Climate Services 32 (2023) 100408

Contents lists available at ScienceDirect

Climate Services

journal homepage: www.elsevier.com/locate/cliser

pyClim-SDM: Service for generation of statistically downscaled climate change projections supporting national adaptation strategies

Alfonso Hernanz [a,*], Carlos Correa [a], Juan Andrés García-Valero [b], Marta Domínguez [a], Esteban Rodríguez-Guisado [a], Ernesto Rodríguez-Camino [a]

[a] Spanish Meteorological Agency (AEMET), Madrid 28040, Spain
[b] AEMET, Murcia 30107, Spain

ARTICLE INFO

Keywords:
Statistical downscaling
Climate service
Climate projections
Graphical user interface
Software

ABSTRACT

The climate change impact and adaptation communities need future scenarios with sufficient high resolution, which are frequently achieved by applying Statistical Downscaling Models (SDMs) over global climate models. A large variety of SDMs exists, and some can be more suitable than others for each specific purpose. For this reason, it is important to develop tools to facilitate the evaluation and generation of downscaled scenarios following different approaches. In this paper we present a service, 'pyClim-SDM', which allows users to generate and evaluate their own downscaled scenarios with a very simple and user-friendly graphical interface. This tool includes a large collection of *state-of-the-art* methods belonging to different families to downscale daily data of the following surface variables: temperature, precipitation, wind, relative humidity and cloud coverage. Additionally, the software is prepared to be applied over any other user-defined target variable. Thus, multivariable indexes can be tackled as target variables themselves, instead of being calculated from the downscaled primary variables. With this possibility, potential intervariable inconsistencies are avoided. An application example for a Fire Weather Index, dependent on temperature, wind, humidity and precipitation, is shown. The service here presented -mainly based on a new downscaling software and a user-friendly graphical interface- is an essential piece for evaluating and generating high-resolution projection data within the Spanish national climate change adaptation strategy which includes, among other elements, a common database for all sectors, viewer and data distribution portal, etc.

Practical implications

There is an increasing demand for high-resolution climate projections for impact and adaptation studies. The Paris Agreement reached in 2015 by the United Nations Framework Convention on Climate Change (UNFCCC, 2015) establishes that Parties should undertake pertinent adaptation measures. Here we present a service to generate high-resolution climate change projections that has the following features:

- It is one of the elements feeding the Spanish National Plan of Adaptation to Climate Change.
- It is being applied also for adaptation purposes in the Central American region.
- It can be easily used by non-experts on climate data.
- It is also prepared for more academic evaluation and comparison exercises.

This tool has been developed and applied to feed the Spanish National Plan of Adaptation to Climate Change, and we consider it might be useful for other countries to accomplish their compromises on adaptation.

Data availability

Data used for this study is freely available (see the Data and Code availability section)

* Corresponding author.
E-mail address: ahernanzl@aemet.es (A. Hernanz).

https://doi.org/10.1016/j.cliser.2023.100408
Received 4 August 2022; Received in revised form 11 July 2023; Accepted 19 September 2023
Available online 24 September 2023

Fig. 1. pyClim-SDM screenshot.

Table 1
Global Climate Models used for the test example.

Model	Institution	References
ACCESS-CM2	Commonwealth Scientific and Industrial Research Org. (CSIRO) and Bureau of Meteorology (BoM), Australia	Bi *et al.* (2020)
CanESM5	Canadian Centre for Climate Modelling and Analysis, Canada	Swart et al. (2019)
EC-Earth3	EC-Earth consortium, Europe	Döscher et al. (2021)
INM-CM5-0	Institute of Numerical Mathematics, Russia	Volodin et al. (2017)
MIROC6	Research Center for Environmental Modeling and Application, Japan	Tatebe *et al.* (2019)
MPI-ESM1-2-HR	Max-Planck-Institut (MPI) for Meteorology, Germany	Müller et al. (2018)
MRI-ESM2-0	Meteorological Research Institute, Tsukuba, Japan	Yukimoto et al. (2019)

Table 2
Predictors used for the test example. tas (mean surface temperature), tasmax (maximum surface temperature), tasmin (minimum surface temperature), pr (precipitation), uas (surface zonal wind component), vas (surface meridional wind component), sfcWind (surface wind speed), hurs (surface relative humidity), clt (cloud coverage), FWI (Fire Weather Index), psl (mean sea level pressure), K_index (stability 'K index'), TT_index (stability 'Total of totals index'). Upper air predictors (at 850 and 500 hPa): ua (zonal wind component), va (meridional wind component), ta (temperature) and hur (relative humidity).

Target variable	Predictors
tas, tasmax, tasmin	tas, ua850, ua500, va850, va500, ta850, ta500
pr	psl, ua850, ua500, va850, va500, hur850, hur500, K_index, TT_index
uas, vas, sfcWind	uas, vas, ua850, ua500, va850, va500, ta850, ta500
hurs	tas, hurs, ta850, ta500, hur850, hur500
clt	clt
FWI	psl, uas, vas, tas, hurs, ua850, ua500, va850, va500, ta850, ta500, hurs850, hurs500, K_index, TT_index

1. Introduction

Global Climate Models (GCMs) are the primary tool to simulate future climate projections, but they have known biases and their resolution is not enough to meet the necessities of the impact and adaptation communities (Charles et al., 2004; Wilby, 2004; Schoof, 2013). Two primary categories of downscaling techniques exist: (1) dynamic downscaling, mostly by nesting a high-resolution Regional Climate Model (RCM) within a GCM and (2) statistical downscaling (SD), based on the existence of statistical relationships between large-scale variables (predictors) and local variables (predictands). Some major advantages of Statistical Downscaling Models (SDMs) are their computational cheapness compared with dynamic downscaling (Trzaska and Schnarr, 2014) and their capability of downscaling to single point scale.

