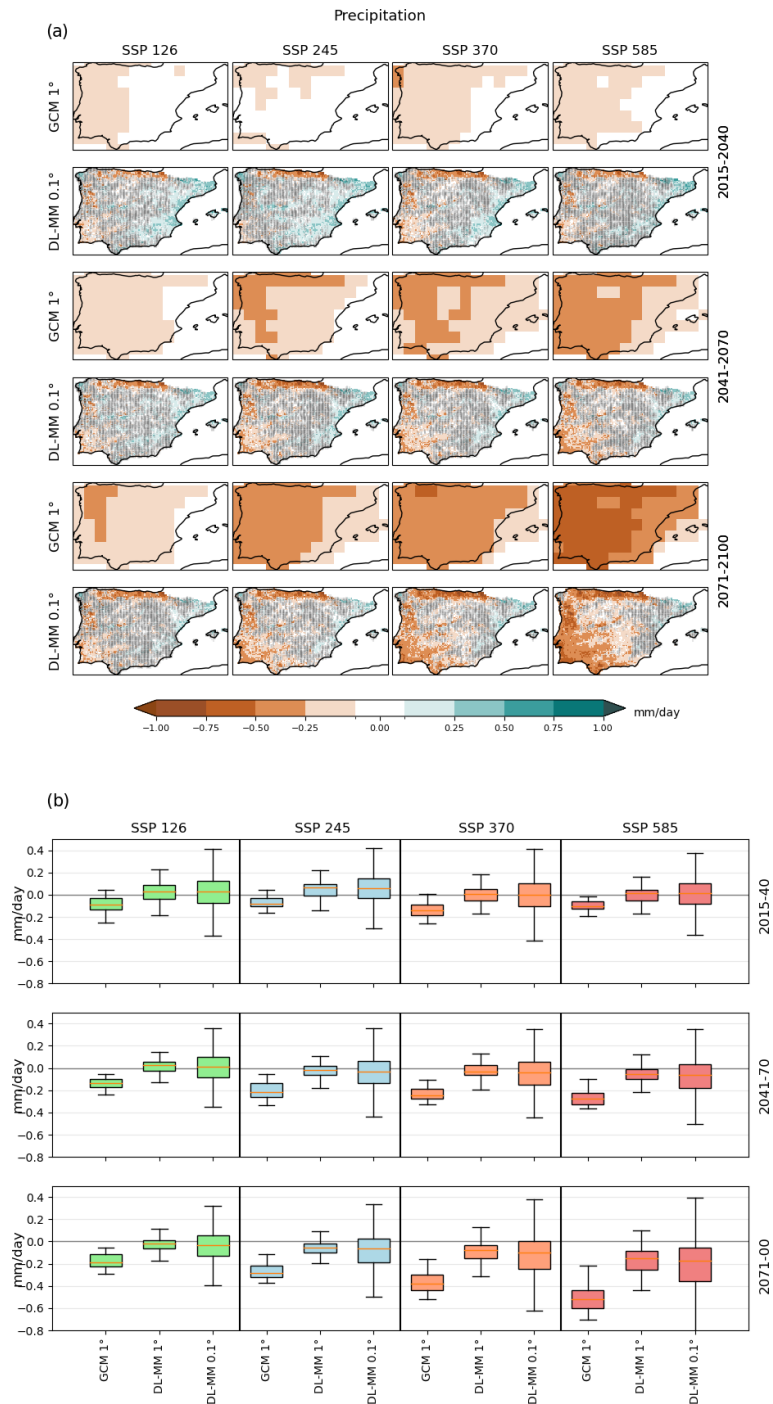


Figure 14. Daily mean precipitation relative changes given by the DL CMIP6 ESGCM multi-model ensemble at 0.1° for SSP1–2.6, SSP2–4.5, SSP3–7.0, and SSP5–8.5 (2015–2040, 2041–2070, 2071–2100 minus 1981–2010)/1981–2100. **(a)** Maps. Gray dots specify grid points where fewer than two-thirds of the DL CMIP6 ESGCM pairs agree on the change signal (no occurrences). **(b)** Box plots. The DL CMIP6 ESGCM multi-model ensembles were interpolated to 1° , and the results are also displayed. As reference, the change signal linked to the ESGCM ensemble at 1° is also shown in panels **(a)** and **(b)**. Each box plot represents the value of all grid points of the output of all CNN models pooled together and forced with all CMIP6 ESGCMs pooled together. The box represents the interval between the 25th and 75th percentiles. The orange line is the median value, and the lower (upper) whisker represents the 10th (90th) percentile.

In general, the future 10th percentile of T_{\min} (Fig. 15) reveals lower warming projections than for T_{\min} (Fig. 12) and also different patterns. The most pronounced warmings are located in the south-central and eastern regions of Iberia (Fig. 14a), which may reach 2 °C (4 °C) in the 2041–2070 (2071–2100) period for the SSP5–8.5 scenario. The remaining scenarios show lower warmings, reaching 1.5, 2.5, and 3.5 °C for SSP1–2.6, SSP2–4.5, and SSP3–7.0, respectively, by the end of the century. As expected, the warming patterns are much more detailed and localized when the DL ensemble is considered.

Similar to extreme T_{\min} , the projections of extreme T_{\max} (Fig. 16) exhibit comparable patterns to those of the mean climate variable (Fig. 13), although with much more pronounced warming values. In particular, over a significant portion of Iberia, the warming reaches over 8 °C by the end of the century for the SSP5–8.5 scenario. The use of a precise and performance-evaluated technique to downscale a large ensemble of ESGCM climate projections at a high resolution provides substantial added value in capturing local climate change for the 90th percentile of T_{\max} , as demonstrated in Fig. 16a. For instance, when considering the 2071–2100 period under SSP5–8.5, both the ESGCM and DL downscaled ensembles project changes exceeding 8 °C. However, the DL downscaled ensemble surpasses this threshold over a wider area, locally exceeding 9 °C and even extending to the southern coast of Iberia, where the projections from the ESGCM ensemble do not surpass 6 °C.

Regarding the extreme Pr (Fig. 17), the DL projections point to reductions in the extreme precipitation across southwestern Iberia, expanding eastward (for part of southern Iberia) throughout the 21st century, and more pronounced for the SSP3–7.0 and SSP5–8.5 scenarios. These decreases can reach more than 3 mm d⁻¹ in these regions. On the other hand, essentially over central, southeastern, and northwestern Iberia, DL projections show an intensification in extreme precipitation in all scenarios and time periods, reaching increases that surpass 3 mm d⁻¹. The ESGCMs projections mostly present decreases in extreme precipitation (for all time slices and scenarios, except the SSP1–2.6 during 2015–2040 and 2041–2070 in the northern half of the peninsula; Fig. 17a). The spatial pattern of changes in extreme precipitation are dissimilar for the DL and ESGCMs projections. The box plots in Fig. 17b summarize the differences between the ESGCM and DL downscaled ensembles, being, nevertheless, slightly affected by the lack of robustness of some of the outcomes, as the DL ensemble shows a larger variability than the ESGCM ensemble.

4 Discussion and conclusions

The Iberian Peninsula, situated in the southwestern tip of the European continent within the Mediterranean region is considered a climate change hotspot, due to the projected fu-

ture warming and drying conditions. These changes can significantly impact the natural environment and human health in the region (Giorgi, 2006; Soares et al., 2017a, b; Cramer et al., 2018; Lionello and Scarascia, 2018; Cardoso et al., 2019; Tuel and Eltahir, 2020; Soares and Lima, 2022; Lima et al., 2023a, b; Soares et al., 2023a). Consequently, there is an urgent need for accurate climate information to assist the planning and development of adaptation strategies. Recent climate change studies focusing on the Iberian Peninsula relied on RCM simulations forced by CMIP5 GCMs (Soares et al., 2017a, b; Cardoso et al., 2019; Lima et al., 2023a, b; Soares et al., 2023a) to project future climate change with increased resolution, accounting for regional features not captured by coarse ESGCMs. However, following the release of the improved CMIP6 global climate simulations and projections (in the context of the most recent IPCC report, AR6; IPCC, 2021), the need for an updated climate change assessment in the Iberia Peninsula arose. The new high-resolution CMIP6 EURO-CORDEX regional climate simulations and projections will become available within the next 1–2 years. In the interim, however, there is a need for high-resolution climate information to accurately assess future projections over Iberia. In this context, an opportunity emerges to explore alternative approaches to downscale the current CMIP6 simulations and projections. Therefore, this study leverages innovative AI methods to evaluate the evolution of mean, minimum, and maximum temperatures, as well as precipitation, across the Iberian Peninsula, throughout the 21st century. The analysis is based on a multi-model, multi-architecture ensemble of CNN-based downscaled projections derived from CMIP6 ESGCMs. The investigation encompasses three distinct future time slices (2015–2040, 2041–2070, and 2071–2100) in line with four SSP–RCP scenarios of SSP1–2.6, SSP2–4.5, SSP3–7.0, and SSP5–8.5.

First, the ability of four DL architectures to reproduce the historical T , T_{\min} , T_{\max} , and Pr climates was evaluated over Iberia during 2010–2014 (Figs. 3 to 6). During this period, all DL architectures, trained using ERA5 data between 1979 and 2004, revealed a good agreement with observations (Iberia01) for the predictand variables (using solely the predictors as input data). Although more complex architectures, such as BMdense, revealed better performance for Pr (lower overall biases and RMSEs and higher ROCSS), a clear distinction between architectures was not meaningful. The results showed that during the 1979–2014 historical period, the DL downscaled ESGCMs were able to represent the Iberia01 reference climate with large increased performance in comparison with the forcing ESGCMs, even when compared with the ERA5 and iERA5 datasets (Figs. 7 to 10). For Pr, nevertheless, the downscaled error metrics were shown to be similar to the reanalysis ones, despite greater differences in the overall variable distributions (as shown by the PSS values). Such disagreement could be related to the singular behavior of Pr, especially considering its extreme events, which can occur under distinct atmospheric synoptic patterns (pre-

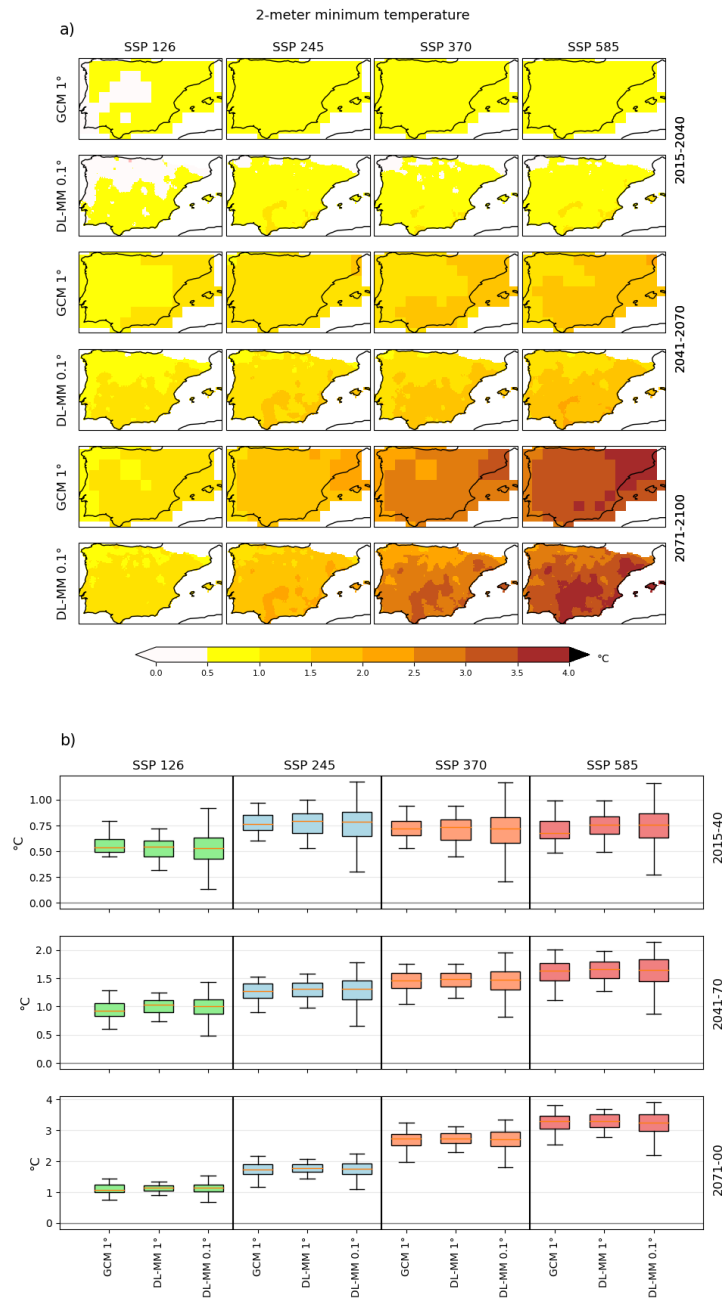


Figure 15. Mean minimum temperature 10th percentile relative changes given by the DL CMIP6 ESGCM multi-model ensemble at 0.1° for SSP1–2.6, SSP2–4.5, SSP3–7.0, and SSP5–8.5 (2015–2040, 2041–2070, 2071–2100 minus 1981–2010)/1981–2010. **(a)** Maps. Gray dots represent grid points where fewer than two-thirds of the DL CMIP6 ESGCM pairs agree on the change signal (no occurrences). As reference, the change signal linked to all ESGCMs at 1° is also shown. **(b)** Box plots. Each box plot represents the value of all grid points of the output of all CNN models pooled together and forced with all CMIP6 ESGCMs pooled together. The box represents the interval between the 25th and 75th percentiles. The orange line is the median value, and the lower (upper) whisker represents the 10th (90th) percentile.

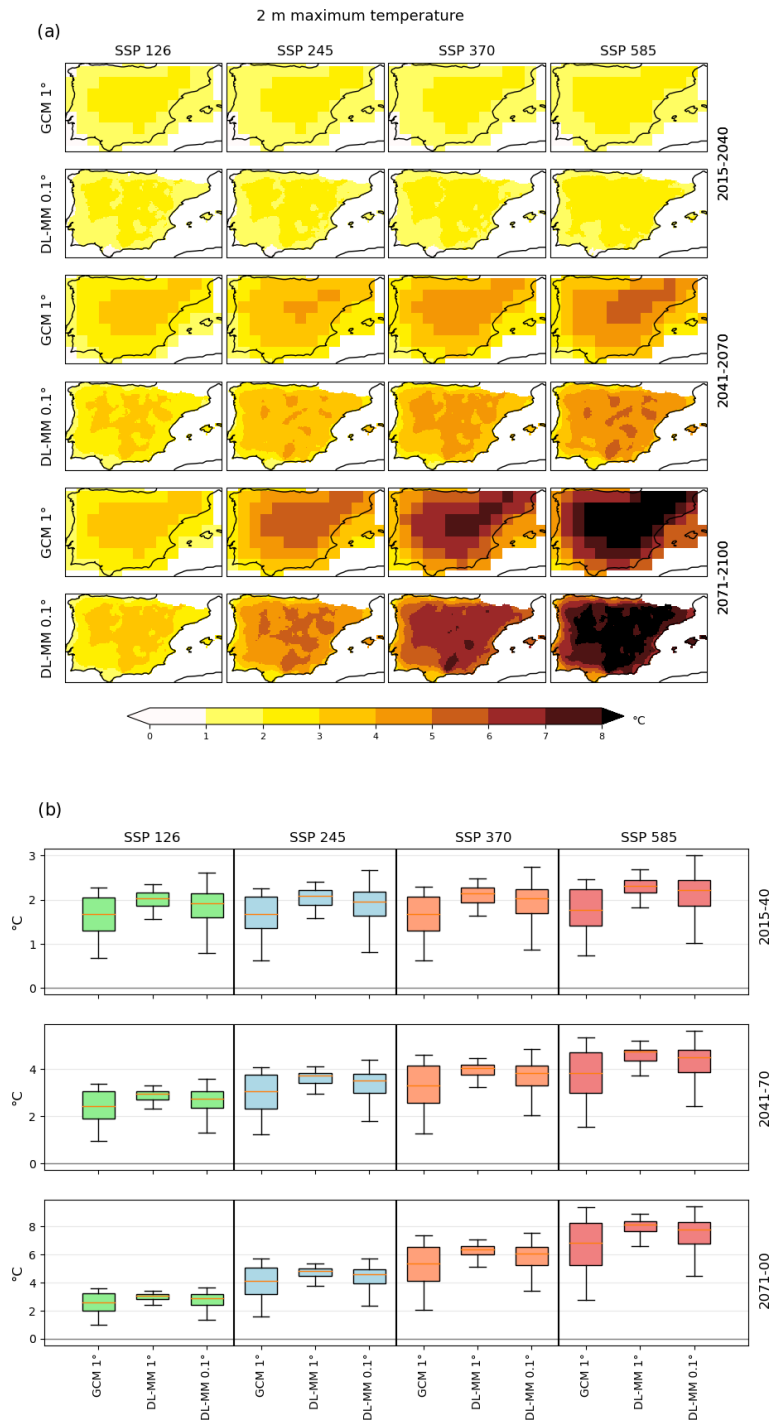


Figure 16. Mean maximum temperature 90th percentile relative changes given by the DL CMIP6 ESGCM multi-model ensemble at 0.1° for SSP1–2.6, SSP2–4.5, SSP3–7.0, and SSP5–8.5 (2015–2040, 2041–2070, 2071–2100 minus 1981–2010)/1981–2010. **(a)** Maps. Gray dots represent grid points where fewer than two-thirds of the DL CMIP6 ESGCM pairs agree on the change signal (no occurrences). As reference, the change signal linked to all ESGCMs at 1° is also shown. **(b)** Box plots. Each box plot represents the value of all grid points of the output of all CNN models pooled together and forced with all CMIP6 ESGCMs pooled together. The box represents the interval between the 25th and 75th percentiles. The orange line is the median value, and the lower (upper) whisker represents the 10th (90th) percentile.