The Paris Agreement reached in 2015 by the United Nations Framework Convention on Climate Change (UNFCCC, 2015), in its Article 7, establishes that Parties should undertake pertinent adaptation measures. For this task, regional information on climate change is needed, and there is an increasing effort to provide high-resolution climate projections around the world. The Coordinated Regional Climate Downscaling Experiment (CORDEX) is a frequent source of high resolution climate projections supporting regional and local climate impact studies and adaptation decisions. Nevertheless, their downscaled projections still lack the needed resolution for some applications over many regions. Thus, Parties often need to additionally generate their own downscaled climate projections, and this is usually done using statistical methods.

There is a huge variety of SDMs belonging to different families and relying on different assumptions. SDMs are based on the assumption of stationarity in the statistical relationships between predictors and predictands. SDMs can be categorized depending on their calibration strategy as Perfect Prognosis (PP) and Model Output Statistics (MOS). The PP approach relies on the assumption that model predictors are unbiased, so these methods are calibrated using observations (reanalysis) and then they are applied to GCMs. On the other hand, the MOS approach assumes imperfections from models, and SDMs are calibrated making use of the GCMs themselves, so their biases are incorporated and

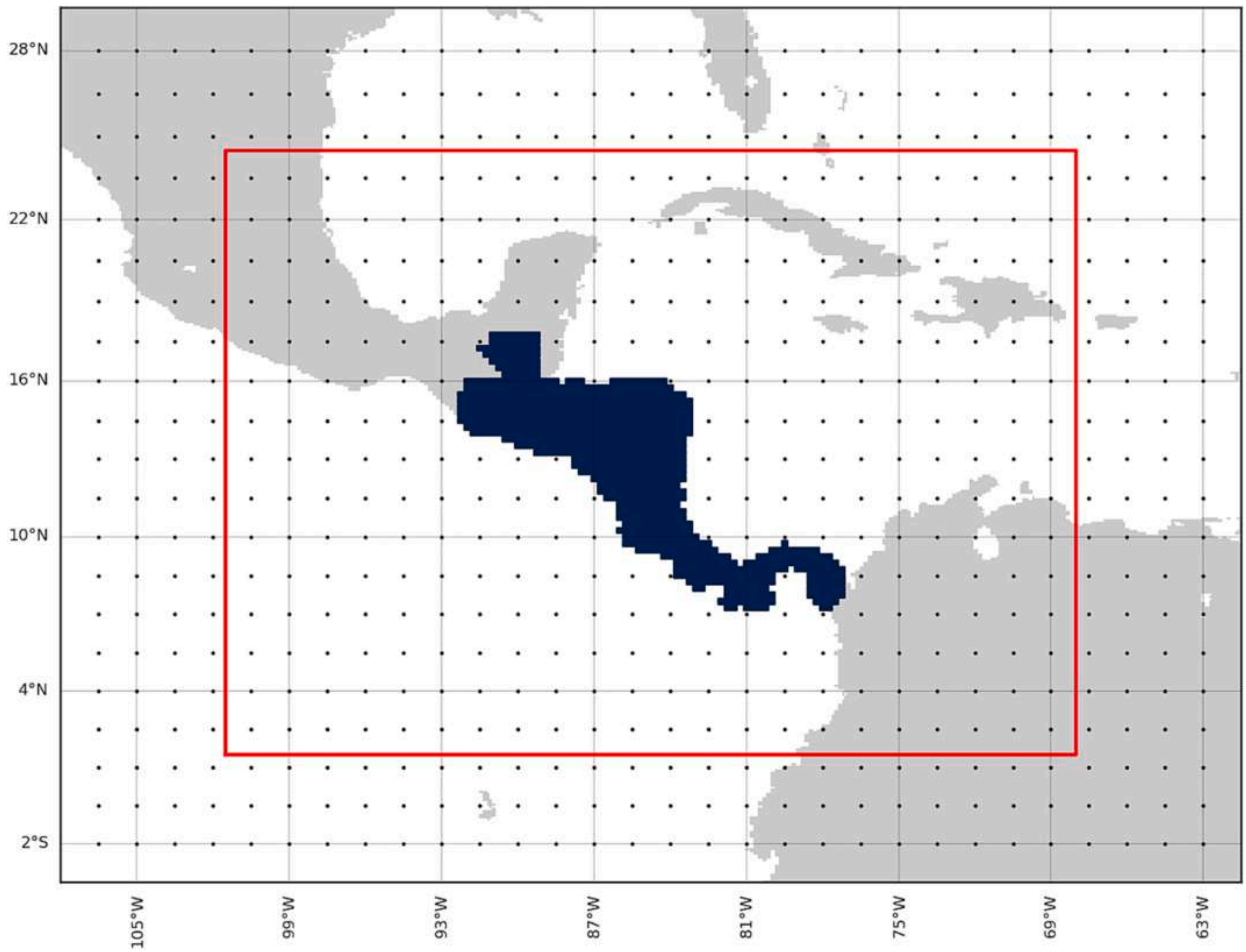

Fig. 2. Area of study. The red rectangle defines the domain used for synoptic analogy. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