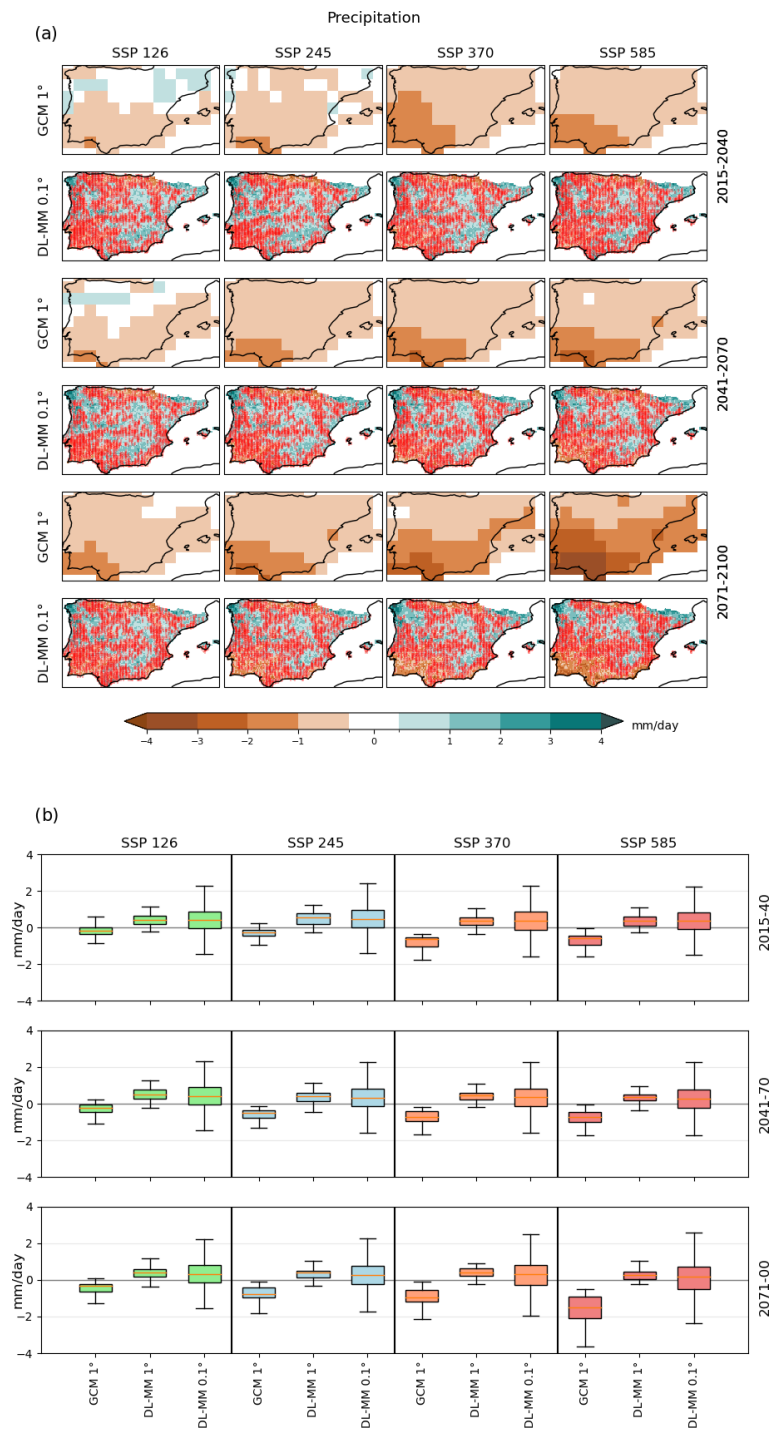


Figure 17. Precipitation 95th percentile relative changes given by the DL CMIP6 ESGCM multi-model ensemble at 0.1° for SSP1–2.6, SSP2–4.5, SSP3–7.0, and SSP5–8.5 (2015–2040, 2041–2070, 2071–2100 minus 1981–2010)/1981–2010. **(a)** Maps. Red dots represent grid points where fewer than two-thirds of the DL CMIP6 ESGCM pairs agree on the change signal. As reference, the change signal linked to all ESGCMs is also shown. **(b)** Box plots. Each box plot represents the value of all grid points of the output of all CNN models pooled together and forced with all CMIP6 ESGCMs pooled together. The box represents the interval between the 25th and 75th percentiles. The orange line is the median value, and the lower (upper) whisker represents the 10th (90th) percentile.

dicator sets), thus becoming challenging for the DL architectures to establish empirical relationships between the vertical atmospheric structure and surface-level precipitation accumulation. Overall, the evaluation of the DL downscaled ESGCMs showed a rather good performance in representing the historical climate (mean, minimum, and maximum temperatures and precipitation), providing the necessary confidence to project the future climate change under different scenarios using this new approach. It should be noted that the DL models trained with predictors from ERA5 were used to generate the ESGCMs output for the historical and future periods. The bias correction procedure applied to the ESGCM predictors is an important asset that may allow their values to better agree with those from ERA5. As a result, we believe that the relationship between the predictors of both ERA5 and ESGCMs and Iberia01 are comparable.

The DL downscaled T projections revealed a projected increase between 1 and 1.5 °C over Iberia (Fig. 11), for all scenarios, during the earliest future period (2011–2040). By the end of the 21st century, however, the DL ensemble projected changes were shown to become more heterogeneous between scenarios and generally varied between 1.5 °C (SSP1–2.6) and 5 °C (SSP5–8.5). In all instances, the DL downscaled T projected changes showed a strong agreement with the original CMIP6 ESGCM ensemble in both the signal and main spatial patterns of climate change. Nevertheless, regional-to-local features are clearly enhanced by the increased resolution. In fact, the most poignant difference between the DL and the original ensembles is the horizontal discretization. Local differences can be identified in the most inland areas of the Iberian Peninsula in the DL results, which are not captured by the original ESGCMs, due to the coarse grid that neglects valleys and other geographically enclosed areas, thus fostering greater horizontal heterogeneities. Similar features were found for T_{\min} (Fig. 12) and T_{\max} (Fig. 13) and for the respective extreme values (10th percentile of T_{\min} and 90th percentile of T_{\max} in Figs. 15 and 16, respectively). Between those, T_{\max} showed larger projected increases than T_{\min} (approximately 2-fold), revealing greater intra-daily temperature ranges to be expected in the future. For the extreme T_{\max} values, even under the SSP1–2.6 “optimistic” scenario, DL downscaled projections revealed increases exceeding 3 °C by the end of the 21st century. For the “less optimistic” ones (SSP3–7.0 and SSP5–8.5), extreme T_{\max} projected changes of up to 7 °C were shown for most of Iberia (Fig. 16). DL downscaled projections for extreme T_{\min} , on the other hand, were shown to be higher in the eastern and southern Iberia (Fig. 15), locally surpassing 3.5 °C (for 2071–2100 under SSP5–8.5). Overall, these projections are aligned with the EURO-CORDEX ensemble projections for Iberia (Soares et al., 2017a; Cardoso et al., 2019; Lima et al., 2023a; Amblar et al., 2017) but with small value differences, which are also linked to the dissimilarities regarding the emission scenarios.

The significance of DL downscaling techniques in the context of Pr projections unveiled further intricacies when com-

pared to T , T_{\min} , and T_{\max} projections. Given that the behavior of daily mean accumulated precipitation is heavily influenced by local topography and other phenomena, particularly owing to convective processes, which can result in local, large precipitation accumulations, projecting Pr was shown to be more complex for DL downscaling methods, considering the widespread continental, mountainous areas of the Iberian Peninsula. Therefore, for both Pr (Fig. 14) and extreme Pr (95th percentile; Fig. 17), the original and DL downscaled ESGCM ensemble projections showed greater discrepancies. While the ESGCM projected changes showed essentially negative values, corresponding to the future large-scale expected drying over Iberia, the DL results revealed a drying trend in western and southwestern regions of Iberia, which are stronger for the upper-end scenarios, and local projected increases, mostly in the central and eastern continental regions. The southern and western precipitation reductions are consistent, with a significant reduction in the large-scale precipitation from frontal activity, due to the northern displacement of the storm tracks (Tamarin and Kaspi, 2017). In fact, the northward expansion of the Hadley cell lead to a northward shift in the storm tracks over the North Atlantic, resulting in a reduction in the large-scale precipitation across southern and western Iberia (Bengtsson et al., 2006; Harvey et al., 2014; Kang and Lu, 2012; Ulbrich et al., 2008). The projected local increases in the precipitation, although non-robust, mostly with fewer than two-thirds of ensemble members in agreement, may be consistent with local-to-regional changes in convective precipitation that are not captured by the original ESGCMs, highlighting potential applications of DL techniques to long-term projections (or even short-term forecasting).

The projected changes in the warming and drying over Iberia, as reported in recent studies using previous CMIP outputs (CMIP5) and in the most recent IPCC report (IPCC, 2021), are consistent with the multi-model, multi-scenario, and multi-architecture DL downscaled ESGCM ensemble projections presented in this study. This behavior is also in accordance with the resulting DL climate change signals from the Baño-Medina et al. (2022) report for Iberia, which showed similar spatial patterns to those obtained from the CMIP5 RCMs but with local-to-regional added value. Previous research has demonstrated that the warming and drying trends over Iberia are more pronounced under high anthropogenic emission scenarios, reflecting the influence of human activities on climate change, compared to the natural variability in the climate system. Our results demonstrated that in the CMIP6 context, within the new set of scenarios encompassing socioeconomic and representative concentration pathways, AI-based DL methods are able to accurately simulate the historical Iberian climate and produce high-resolution scenario-based projections, consistent with each other and with previous studies, using (coarse) GCM forcing and a high-resolution training database. Thus, the present study highlighted the substantial advantages of employing

novel approaches based on DL to obtain efficiently up-to-date, high-resolution climate information at a local scale that is specifically for Iberia. This is crucial for supporting and designing mitigation and adaptation strategies.

Code and data availability. The datasets generated and/or analyzed during the current study are available through the Earth System Grid Federation (ESGF) Lawrence Livermore National Laboratory (LLNL) repository for CMIP6 ESGCMs (<https://esgf-node.llnl.gov/projects/esgf-llnl/>, Cinquini et al., 2014) and the Copernicus Climate Change Service (C3S) and Copernicus Data Store (CDS) repository for ERA5 (<https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels?tab=overview>, last access: June 2023; <https://doi.org/10.24381/cds.adbb2d47>, Hersbach et al., 2023). The Iberia01 dataset is publicly available through the DIGITAL.CSIC open-science service (Herrera et al., 2019; Gutiérrez et al., 2019b, <https://doi.org/10.20350/digitalCSIC/8641>). The DL configuration and both input and output data are freely available as Zenodo repositories (<https://doi.org/10.5281/zenodo.8314980>, Soares et al., 2023b; <https://doi.org/10.5281/zenodo.8338468>, Soares et al., 2023c; <https://doi.org/10.5281/zenodo.8340234>, Soares et al., 2023d; <https://doi.org/10.5281/zenodo.8340250>, Soares et al., 2023e; <https://doi.org/10.5281/zenodo.8340266>, Soares et al., 2023f; <https://doi.org/10.5281/zenodo.8340274>, Soares et al., 2023g; <https://doi.org/10.5281/zenodo.8340279>, Soares et al., 2023h; <https://doi.org/10.5281/zenodo.8340287>, Soares et al., 2023i; <https://doi.org/10.5281/zenodo.8340297>, Soares et al., 2023j; <https://doi.org/10.5281/zenodo.8340318>, Soares et al., 2023k; <https://doi.org/10.5281/zenodo.8340338>, Soares et al., 2023l).

Author contributions. The conceptualization was done by PMMS. CMIP6 ESGCM data were acquired and processed by DCAL and GL. ERA5 and Iberia01 data were acquired by FJ. The formal analysis and visualization were done by FJ. All authors contributed to the writing, reviewing, and editing of the paper.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. The authors would like to acknowledge the financial support of the Portuguese Fundação para a Ciência e a Tecnologia (FCT) I.P./MCTES through national funds (PIDDAC) – UIDB/50019/2020 (<https://doi.org/10.54499/UIDB/50019/2020>), UIDP/50019/2020 (<https://doi.org/10.54499/UIDP/50019/2020>),

and LA/P/0068/2020 (<https://doi.org/10.54499/LA/P/0068/2020>). The authors also acknowledge the EEA Financial Mechanism 2014–2021 and the Portuguese Environment Agency through Pre defined Project 2 National Roadmap for Adaptation XXI (PDP 2) and project DHEFEUS 2022.09185.PTDC. Frederico Johannsen has been supported by FCT under the research grant UI/BD/151498/2021. Daniela C. A. Lima has been supported by the Scientific Employment Stimulus 5th edition from Fundação para a Ciência e a Tecnologia (FCT, 2022.03183.CEECIND).

Financial support. This research has been funded by the Portuguese Fundação para a Ciência e a Tecnologia (FCT) I.P./MCTES through national funds (PIDDAC) – UIDB/50019/2020 (<https://doi.org/10.54499/UIDB/50019/2020>), UIDP/50019/2020 (<https://doi.org/10.54499/UIDP/50019/2020>), and LA/P/0068/2020 (<https://doi.org/10.54499/LA/P/0068/2020>).

Review statement. This paper was edited by Lele Shu and reviewed by Carolyn Begeman and one anonymous referee.

References

- Alzubi, J., Nayyar, A., and Kumar, A.: Machine Learning from Theory to Algorithms: An Overview, *J. Phys. Conf. Ser.*, 1142, 012012, <https://doi.org/10.1088/1742-6596/1142/1/012012>, 2018.
- Amblar, M. P., Casado Calle, M. J., Pastor Saavedra, M. A., Ramos Calzado, P., and Rodríguez Camino, E.: Guía de escenarios regionalizados de cambio climático sobre España a partir de los resultados del IPCC-AR5, Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente Agencia Estatal de Meteorología, Madrid, <https://doi.org/10.31978/014-17-010-8>, 2017.
- Argüeso, D., Hidalgo-Muñoz, J. M., Gámiz-Fortis, S. R., Esteban-Parra, M. J., and Castro-Díez, Y.: High-resolution projections of mean and extreme precipitation over Spain using the WRF model (2070–2099 versus 1970–1999), *J. Geophys. Res.-Atmos.*, 117, D12108, <https://doi.org/10.1029/2011JD017399>, 2012.
- Baño-Medina, J., Manzanar, R., and Gutiérrez, J. M.: Configuration and intercomparison of deep learning neural models for statistical downscaling, *Geosci. Model Dev.*, 13, 2109–2124, <https://doi.org/10.5194/gmd-13-2109-2020>, 2020.
- Baño-Medina, J., Manzanar, R., Cimadevilla, E., Fernández, J., González-Abad, J., Cofiño, A. S., and Gutiérrez, J. M.: Downscaling multi-model climate projection ensembles with deep learning (DeepESD): contribution to CORDEX EUR-44, *Geosci. Model Dev.*, 15, 6747–6758, <https://doi.org/10.5194/gmd-15-6747-2022>, 2022.
- Bauer, P., Dueben, P. D., Hoefler, T., Quintino, T., Schulthess, T. C., and Wedi, N. P.: The digital revolution of Earth-system science, *Nature Computational Science*, 1, 104–113, <https://doi.org/10.1038/s43588-021-00023-0>, 2021.
- Bengtsson, L., Hodges, K. I., and Roeckner, E.: Storm Tracks and Climate Change, *J. Climate*, 19, 3518–3543, <https://doi.org/10.1175/JCLI3815.1>, 2006.
- Bento, V. A., Ribeiro, A. F. S., Russo, A., Gouveia, C. M., Cardoso, R. M., and Soares, P. M. M.: The impact of climate change