taken into account. Most MOS methods (see for example Maraun, 2016) adjust different aspects (mean value, variance, quantiles, etc.) of the simulated series to match those from observations. The two major drawbacks of MOS methods are the assumption of stationarity of model biases in the future and the impact that different adjustments may have on trends given by GCMs. However, some MOS methods are able to preserve trends, and some do not even assume stationarity (see Sect. 2.1). Transfer Function methods are based on the existence of statistical relationships between large scale predictors and local predictands (Sailor and Li, 1999; Wilby et al., 2002), which are detected and calibrated in the present and applied over future simulations. These relationships can go from simple linear models (LIN) such as Multiple Linear Regression (MLR) and Generalized Linear Models (GLM) (Wilby et al., 2002) to complex nonlinear relationships based on Machine Learning (ML) algorithms such as Artificial Neural Networks (ANN; McCulloch and Pitts, 1943; Rosenblatt, 1958) or Support Vector Machines (SVM; Boser et al., 1992; Cortes and Vapnik, 1995; Vapnik, 1995). Another family of SDMs, Analog (ANA) and Weather Typing (WT), is based on the assumption of similar local conditions under similar synoptic situations (Lorenz, 1969; Zorita and von Storch, 1999). These methods search for analog synoptic conditions that occurred in the past to those projected by GCMs and their major drawback is their limitation to predict values inside of the observed range. And finally, Weather Generators (WG) are stochastic models able to produce synthetic series matching their marginal and temporal aspects with climatological statistics conditioned on properties given by GCMs simulations (Wilks and Wilby, 1999). As it has been said, all the statistical relationships found in the present (regression coefficients, model parameters for ML methods or link between weather types and local predictands) are assumed to be maintained under future climate change. Although these are the main categories, there are also hybrid methods which combine features of the different families. For a more detailed description of these families and approaches see, e.g., Maraun and Widmann (2018).

Furthermore, there is not a clear best approach. Some SDMs appear to be more suitable for some purposes and some SDMs get better results for others. See the intercomparison of a large ensemble of SDMs performed in Gutiérrez et al. (2019), in which the main strengths and limitations of each family can be found, for more information. For this reason, there is a user need for availability of downscaled climate projections using different SDMs.

There are some available packages aiming to provide users with simple tools to generate their own downscaled projections. For example, Wilby et al. (2002) published their famous 'Statistical DownScaling Model' (SDSM), a tool for rapid development of multiple, low-cost, single-site scenarios of daily surface weather variables under current and future regional climate forcing. More recently other very complete packages incorporating a large set of *state-of-the-art* SDMs from different

Table 3
List of Statistical Downscaling Models available at pyClim-SDM and their disk consumption and execution times for the test example. Downscaling times correspond to one GCM and SSP5-8.5. Disk consumption refers to calibrated models, not to the output itself. Be aware that both disk and execution times strongly depend on the number of target points and the number of predictors. ip has been used for disk consumption lower than 20 Mb and execution times lower than 5 min. For other variables, disk consumption and execution times are similar to temperature. Some methods are not common to temperature and precipitation, and others do not need training. Both cases are represented by '-'.

Family	Method	Temperature			Precipitation		
		Training Disk	Training Time	Downscaling Time	Training Disk	Training Time	Downscaling Time
RAW (interpolation)	RAW	–	–	ip	–	–	ip
	RAW-BIL	-	-	ip	-	-	ip
MOS (Model Output Statistics)	QM	-	-	ip	-	-	ip
	DQM	-	-	ip	-	-	ip
	QDM	-	-	ip	-	-	ip
	PSDM	-	-	ip	-	-	ip
ANA / WT (Analogs/Weather Typing)	ANA-SYN-1NN	-	-	10′	-	-	10′
	ANA-SYN-kNN	-	-	10′	-	-	10′
	ANA-SYN-rand	-	-	10′	-	-	10′
	ANA-LOC-1NN	ip	ip	20′	ip	ip	20′
	ANA-LOC-kNN	ip	ip	20′	ip	ip	20′
	ANA-LOC-rand	ip	ip	20′	ip	ip	20′
	ANA-VAR-1NN	-	-	10′	-	-	10′
	ANA-VAR-kNN	-	-	10′	-	-	10′
	ANA-VAR-rand	-	-	10′	-	-	10′
Linear	MLR	ip	ip	ip	-	-	-
	MLR-ANA	-	-	20′	-	-	-
	MLR-WT	ip	ip	10′	-	-	-
	GLM-LIN	-	-	-	ip	ip	ip
	GLM-EXP	-	-	-	ip	ip	ip
	GLM-CUB	-	-	-	ip	ip	ip
Machine Learning	SVM	280 Mb	1 h30′	ip	1.2 Gb	2 h	20′
	LS-SVM	270 Mb	20′	ip	502 Mb	1 h	20′
	RF	59 Gb	ip	ip	48 Gb	40′	ip
	XGB	250 Mb	ip	ip	399 Mb	ip	ip
	ANN	98 Mb	ip	ip	201 Mb	ip	ip
	CNN	314 Mb	ip	ip	1.7 Gb	ip	ip
WG (Weather Generators)	WG-NMM	-	-	-	ip	10′	ip
	WG-PDF	ip	ip	ip	ip	ip	ip

families have appeared. One example is the 'esd' R-package (Benestad et al., 2007), freely available from the Norwegian Meteorological Institute (https://github.com/metno/esd). This package is designed for climate and weather data analysis, empirical-statistical downscaling of monthly and daily data, and visualization, and it incorporates several SDMs such as EOF analysis, regression, canonical correlation analysis, multivariate regression, and weather generators. Another example is the 'downscaleR' package (Bedia et al., 2020), freely available from the Santander Meteorology Group (https://github.com/SantanderMetGroup/downscaleR). This package is designed for empirical-statistical downscaling of daily data and it includes several SDMs such as quantile mapping, regression, analogs and neural networks. And there are several more specific packages, focused on particular families or approaches, such as the 'Rglimclim' package (https://www.ucl.ac.uk/~ucakarc/work/glimclim.html) with Generalized Linear Models, 'scikit-downscale' (https://github.com/pangeo-data/scikit-downscale) with methods based on Quantile Mapping, Linear Models and Analogs, 'pyCAT' (https://github.com/wegener-center/pyCAT) with different versions of Quantile Mapping or 'ClimDown' (https://github.com/pacificclimate/ClimDown) with methods based on Quantile Mapping and Analogs.