- in wheat and barley yields in the Iberian Peninsula, *Scientific Reports*, 11, 15484, <https://doi.org/10.1038/s41598-021-95014-6>, 2021.
- Bento, V. A., Russo, A., Gouveia, C. M., and DaCamara, C. C.: Recent change of burned area associated with summer heat extremes over Iberia, *Int. J. Wildland Fire*, 31, 658–669, 2022.
- Bento, V. A., Russo, A., Vieira, I., and Gouveia, C. M.: Identification of forest vulnerability to droughts in the Iberian Peninsula, *Theor. Appl. Climatol.*, 152, 559–579, <https://doi.org/10.1007/s00704-023-04427-y>, 2023.
- Bi, D., Dix, M., Marsland, S., O'Farrell, S., Rashid, H., Uotila, P., Hirst, A., Kowalczyk, E., Golebiewski, M., Sullivan, A., Yan, H., Hannah, N., Franklin, C., Sun, Z., Vohralik, P., Watterson, I., Zhou, X., Fiedler, R., Collier, M., Ma, Y., Noonan, J., Stevens, L., Uhe, P., Zhu, H., Griffies, S., Hill, R., Harris, C., and Puri, K.: The ACCESS coupled model: description, control climate and evaluation, *Aust. Meteor. and Ocean. J.*, 63, 41–64, 2013.
- Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., de Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharme, A., Dufresne, J.-L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M.-A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.-Y., Guenet, B., Guez, L., E., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J.-B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and Evaluation of the IPSL-CM6A-LR Climate Model, *J. Adv. Model. Earth Sy.*, 12, e2019MS002010, <https://doi.org/10.1029/2019MS002010>, 2020.
- Cannon, A. J.: Probabilistic Multisite Precipitation Downscaling by an Expanded Bernoulli Gamma Density Network, *J. Hydrometeorol.*, 9, 1284–1300, 2008.
- Cardoso, R. M., Soares, P. M. M., Lima, D. C. A., and Miranda, P. M. A.: Mean and extreme temperatures in a warming climate: EURO CORDEX and WRF regional climate high-resolution projections for Portugal, *Clim. Dynam.*, 52, 129–157, <https://doi.org/10.1007/s00382-018-4124-4>, 2019.
- Careto, J. A. M., Soares, P. M. M., Cardoso, R. M., Herrera, S., and Gutiérrez, J. M.: Added value of EURO-CORDEX high-resolution downscaling over the Iberian Peninsula revisited – Part 1: Precipitation, *Geosci. Model Dev.*, 15, 2635–2652, <https://doi.org/10.5194/gmd-15-2635-2022>, 2022a.
- Careto, J. A. M., Soares, P. M. M., Cardoso, R. M., Herrera, S., and Gutiérrez, J. M.: Added value of EURO-CORDEX high-resolution downscaling over the Iberian Peninsula revisited – Part 2: Max and min temperature, *Geosci. Model Dev.*, 15, 2653–2671, <https://doi.org/10.5194/gmd-15-2653-2022>, 2022b.
- Carter, B., Mueller, J., Jain, S., and Gifford, D. K.: What made you do this? Understanding black-box decisions with sufficient input subsets, *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics (AISTATS) 2019*, arXiv [preprint], <https://doi.org/10.48550/arXiv.1810.03805>, 2018.
- Chantry, M., Hatfield, S., Dueben, P., Polichtchouk, I., and Palmer, T.: Machine Learning Emulation of Gravity Wave Drag in Numerical Weather Forecasting, *J. Adv. Model. Earth Sy.*, 13, e2021MS002477, <https://doi.org/10.1029/2021MS002477>, 2021a.
- Chantry, M., Christensen, H., Dueben, P., and Palmer, T.: Opportunities and challenges for machine learning in weather and climate modelling: hard, medium and soft AI, *Philos. T. Roy. Soc. A*, 379, 20200083, <https://doi.org/10.1098/rsta.2020.0083>, 2021b.
- Christensen, J. H., Hewitson, B., Busuioc, A., Chen, A., Gao, X., Held, I., Jones, R., Kolli, R. K., Kwon, W.-T., Laprise, R., Magaña Rueda, V., Mearns, L., Menéndez, C. G., Räisänen, J., Rinke, A., Sarr, A., and Whetton, P.: Regional Climate Projections, in: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, <https://www.ipcc.ch/site/assets/uploads/2018/02/ar4-wg1-chapter11-1.pdf> (last access: January 2024), 2007.
- Cinquini, L., Crichton, D., Mattmann, C., Harney, J., Shipman, G., Wang, F., Ananthakrishnan, R., Miller, N., Denvil, S., Morgan, M., Pobre, Z., Bell, G. M., Doutriaux, C., Drach, R., Williams, D., Kershaw, P., Pascoe, S., Gonzalez, E., Fiore, S., and Schweitzer, R.: The Earth System Grid Federation: An open infrastructure for access to distributed geospatial data, *Future Gener. Comp. Sy.*, 36, 400–417, <https://doi.org/10.1016/j.future.2013.07.002>, 2014 (data available at: <https://esgf-node.llnl.gov/projects/esgf-llnl/>, last access: January 2024).
- Coppola, E., Sobolowski, S., Pichelli, E., Raffaele, F., Ahrens, B., Anders, I., Ban, N., Bastin, S., Belda, M., Belusic, D., Caldas-Alvarez, A., Cardoso, R. M., Davolio, S., Dobler, A., Fernandez, J., Fita, L., Fumiere, Q., Giorgi, F., Goergen, K., Güttler, I., Halenka, T., Heinzeller, D., Hodnebrog, Ø., Jacob, D., Kartsios, S., Katragkou, E., Kendon, E., Khodayar, S., Kunstmann, H., Knist, S., Lavín-Gullón, A., Lind, P., Lorenz, T., Maraun, D., Marelle, L., van Meijgaard, E., Milovac, J., Myhr, G., Panitz, H.-J., Piazza, M., Raffa, M., Raub, T., Rockel, B., Schär, C., Sieck, K., Soares, P. M. M., Somot, S., Srncic, L., Stocchi, P., Tölle, M. H., Truhetz, H., Vautard, R., de Vries, H., and Warrach-Sagi, K.: A first-of-its-kind multi-model convection permitting ensemble for investigating convective phenomena over Europe and the Mediterranean, *Clim. Dynam.*, 55, 3–34, <https://doi.org/10.1007/s00382-018-4521-8>, 2020.
- Cos, J., Doblas-Reyes, F., Jury, M., Marcos, R., Bretonnière, P.-A., and Samsó, M.: The Mediterranean climate change hotspot in the CMIP5 and CMIP6 projections, *Earth Syst. Dynam.*, 13, 321–340, <https://doi.org/10.5194/esd-13-321-2022>, 2022.
- Cramer, W., Guiot, J., Fader, M., Garrabou, J., Gattuso, J.-P., Iglesias, A., Lange, M. A., Lionello, P., Llasat, M. C., Paz, S., Peñuelas, J., Snoussi, M., Toreti, A., Tsimplis, M. N., and Xoplaki, E.: Climate change and interconnected risks to sustainable development in the Mediterranean, *Nat. Clim. Change*, 8, 972–980, <https://doi.org/10.1038/s41558-018-0299-2>, 2018.
- Dickinson, R. E., Errico, R. M., Giorgi, F., and Bates, G. T.: A regional climate model for the western United States, *Climatic*

- Change, 15, 383–422, <https://doi.org/10.1007/BF00240465>, 1989.
- Diffenbaugh, N. S. and Giorgi, F.: Climate change hotspots in the CMIP5 global climate model ensemble, *Climatic Change*, 114, 813–822, <https://doi.org/10.1007/s10584-012-0570-x>, 2012.
- Feser, F., Rockel, B., Storch, H. von, Winterfeldt, J., and Zahn, M.: Regional Climate Models Add Value to Global Model Data: A Review and Selected Examples, *B. Am. Meteor. Soc.*, 92, 1181–1192, <https://doi.org/10.1175/2011BAMS3061.1>, 2011.
- Fowler, H. J., Blenkinsop, S., and Tebaldi, C.: Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling, *Int. J. Climatol.*, 27, 1547–1578, <https://doi.org/10.1002/joc.1556>, 2007.
- Giorgi, F.: Climate change hot-spots, *Geophys. Res. Lett.*, 33, L08707, <https://doi.org/10.1029/2006GL025734>, 2006.
- Giorgi, F. and Bates, G. T.: The Climatological Skill of a Regional Model over Complex Terrain, *Mon. Weather Rev.*, 117, 2325–2347, [https://doi.org/10.1175/1520-0493\(1989\)117<2325:TCOAR>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<2325:TCOAR>2.0.CO;2), 1989.
- Giorgi, F. and Mearns, L. O.: Approaches to the simulation of regional climate change: A review, *Rev. Geophys.*, 29, 191–216, <https://doi.org/10.1029/90RG02636>, 1991.
- Giorgi, F. and Mearns, L. O.: Introduction to special section: Regional Climate Modeling Revisited, *J. Geophys. Res.-Atmos.*, 104, 6335–6352, <https://doi.org/10.1029/98JD02072>, 1999.
- Giorgi, F., Torma, C., Coppola, E., Ban, N., Schär, C., and Somot, S.: Enhanced summer convective rainfall at Alpine high elevations in response to climate warming, *Nat. Geosci.*, 9, 584–589, <https://doi.org/10.1038/ngeo2761>, 2016.
- Gutiérrez, J. M., Maraun, D., Widmann, M., Huth, R., Hertig, E., Benestad, R., Roessler, O., Wibig, J., Wilcke, R., Kotlarski, S., San Martín, D., Herrera, S., Bedia, J., Casanueva, A., Manzanas, R., Iturbide, M., Vrac, M., Dubrovsky, M., Ribalaygua, J., Pórtoles, J., Rätty, O., Räisänen, J., Hingray, B., Raynaud, D., Casado, M. J., Ramos, P., Zerener, T., Turco, M., Bosshard, T., Štěpánek, P., Bartholy, J., Pongracz, R., Keller, D. E., Fischer, A. M., Cardoso, R. M., Soares, P. M. M., Czernecki, B., and Pagé, C.: An intercomparison of a large ensemble of statistical downscaling methods over Europe: Results from the VALUE perfect predictor cross-validation experiment, *Int. J. Climatol.*, 39, 3750–3785, <https://doi.org/10.1002/joc.5462>, 2019a.
- Gutiérrez, J. M., Herrera, S., Cardoso, R. M., Matos Soares, P., Espírito-Santo, F., and Viterbo, P.: Iberia01: Daily gridded (0.1° resolution) dataset of precipitation and temperatures over the Iberian Peninsula, DIGITAL.CSIC [data set], <https://doi.org/10.20350/digitalCSIC/8641>, 2019b.
- Gutjahr, O., Putrasahan, D., Lohmann, K., Jungclaus, J. H., von Storch, J.-S., Brüggemann, N., Haak, H., and Stössel, A.: Max Planck Institute Earth System Model (MPI-ESM1.2) for the High-Resolution Model Intercomparison Project (HighResMIP), *Geosci. Model Dev.*, 12, 3241–3281, <https://doi.org/10.5194/gmd-12-3241-2019>, 2019.
- Harvey, B. J., Shaffrey, L. C., and Woollings, T. J.: Equator-to-pole temperature differences and the extra-tropical storm track responses of the CMIP5 climate models, *Clim. Dynam.*, 43, 1171–1182, <https://doi.org/10.1007/s00382-013-1883-9>, 2014.
- Haupt, S. E., Chapman, W., Adams, S. V., Kirkwood, C., Hosking, J. S., Robinson, N. H., Lerch, S., and Subramanian, A. C.: Towards implementing artificial intelligence post-processing in weather and climate: proposed actions from the Oxford 2019 workshop, *Philos. T. Roy. Soc. A*, 379, 20200091, <https://doi.org/10.1098/rsta.2020.0091>, 2021.
- Hernanz, A., García-Valero, J. A., Domínguez, M., and Rodríguez-Camino, E.: A critical view on the suitability of machine learning techniques to downscale climate change projections: Illustration for temperature with a toy experiment, *Atmos. Sci. Lett.*, 23, e1087, <https://doi.org/10.1002/asl.1087>, 2022.
- Herrera, S., Cardoso, R. M., Soares, P. M., Espírito-Santo, F., Viterbo, P., and Gutiérrez, J. M.: Iberia01: a new gridded dataset of daily precipitation and temperatures over Iberia, *Earth Syst. Sci. Data*, 11, 1947–1956, <https://doi.org/10.5194/essd-11-1947-2019>, 2019.
- Herrera, S., Soares, P. M. M., Cardoso, R. M., and Gutiérrez, J. M.: Evaluation of the EURO-CORDEX Regional Climate Models Over the Iberian Peninsula: Observational Uncertainty Analysis, *J. Geophys. Res.-Atmos.*, 125, e2020JD032880, <https://doi.org/10.1029/2020JD032880>, 2020.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, *Q. J. Roy. Meteor. Soc.*, 146, 1999–2049, <https://doi.org/10.1002/qj.3803>, 2020.
- Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 hourly data on single levels from 1940 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set], <https://doi.org/10.24381/cds.adbb2d47>, 2023.
- Hertig, E., Maraun, D., Bartholy, J., Pongracz, R., Vrac, M., Mares, I., Gutiérrez, J. M., Wibig, J., Casanueva, A., and Soares, P. M. M.: Comparison of statistical downscaling methods with respect to extreme events over Europe: Validation results from the perfect predictor experiment of the COST Action VALUE, *Int. J. Climatol.*, 39, 3846–3867, <https://doi.org/10.1002/joc.5469>, 2019.
- Hoerling, M., Eischeid, J., Perlwitz, J., Quan, X., Zhang, T., and Peignon, P.: On the Increased Frequency of Mediterranean Drought, *J. Climate*, 25, 2146–2161, <https://doi.org/10.1175/JCLI-D-11-00296.1>, 2012.
- IPCC: Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, <https://www.ipcc.ch/report/ar6/wg1/> (last access: January 2024), 2021.
- Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., and Yiou, P.: EURO-

- CORDEX: new high-resolution climate change projections for European impact research, *Reg. Environ. Change*, 14, 563–578, <https://doi.org/10.1007/s10113-013-0499-2>, 2014.
- Jacob, D., Teichmann, C., Sobolowski, S., Katragkou, E., Anders, I., Belda, M., Benestad, R., Boberg, F., Buonomo, E., Cardoso, R. M., Casanueva, A., Christensen, O. B., Christensen, J. H., Coppola, E., De Cruz, L., Davin, E. L., Dobler, A., Domínguez, M., Fealy, R., Fernandez, J., Gaertner, M. A., García-Díez, M., Giorgi, F., Gobiet, A., Goergen, K., Gómez-Navarro, J. J., Alemán, J. J. G., Gutiérrez, C., Gutiérrez, J. M., Güttler, I., Haensler, A., Halenka, T., Jerez, S., Jiménez-Guerrero, P., Jones, R. G., Keuler, K., Kjellström, E., Knist, S., Kotlarski, S., Maraun, D., van Meijgaard, E., Mercogliano, P., Montávez, J. P., Navarra, A., Nikulin, G., de Noblet-Ducoudré, N., Panitz, H.-J., Pfeifer, S., Piazza, M., Pichelli, E., Pietikäinen, J.-P., Prein, A. F., Preuschmann, S., Rechid, D., Rockel, B., Romera, R., Sánchez, E., Sieck, K., Soares, P. M. M., Somot, S., Srncic, L., Sørland, S. L., Termonia, P., Truhetz, H., Vautard, R., Warrach-Sagi, K., and Wulfmeyer, V.: Regional climate downscaling over Europe: perspectives from the EURO-CORDEX community, *Reg. Environ. Change*, 20, 51, <https://doi.org/10.1007/s10113-020-01606-9>, 2020.
- Kang, S. M. and Lu, J.: Expansion of the Hadley Cell under Global Warming: Winter versus Summer, *J. Climate*, 25, 8387–8393, <https://doi.org/10.1175/JCLI-D-12-00323.1>, 2012.
- Laprise, R., de Elía, R., Caya, D., Biner, S., Lucas-Picher, P., Diaconescu, E., Leduc, M., Alexandru, A., Separovic, L., and Canadian Network for Regional Climate Modelling and Diagnostics: Challenging some tenets of Regional Climate Modelling, *Meteorology and Atmospheric Physics*, 100, 3–22, <https://doi.org/10.1007/s00703-008-0292-9>, 2008.
- Lecun, Y. and Bengio, Y.: Convolutional Networks for Images, Speech and Time Series, in: *The Handbook of Brain Theory and Neural Networks*, The MIT Press, 255–258, 1995.
- Le Roux, R., Katurji, M., Zawar-Reza, P., Quénot, H., and Sturman, A.: Comparison of statistical and dynamical downscaling results from the WRF model, *Environ. Modell. Softw.*, 100, 67–73, <https://doi.org/10.1016/j.envsoft.2017.11.002>, 2018.
- Lima, D. C. A., Lemos, G., Bento, V. A., Nogueira, M., and Soares, P. M. M.: A multi-variable constrained ensemble of regional climate projections under multi-scenarios for Portugal – Part I: An overview of impacts on means and extremes, *Climate Services*, 30, 100351, <https://doi.org/10.1016/j.cliser.2023.100351>, 2023a.
- Lima, D. C. A., Bento, V. A., Lemos, G., Nogueira, M., and Soares, P. M. M.: A multi-variable constrained ensemble of regional climate projections under multi-scenarios for Portugal – Part II: Sectoral climate indices, *Climate Services*, 30, 100377, <https://doi.org/10.1016/j.cliser.2023.100377>, 2023b.
- Lionello, P. and Scarascia, L.: The relation between climate change in the Mediterranean region and global warming, *Reg. Environ. Change*, 18, 1481–1493, <https://doi.org/10.1007/s10113-018-1290-1>, 2018.
- Lucas-Picher, P., Laprise, R., and Winger, K.: Evidence of added value in North American regional climate model hindcast simulations using ever-increasing horizontal resolutions, *Clim. Dynam.*, 48, 2611–2633, <https://doi.org/10.1007/s00382-016-3227-z>, 2017.
- Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Theißel, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, *Rev. Geophys.*, 48, RG3003, <https://doi.org/10.1029/2009RG000314>, 2010.
- Maraun, D., Shepherd, T. G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J. M., Hagemann, S., Richter, I., Soares, P. M. M., Hall, A., and Mearns, L. O.: Towards process-informed bias correction of climate change simulations, *Nat. Clim. Change*, 7, 764–773, <https://doi.org/10.1038/nclimate3418>, 2017.
- Maraun, D., Widmann, M., and Gutiérrez, J. M.: Statistical downscaling skill under present climate conditions: A synthesis of the VALUE perfect predictor experiment, *Int. J. Climatol.*, 39, 3692–3703, <https://doi.org/10.1002/joc.5877>, 2019.
- McGregor, J. L.: Regional climate modelling, *Meteorol. Atmos. Phys.*, 63, 105–117, <https://doi.org/10.1007/BF01025367>, 1997.
- Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bitner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., Ilyina, T., Kleine, T., Kornblüeh, L., Li, H., Modali, K., Notz, D., Pohlmann, H., Roeckner, E., Stemmler, I., Tian, F., and Marotzke, J.: A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR), *J. Adv. Model. Earth Sy.*, 10, 1383–1413, <https://doi.org/10.1029/2017MS001217>, 2018.
- Nikulin, G., Lennard, C., Dosio, A., Kjellström, E., Chen, Y., Hänsler, A., Kupiainen, M., Laprise, R., Mariotti, L., Maule, C. F., van Meijgaard, E., Panitz, H.-J., Scinocca, J. F., and Somot, S.: The effects of 1.5 and 2 degrees of global warming on Africa in the CORDEX ensemble, *Environ. Res. Lett.*, 13, 065003, <https://doi.org/10.1088/1748-9326/aab1b1>, 2018.
- Palma, J. H. N., Paulo, J. A., Faias, S. P., Garcia-Gonzalo, J., Borges, J. G., and Tomé, M.: Adaptive management and debarking schedule optimization of *Quercus suber* L. stands under climate change: case study in Chamusca, Portugal, *Reg. Environ. Change*, 15, 1569–1580, <https://doi.org/10.1007/s10113-015-0818-x>, 2015.
- Palma, J. H. N., Cardoso, R. M., Soares, P. M. M., Oliveira, T. S., and Tomé, M.: Using high-resolution simulated climate projections in forest process-based modelling, *Agr. Forest Meteorol.*, 263, 100–106, <https://doi.org/10.1016/j.agrformet.2018.08.008>, 2018.
- Páscoa, P., Russo, A., Gouveia, C. M., Soares, P. M. M., Cardoso, R. M., Careto, J. A. M., and Ribeiro, A. F. S.: A high-resolution view of the recent drought trends over the Iberian Peninsula, *Weather and Climate Extremes*, 32, 100320, <https://doi.org/10.1016/j.wace.2021.100320>, 2021.
- Pereira, C. and Coelho, C.: Mapping erosion risk under different scenarios of climate change for Aveiro coast, Portugal, *Nat. Hazards*, 69, 1033–1050, <https://doi.org/10.1007/s11069-013-0748-1>, 2013.
- Perkins, S. E., Pitman, A. J., Holbrook, N. J., and McAneney, J.: Evaluation of the AR4 Climate Models' Simulated Daily Maximum Temperature, Minimum Temperature, and Precipitation over Australia Using Probability Density Functions, *J. Climate*, 20, 4356–4376, <https://doi.org/10.1175/JCLI4253.1>, 2007.
- Pichelli, E., Coppola, E., Sobolowski, S., Ban, N., Giorgi, F., Stocchi, P., Alias, A., Belušić, D., Berthou, S., Caillaud, C., Cardoso, R. M., Chan, S., Christensen, O. B., Dobler, A., de Vries, H.,