These packages provide the user with routines and libraries and are meant for advanced users with programming knowledge. On the contrary, in this paper we present a new tool, pyClim-SDM, with a user-friendly graphical interface that makes its use simple and intuitive. Due to the lower computational expense of the statistical methods compared to RCMs, this software allows the generation of large ensembles to explore uncertainties associated with different Global Climate Models and/or emission scenarios, which is very useful especially in regions where RCMs simulations are scarce. pyClim-SDM includes a large set of *state-of-the-art* statistical methods developed and evaluated in a number of projects and initiatives including the most recent approaches based on machine learning techniques. It is prepared to downscale daily data of surface temperature, precipitation, wind, humidity and cloud cover. Additionally, any user-defined climate index can be downscaled and evaluated. An application example for a Fire Weather Index (FWI) dependent on different fundamental variables is presented.

Previous versions of the software have already been used for different purposes. The Spanish Meteorological Agency (AEMET) is responsible for the elaboration of downscaled climate projections over Spain to feed the National Plan of Adaptation to Climate Change (PNACC) and the Spanish Adaptation Viewer 'AdapteCCa' (https://escenarios.adaptecca.es). In this context, different methods have been thoroughly evaluated in the region using a previous version of the software (García-Valero, 2021; Hernanz *et al.*, 2021a; Hernanz *et al.*, 2021b; Hernanz et al., 2022a) in order to be employed over the CMIP6 generation. And in the same context, previous versions of the software were used to generate downscaled projections in Spain using the CMIP5 generation (Amblar-Francés et al., 2017). It has been also used for the elaboration of high-resolution projections in the Pyrenees region (Amblar-Francés et al., 2020) under the CLIMPY project (https://www.opcc-ctp.org/es/climpy) and to feed the associated viewer (https://www.opcc-ctp.org/en/geoportal). Previous versions of the software have been used within the EUROCLIMA + programme (https://euroclimaplus.org/) to generate climate change regional information for Central America, available at https://centroclima.org/escenarios-cambio-climatico/. In this context, the current version of the software has been co-designed with personnel of the meteorological and hydrological national services of Guatemala, Honduras, El Salvador, Panamá, Costa Rica and Nicaragua. This co-design process has been carried out through a series of training workshops in which these users were guided

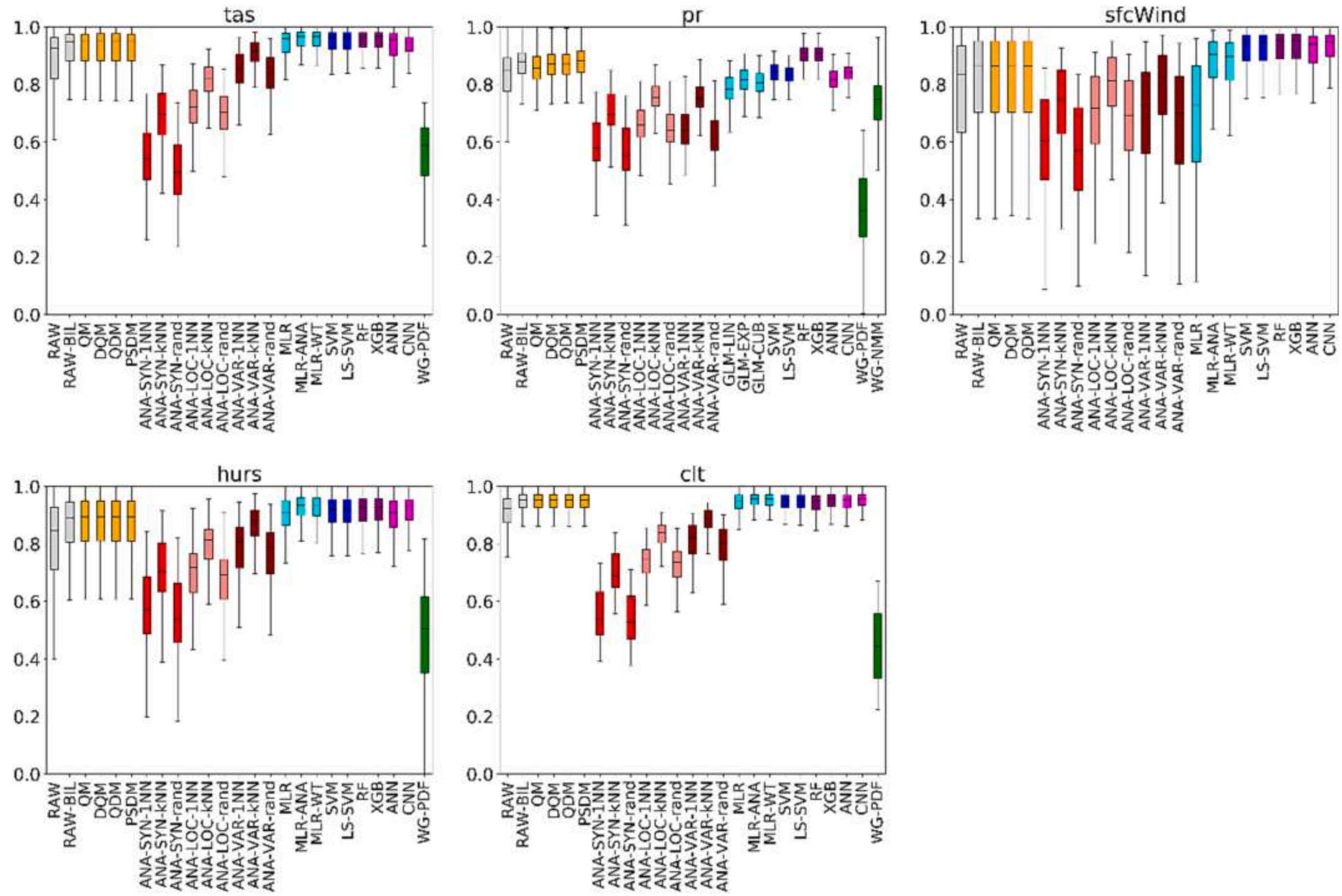

Fig. 3. Correlation for daily downscaled and observed series of temperature (tas, top-left), precipitation (pr, top-center), surface wind speed (sfcWind, top-right), relative humidity (hurs, bottom-left) and cloud cover (clt, bottom-center). SDMs have been colored by families: RAW (gray), MOS (orange), ANA / WT (red, light red and dark red), LIN (cyan), ML (dark blue, purple and pink) and WG (green). Each box contains the quartiles of all grid points (828 values) and the whiskers extend to a maximum of 1.5 times the interquartile range. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

on the use of the software to produce their own downscaled climate projections, and their requests, mainly related to the intuitiveness of the software, were implemented. An adaptation of the software has also been used for the development of downstream climate services based on seasonal forecasts (Sanchez-García *et al.*, 2021). Additionally, the software has been used for more academic purposes (Hernanz et al., 2022b; Hernanz et al., 2022c) and has the potential to be used in different projects and regions. It should be noticed that potential users of pyClim-SDM are not the final consumers of climate information, but intermediate users facing the need for high resolution climate change scenarios.

The main objective of the paper is to present pyClim-SDM, a user friendly tool for statistical downscaling of climate change projections, and the paper is organized as follows. First, we provide a description of the downscaling methods in Section 2. Then, in Section 3, we perform a test example, providing evaluation results, a description of the technical specifications of the environment used for it and the execution times, as well as a real application example. And finally, in Section 4, we include the main conclusions.

2. Downscaling methods

pyClim-SDM incorporates a large set of *state-of-the-art* SDMs belonging to different families, easily usable through its Graphical User Interface (Fig. 1):

Raw models (no downscaling):

- **RAW**: nearest grid point.
- **RAW-BIL**: bilinear interpolation.

Model Output Statistics (MOS):

- **QM**: basic Empirical Quantile Mapping (Themeßl et al., 2011). The simulated series are adjusted by quantiles to match the observed distributions, and then those corrections are applied to the future. Quantiles for future values are derived from the past distribution.
- **DQM**: Detrended Quantile Mapping (Cannon et al., 2015), adapted from the ‘ClimDown’ software (https://github.com/pacificclimate/ClimDown). The long-term trend of the observed and simulated series is first removed and then the adjustment by quantiles is done. This method preserves the signal of change in the mean values given by GCMs. Quantiles for future values are derived from the past distribution.
- **QDM**: Quantile Delta Mapping (Cannon et al., 2015), adapted from the ‘ClimDown’ software (https://github.com/pacificclimate/ClimDown). A delta change is applied for each quantile of the simulated and observed series so trends are preserved in all quantiles and no assumption on the transferability of biases is made. Quantiles for future values are derived from the future distribution.
- **PSDM**: (Parametric) Scaled Distribution Mapping (Switanek *et al.*, 2021), adapted from the ‘pyCAT’ software (https://github.com/wegener-center/pyCAT). This method is similar to QDM, but it performs the quantiles adjustment over a parametric distribution fitting