- Goergen, K., Kendon, E. J., Keuler, K., Lenderink, G., Lorenz, T., Mishra, A. N., Panitz, H.-J., Schär, C., Soares, P. M. M., Truhetz, H., and Vergara-Temprado, J.: The first multi-model ensemble of regional climate simulations at kilometer-scale resolution part 2: historical and future simulations of precipitation, *Clim. Dynam.*, 56, 3581–3602, <https://doi.org/10.1007/s00382-021-05657-4>, 2021.
- Planton, S., Lionello, P., Artale, V., Aznar, R., Carrillo, A., Colin, J., Congedi, L., Dubois, C., Elizalde, A., Gualdi, S., Hertig, E., Jacobeit, J., Jordà, G., Li, L., Mariotti, A., Piani, C., Ruti, P., Sanchez-Gomez, E., Sannino, G., Sevault, F., Somot, S., and Tsimplis, M.: 8 – The Climate of the Mediterranean Region in Future Climate Projections, in: *The Climate of the Mediterranean Region*, edited by: Lionello, P., Elsevier, Oxford, 449–502, <https://doi.org/10.1016/B978-0-12-416042-2.00008-2>, 2012.
- Prein, A. F., Gobiet, A., Suklitsch, M., Truhetz, H., Awan, N. K., Keuler, K., and Georgievski, G.: Added value of convection permitting seasonal simulations, *Clim. Dynam.*, 41, 2655–2677, <https://doi.org/10.1007/s00382-013-1744-6>, 2013.
- Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F., Brisson, E., Kollet, S., Schmidli, J., van Lipzig, N. P. M., and Leung, R.: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges, *Rev. Geophys.*, 53, 323–361, <https://doi.org/10.1002/2014RG000475>, 2015.
- 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.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, *Geosci. Model Dev.*, 9, 3461–3482, <https://doi.org/10.5194/gmd-9-3461-2016>, 2016.
- Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R. J., Sumi, A., and Taylor, K. E.: Climate models and their evaluation, in: *IPCC, 2007: Climate Change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., Cambridge University Press, Cambridge, UK, <http://hdl.handle.net/102.100.100/124621?index=1> (last access: November 2023), 2007.
- Rasp, S. and Lerch, S.: Neural Networks for Postprocessing Ensemble Weather Forecasts, *Mon. Weather Rev.*, 146, 3885–3900, <https://doi.org/10.1175/MWR-D-18-0187.1>, 2018.
- Rasp, S., Dueben, P. D., Scher, S., Weyn, J. A., Mouatadid, S., and Thuerey, N.: WeatherBench: A Benchmark Data Set for Data-Driven Weather Forecasting, *J. Adv. Model. Earth Sy.*, 12, e2020MS002203, <https://doi.org/10.1029/2020MS002203>, 2020.
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, *Nature*, 566, 195–204, <https://doi.org/10.1038/s41586-019-0912-1>, 2019.
- Rios-Entenza, A., Soares, P. M. M., Trigo, R. M., Cardoso, R. M., and Miguez-Macho, G.: Moisture recycling in the Iberian Peninsula from a regional climate simulation: Spatiotemporal analysis and impact on the precipitation regime, *J. Geophys. Res.-Atmos.*, 119, 5895–5912, <https://doi.org/10.1002/2013JD021274>, 2014.
- Rössler, O., Fischer, A. M., Huebener, H., Maraun, D., Benestad, R. E., Christodoulides, P., Soares, P. M. M., Cardoso, R. M., Pagé, C., Kanamaru, H., Kreienkamp, F., and Vlachogiannis, D.: Challenges to link climate change data provision and user needs: Perspective from the COST-action VALUE, *Int. J. Climatol.*, 39, 3704–3716, <https://doi.org/10.1002/joc.5060>, 2019.
- Rozenberg, J., Guivarch, C., Lempert, R., and Hallegatte, S.: Building SSP for climate policy analysis: a scenario elicitation methodology to map the space of possible future challenges to mitigation and adaptation, *Climatic Change*, 122, 509–522, <https://doi.org/10.1007/s10584-013-0904-3>, 2014.
- Rummukainen, M.: State-of-the-art with regional climate models, *WIREs Climate Change*, 1, 82–96, <https://doi.org/10.1002/wcc.8>, 2010.
- Rummukainen, M.: Added value in regional climate modeling, *WIREs Climate Change*, 7, 145–159, <https://doi.org/10.1002/wcc.378>, 2016.
- Russo, A., Gouveia, C. M., Dutra, E., Soares, P. M. M., and Trigo, R. M.: The synergy between drought and extremely hot summers in the Mediterranean, *Environ. Res. Lett.*, 14, 014011, <https://doi.org/10.1088/1748-9326/aaf09e>, 2019.
- Schmidhuber, J.: Deep learning in neural networks: An overview, *Neural Networks*, 61, 85–117, <https://doi.org/10.1016/j.neunet.2014.09.003>, 2015.
- Schultz, M. G., Betancourt, C., Gong, B., Kleinert, F., Langguth, M., Leufen, L. H., Mozaffari, A., and Stadler, S.: Can deep learning beat numerical weather prediction?, *Philos. T. Roy. Soc. A*, 379, 20200097, <https://doi.org/10.1098/rsta.2020.0097>, 2021.
- Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voltaire, A., Colin, J., Decharme, B., Delire, C., Berthet, S., Chevallier, M., Sénési, S., Franchisteguy, L., Vial, J., Mallet, M., Joetzjer, E., Geoffroy, O., Guérémy, J.-F., Moine, M.-P., Msadek, R., Ribes, A., Rocher, M., Roehrig, R., Salas-y-Méllia, D., Sanchez, E., Terray, L., Valcke, S., Waldman, R., Aumont, O., Bopp, L., Deshayes, J., Éthé, C., and Madec, G.: Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate, *J. Adv. Model. Earth Sy.*, 11, 4182–4227, <https://doi.org/10.1029/2019MS001791>, 2019.
- Seland, Ø., Bentsen, M., Olivie, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y.-C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations, *Geosci. Model Dev.*, 13, 6165–6200, <https://doi.org/10.5194/gmd-13-6165-2020>, 2020.
- Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., de Mora, L., Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews, T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N., Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O.,