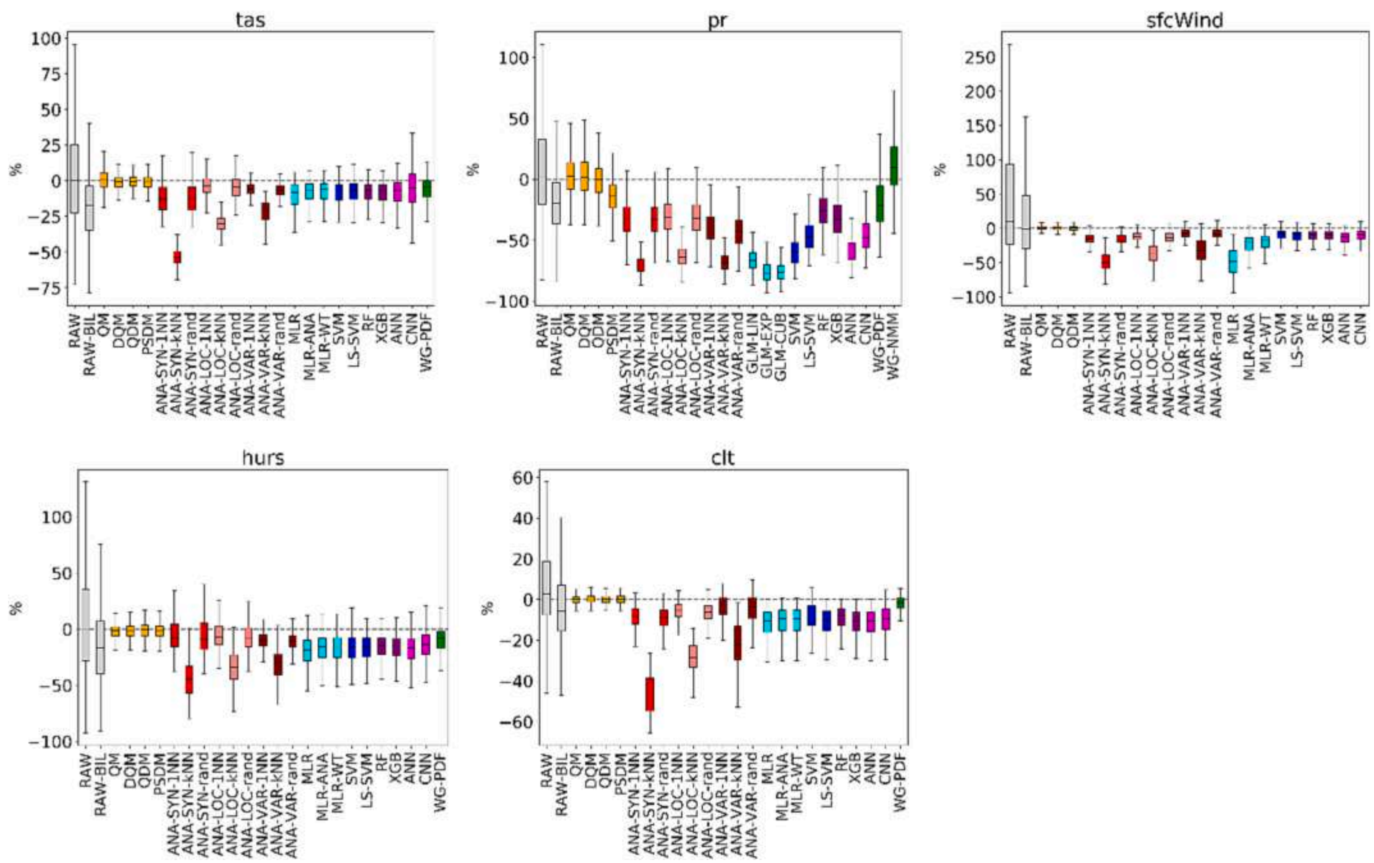

Fig. 4. Same as Fig. 3, but for relative bias of variance for downscaled daily series.

the observed and simulated series. For precipitation it uses a gamma distribution and explicitly adjusts the precipitation occurrence. For other target variables a normal distribution is used. Quantiles for future values are derived from the future distribution.

Analog/Weather Typing methods (ANA/WT):

- **ANA-SYN**: Analog based on synoptic analogy. The similarity is established using the user
- defined synoptic analogy fields, measured by the Euclidean distance. Previously
- these fields have been reduced to principal components preserving the 95%
- (configurable) of the variance. **1NN**: Nearest analog. For each target day, a
- search of similar days in the past is done, and it takes the values of the observations
- corresponding to the closest analog. **kNN**: k-nearest analogs. Similar to ANA-SYN-1NN but instead of taking values of the closest
- analog, it averages the k (configurable) closest analogs. **rand**: random analog with probabilities given by their analogy. Similar to ANA-SYN-kNN but instead of averaging the k closest analogs, it takes one of them randomly, having each analog an associated probability depending on the similarity. See Hernanz *et al.* (2021a).
- **ANA-LOC (1NN/kNN/rand)**: Same as ANA-SYN but using a combination of synoptic + local analogy. Local analogy is given by the similarity (measured by the Euclidean distance) of a set of significant predictors for each grid point and weather type. The initial set of potential predictors is defined by the user. See Petisco de Lara (2008a), Amblar-Francés et al. (2017) and Hernanz *et al.* (2021a).
- **ANA-VAR (1NN/kNN/rand)**: Same as ANA-SYN but using the spatial pattern of the target variable itself.

Linear methods (LIN): The physical variables used as predictors are defined by the user.

- **MLR**: multiple linear regression, not available for precipitation. See Amblar-Francés et al., (2017) and Hernanz et al. (2021). Based on the SDSM (Wilby et al., 2002). **MLR-ANA**: multiple linear regression based on analogs. See Petisco de Lara (2008b), Amblar-Francés et al. (2017) and Hernanz *et al.* (2021a). First, analog days are selected using the Euclidean distance of the analogy fields as similarity metric, and then, for each grid point, a multiple linear regression is calibrated using those analog days. **MLR-WT**: multiple linear regression based on weather types. Similar to ANA-MLR but using precalibrated regressions for each weather type. Based on Petisco de Lara (2008b).
- **GLM**: Generalized Linear Model, only available for precipitation. A combination of logistic regression for the wet/dry classification with a multiple linear regression for the precipitation amount (**LIN**), with the possibility of using transformed data (**EXP** for exponential and **CUB** for cubic transformations). See Amblar-Francés et al. (2017) and Hernanz *et al.* (2021a). Based on the SDSM (Wilby et al., 2002).

Machine Learning (ML) methods: The physical variables used as predictors are defined by the user.

- **SVM**: nonlinear machine learning method based on Support Vector Machines (Drucker et al., 1997), used for a classification task for the precipitation occurrence and for a regression task for the precipitation intensity and the other target variables. See Hernanz *et al.* (2021a).
- **LS-SVM**: same as SVM but using a Least-Square Support Vector Machines (Suykens and Vandewalle, 1999). See Hernanz *et al.* (2021a).

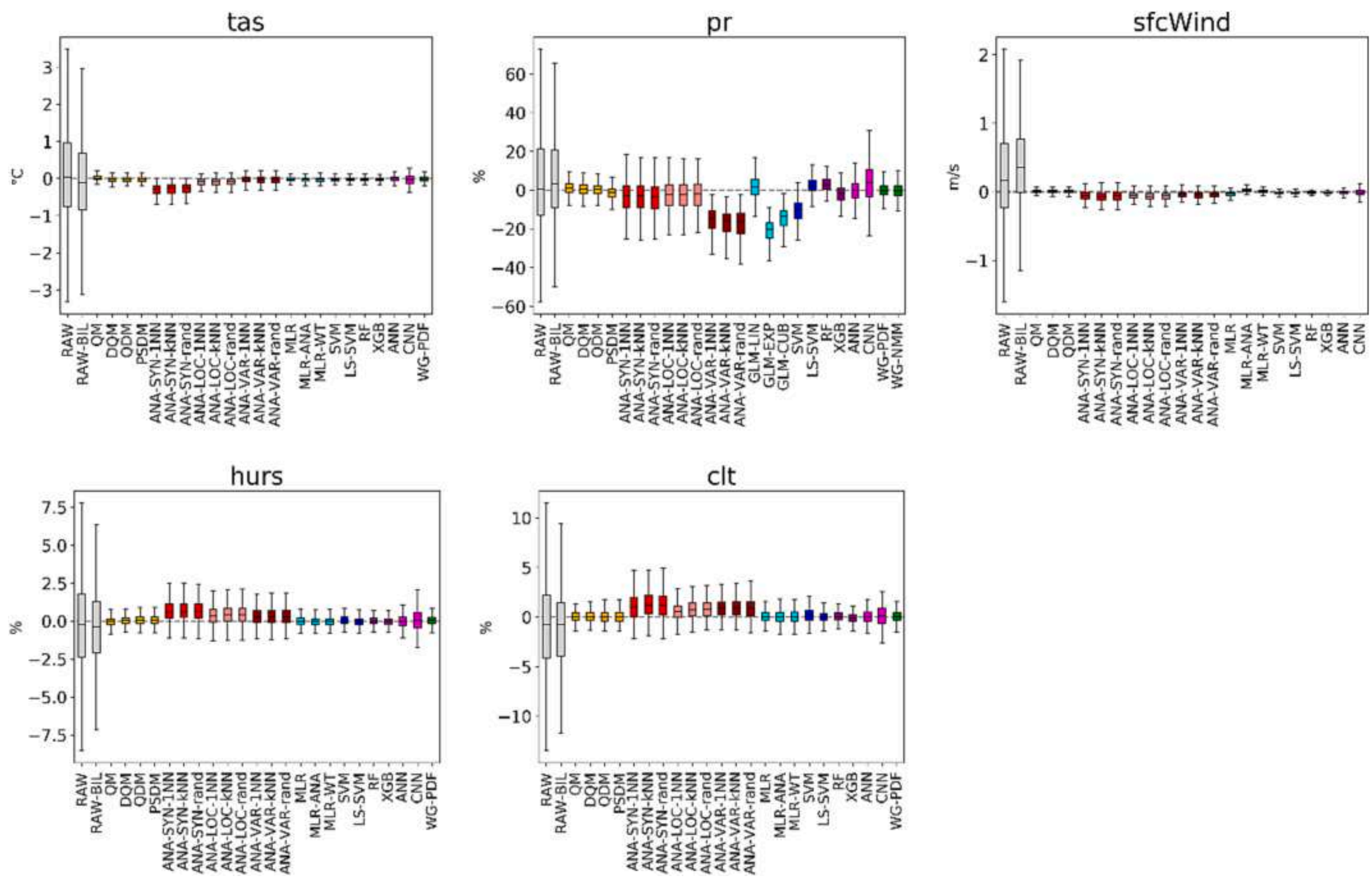

Fig. 5. Same as Fig. 3 but for mean value bias of downscaled series, relative for precipitation and absolute for the rest.

- **RF**: same as SVM but using a Random Forest (Breiman, 2001).
- **XGB**: eXtreme Gradient Boosting (Chen and Guestrin, 2016).
- **ANN**: same as SVM but using Artificial Neural Networks based on the multilayer perceptron (Rosenblatt, 1958). See García-Valero (2021) and Hernanz *et al.* (2021a).
- **CNN**: same as ANN but using Convolutional Neural Networks (see for example Gu et al., 2018).

Weather Generators:

- **WG-NMM**: Non-homogeneous Markov Model consisting of a non-parametric Weather Generator following a first-order two-state (wet/dry) Markov chain. Both the transition probabilities and the empirical distributions used for the intensity are conditioned on the precipitation given by the reanalysis/models. See Richardson (1981). Only available for precipitation.
- **WG-PDF**: weather generator based on downscaling parameters of the distributions instead of downscaling daily data. See Erlandsen et al. (2020) and Benestad (2021). Monthly means, standard deviations and rain frequencies are downscaled for each grid point using a linear regression and then daily data is randomly generated using an exponential distribution for precipitation and a normal distribution for the other target variables.

Predictors for almost all methods are taken from the four nearest coarse resolution gridpoints bilinearly interpolated. For RAW, the nearest gridpoint is used instead. And for Analog methods, the synoptic analogy is measured over a large spatial domain defined by the user. All methods, except MOS methods, use a reanalysis for the search of statistical relationships with observations, *i.e.* they follow a Perfect Prognosis approach. It must be noticed that the term MOS is being used for two different purposes. While the MOS approach refers, as explained at section 1, to a calibration strategy, the same term has been used to define a group of the SDMs available at pyClim-SDM.

3. Test example

In the following subsections a test example is shown. In Section 3.1 a description of the data used is provided. Then, an evaluation of basic variables (temperature, precipitation, wind speed, relative humidity, and cloud cover) is shown in Section 3.2. And finally, a more sophisticated use of the software possibilities is shown in Section 3.3, tackling a customized climate index. In this case a Fire Weather Index has been used. This last example is described in some detail, so it can be used as a guide. pyClim-SDM is ready to be run both in serial processing and in parallel, using High-Performance Computers (HPCs) under the popular SLURM Workload Manager (Yoo, Jette and Grondona, 2003). There is no need for HPC, everything can be done in serial processing. Nevertheless, for large datasets and computational demanding downscaling methods, processing times in serial can be too long.

This test example has been performed combining the following two environments: (1) For preprocessing and postprocessing steps we have used a virtual machine with 32 CPUs and 64 GB of memory under a virtual infrastructure VMware® 7.0. SO CentOS Linux release 7.6.1810. Virtual servers: 18xFujitsu Primergy RX2540 M5 with 2xIntel Xeon Gold 5218 16cores at 2,3 GHz y 768 GB of memory DDR$-2933 MHz. (2) For training methods and downscaling we have used a High-Performance Computer consisting of two clusters of 140 nodes. Each node has 2 processors AMD EPYC™7742 of 64 cores, 256 GB DDR4-3200 of memory and 1 SSD of 240 GB. SO Red Hat Enterprise Linux.