- Parker, R. J., Predoi, V., Robertson, E., Siahhan, A., Smith, R. S., Swaminathan, R., Woodhouse, M. T., Zeng, G., and Zerroukat, M.: UKESM1: Description and Evaluation of the U.K. Earth System Model, *J. Adv. Model. Earth Sy.*, 11, 4513–4558, <https://doi.org/10.1029/2019MS001739>, 2019.
- Soares, P. M. M. and Lima, D. C. A.: Water scarcity down to earth surface in a Mediterranean climate: The extreme future of soil moisture in Portugal, *J. Hydrol.*, 615, 128731, <https://doi.org/10.1016/j.jhydrol.2022.128731>, 2022.
- Soares, P. M. M., Cardoso, R. M., Miranda, P. M. A., de Medeiros, J., Belo-Pereira, M., and Espirito-Santo, F.: WRF high resolution dynamical downscaling of ERA-Interim for Portugal, *Clim. Dynam.*, 39, 2497–2522, <https://doi.org/10.1007/s00382-012-1315-2>, 2012.
- Soares, P. M. M., Cardoso, R. M., Lima, D. C. A., and Miranda, P. M. A.: Future precipitation in Portugal: high-resolution projections using WRF model and EURO-CORDEX multi-model ensembles, *Clim. Dynam.*, 49, 2503–2530, <https://doi.org/10.1007/s00382-016-3455-2>, 2017a.
- Soares, P. M. M., Lima, D. C. A., Cardoso, R. M., and Semedo, A.: High resolution projections for the western Iberian coastal low level jet in a changing climate, *Clim. Dynam.*, 49, 1547–1566, <https://doi.org/10.1007/s00382-016-3397-8>, 2017b.
- Soares, P. M. M., Maraun, D., Brands, S., Jury, M. W., Gutiérrez, J. M., San-Martín, D., Hertig, E., Huth, R., Belušić, Vozila, A., Cardoso, R. M., Kotlarski, S., Drobinski, P., and Obermann-Hellhund, A.: Process-based evaluation of the VALUE perfect predictor experiment of statistical downscaling methods, *Int. J. Climatol.*, 39, 3868–3893, <https://doi.org/10.1002/joc.5911>, 2019.
- Soares, P. M. M., Careto, J. A. M., Cardoso, R. M., Goergen, K., Katragkou, E., Sobolowski, S., Coppola, E., Ban, N., Belušić, D., Berthou, S., Caillaud, C., Dobler, A., Hodnebrog, Ø., Kartios, S., Lenderink, G., Lorenz, T., Milovac, J., Feldmann, H., Pichelli, E., Truhetz, H., Demory, M. E., de Vries, H., Warrach-Sagi, K., Keuler, K., Raffa, M., Tölle, M., Sieck, K., and Bastin, S.: The added value of km-scale simulations to describe temperature over complex orography: the CORDEX FPS-Convection multi-model ensemble runs over the Alps, *Clim. Dynam.*, <https://doi.org/10.1007/s00382-022-06593-7>, 2022.
- Soares, P. M. M., Careto, J. A. M., Russo, A., and Lima, D. C. A.: The future of Iberian droughts: a deeper analysis based on multi-scenario and a multi-model ensemble approach, *Nat. Hazards*, 117, 2001–2028, <https://doi.org/10.1007/s11069-023-05938-7>, 2023a.
- Soares, P. M., Johannsen, F., Lima, D. C., Lemos, G., Bento, V., and Bushenkova, A.: High resolution downscaling of CMIP6 Earth System and Global Climate Models using deep learning for Iberia Part 1 (Version 1), Zenodo [data set], <https://doi.org/10.5281/zenodo.8314980>, 2023b.
- Soares, P. M., Johannsen, F., Lima, D. C., Lemos, G., Bento, V., and Bushenkova, A.: High resolution downscaling of CMIP6 Earth System and Global Climate Models using deep learning for Iberia Part 2 (Version 1), Zenodo [data set], <https://doi.org/10.5281/zenodo.8338468>, 2023c.
- Soares, P. M., Johannsen, F., Lima, D. C., Lemos, G., Bento, V., and Bushenkova, A.: High resolution downscaling of CMIP6 Earth System and Global Climate Models using deep learning for Iberia Part 3 (Version 1), Zenodo [data set], <https://doi.org/10.5281/zenodo.8340234>, 2023d.
- Soares, P. M., Johannsen, F., Lima, D. C., Lemos, G., Bento, V., and Bushenkova, A.: High resolution downscaling of CMIP6 Earth System and Global Climate Models using deep learning for Iberia Part 4 (Version 1), Zenodo [data set], <https://doi.org/10.5281/zenodo.8340250>, 2023e.
- Soares, P. M., Johannsen, F., Lima, D. C., Lemos, G., Bento, V., and Bushenkova, A.: High resolution downscaling of CMIP6 Earth System and Global Climate Models using deep learning for Iberia Part 5 (Version 1), Zenodo [data set], <https://doi.org/10.5281/zenodo.8340266>, 2023f.
- Soares, P. M., Johannsen, F., Lima, D. C., Lemos, G., Bento, V., and Bushenkova, A.: High resolution downscaling of CMIP6 Earth System and Global Climate Models using deep learning for Iberia Part 6 (Version 1), Zenodo [data set], <https://doi.org/10.5281/zenodo.8340274>, 2023g.
- Soares, P. M., Johannsen, F., Lima, D. C., Lemos, G., Bento, V., and Bushenkova, A.: High resolution downscaling of CMIP6 Earth System and Global Climate Models using deep learning for Iberia Part 7 (Version 1), Zenodo [data set], <https://doi.org/10.5281/zenodo.8340279>, 2023h.
- Soares, P. M., Johannsen, F., Lima, D. C., Lemos, G., Bento, V., and Bushenkova, A.: High resolution downscaling of CMIP6 Earth System and Global Climate Models using deep learning for Iberia Part 8 (Version 1), Zenodo [data set], <https://doi.org/10.5281/zenodo.8340287>, 2023i.
- Soares, P. M., Johannsen, F., Lima, D. C., Lemos, G., Bento, V., and Bushenkova, A.: High resolution downscaling of CMIP6 Earth System and Global Climate Models using deep learning for Iberia Part 9 (Version 1), Zenodo [data set], <https://doi.org/10.5281/zenodo.8340297>, 2023j.
- Soares, P. M., Johannsen, F., Lima, D. C., Lemos, G., Bento, V., and Bushenkova, A.: High resolution downscaling of CMIP6 Earth System and Global Climate Models using deep learning for Iberia Part 10 (Version 1), Zenodo [data set], <https://doi.org/10.5281/zenodo.8340318>, 2023k.
- Soares, P. M., Johannsen, F., Lima, D. C., Lemos, G., Bento, V., and Bushenkova, A.: High resolution downscaling of CMIP6 Earth System and Global Climate Models using deep learning for Iberia Part 11 (Version 1), Zenodo [data set], <https://doi.org/10.5281/zenodo.8340338>, 2023l.
- Sobolowski, S., Somot, S., Fernandez, J., Evin, G., Maraun, D., Kotlarski, S., Jury, M., Benestad, R. E., Teichmann, C., Christensen, O. B., Katharina, B., Buonomo, E., Katragkou, E., Steger, C., Sørland, S., Nikulin, G., McSweeney, C., Dobler, A., Palmer, T., Wilke, R., Boé, J., Brunner, L., Ribes, A., Qasmi, S., Nabat, P., Sevault, F., Oudar, T., and Brands, S. S.: EURO CORDEX CMIP6 GCM Selection & Ensemble Design: Best Practices and Recommendations, Zenodo, <https://doi.org/10.5281/zenodo.7673400>, 2023.
- Tamarin, T. and Kaspi, Y.: The poleward shift of storm tracks under global warming: A Lagrangian perspective, *Geophys. Res. Lett.*, 44, 10666–10674, <https://doi.org/10.1002/2017GL073633>, 2017.
- Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., Sudo, K., Sekiguchi, M., Abe, M., Saito, F., Chikira, M., Watanabe, S., Mori, M., Hirota, N., Kawatani, Y., Mochizuki, T., Yoshimura, K., Takata, K., O'ishi, R., Yamazaki, D., Suzuki,

- T., Kurogi, M., Kataoka, T., Watanabe, M., and Kimoto, M.: Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6, *Geosci. Model Dev.*, 12, 2727–2765, <https://doi.org/10.5194/gmd-12-2727-2019>, 2019.
- Trigo, R. M. and Palutikof, J. P.: Simulation of daily temperatures for climate change scenarios over Portugal, *Clim. Res.*, 13, 45–59, 1999.
- Tuel, A. and Eltahir, E. A. B.: Why Is the Mediterranean a Climate Change Hot Spot?, *J. Climate*, 33, 5829–5843, <https://doi.org/10.1175/JCLI-D-19-0910.1>, 2020.
- Turco, M., Palazzi, E., von Hardenberg, J., and Provenzale, A.: Observed climate change hotspots, *Geophys. Res. Lett.*, 42, 3521–3528, <https://doi.org/10.1002/2015GL063891>, 2015.
- Ulbrich, U., Pinto, J. G., Kupfer, H., Leckebusch, G. C., Spanghel, T., and Reyers, M.: Changing Northern Hemisphere Storm Tracks in an Ensemble of IPCC Climate Change Simulations, *J. Climate*, 21, 1669–1679, <https://doi.org/10.1175/2007JCLI1992.1>, 2008.
- Vrac, M. and Ayar, P. V.: Influence of Bias Correcting Predictors on Statistical Downscaling Models, *J. Appl. Meteorol. Clim.*, 56, 5–26, <https://doi.org/10.1175/JAMC-D-16-0079.1>, 2017.
- Widmann, M., Bedia, J., Gutiérrez, J. M., Bosshard, T., Hertig, E., Maraun, D., Casado, M. J., Ramos, P., Cardoso, R. M., Soares, P. M. M., Ribalaygua, J., Pagé, C., Fischer, A. M., Herrera, S., and Huth, R.: Validation of spatial variability in downscaling results from the VALUE perfect predictor experiment, *Int. J. Climatol.*, 39, 3819–3845, <https://doi.org/10.1002/joc.6024>, 2019.
- Wilby, R. L. and Wigley, T. M. L.: Downscaling general circulation model output: a review of methods and limitations, *Prog. Phys. Geog.*, 21, 530–548, <https://doi.org/10.1177/030913339702100403>, 1997.
- Wilby, R. L., Wigley, T. M. L., Conway, D., Jones, P. D., Hewitson, B. C., Main, J., and Wilks, D. S.: Statistical downscaling of general circulation model output: A comparison of methods, *Water Resour. Res.*, 34, 2995–3008, <https://doi.org/10.1029/98WR02577>, 1998.