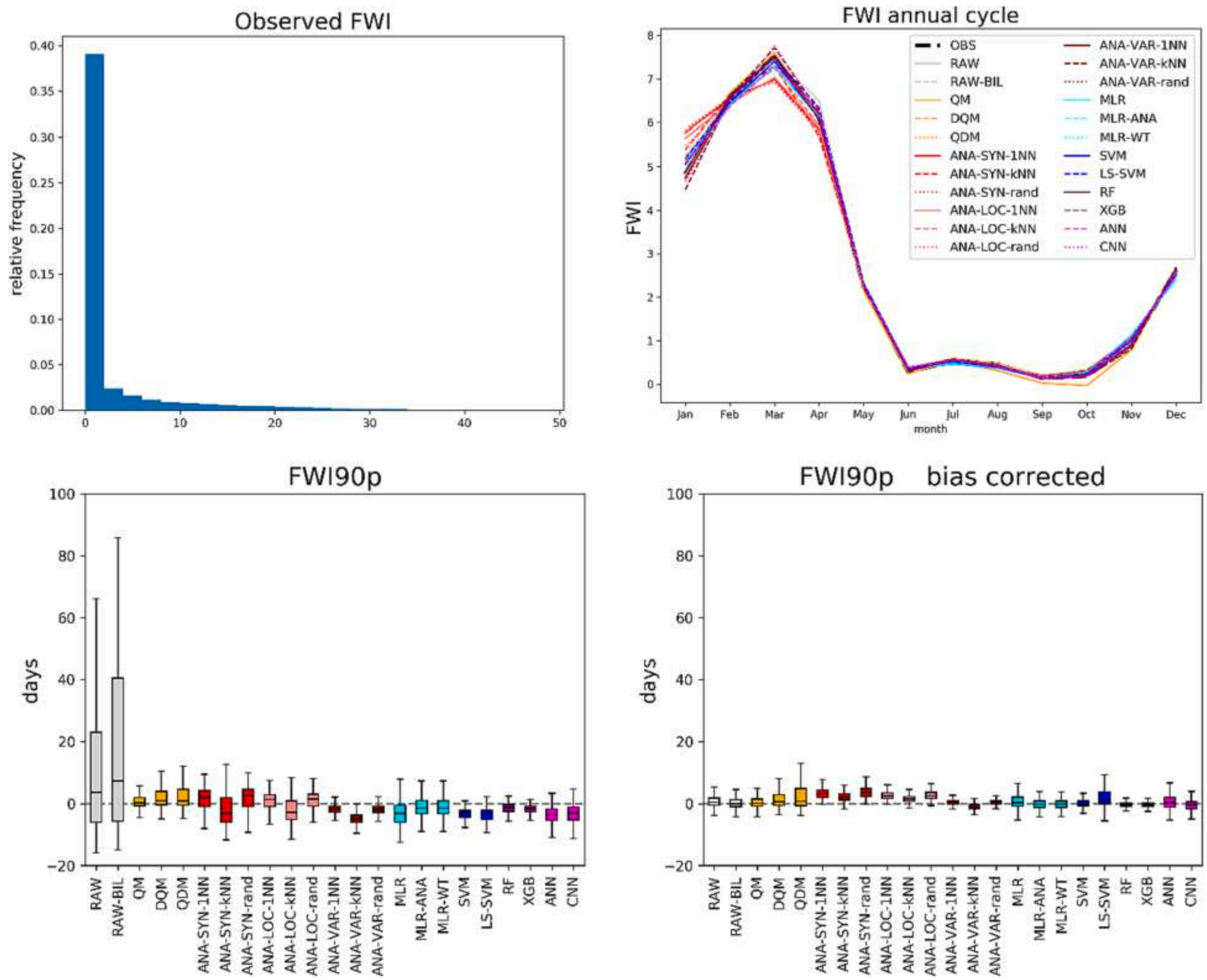

Fig. 6. Histogram of observed FWI (top-left), annual cycle of observed and downscaled FWI (top-right), bias for the number of days over the 90th percentile of FWI in the dry season by SDMs (bottom-left) and SDMs bias corrected using QM by seasons (bottom-right). SDMs have been colored by families: RAW (gray), MOS (orange), ANA / WT (red, light red and dark red), LIN (cyan) and ML (dark blue, purple and pink). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.1. Data

Predictands have been taken from the reanalysis ERA5 (Hersbach et al., 2020) of the European Centre for Medium-Range Weather Forecasts (ECMWF) with a resolution of 0.25° over land (828 points). The election of a reanalysis instead of actual observations has been done due to the lack of availability of observations for the different meteorological variables analyzed. It should be noticed that ERA5 precipitation is not assimilated, and therefore substantial biases can exist. Thus, results here shown for precipitation (and for all variables in general) might differ when using actual observations. Nonetheless, presenting an accurate evaluation of the SDMs is not the purpose of this paper, for their application to different regions and/or datasets is unique and SDMs must be evaluated for each particular application. Predictors have been taken from the same reanalysis but with a resolution of 1.5° and from the seven CMIP6 (Eyring et al., 2016) models listed in Table 1.

For each variable, predictors listed in Table 2 have been used, with a spatial resolution of 1.5° × 1.5° (GCMs outputs have been previously interpolated with bilinear interpolation from their original resolutions). All predictors have been standardized using their own mean and standard deviation over the reference period 1979–2005. For the synoptic analogy, zonal and meridional wind components at 850 and 500 hPa and relative humidity at 700 hPa have been used, based on Ribalaygua et al. (2018), in the domain (23.5°N, 100.5°W, 1°N, 70.5°W). See area of study at Fig. 2.

All these inputs need to be prepared by the user in a specific but standard format, and storaged in the corresponding folders (see the User Manual accompanying the software for detailed information).

SDMs have been trained in 1979–2005 and evaluated over reanalysis in 2006–2020. Projections correspond to run r1i1p1f1 of SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5 scenarios (see IPCC, 2021) in 2015–2100, plus the historical scenario in 1951–2014. See (Table 3).

3.2. Evaluation

In this subsection, SDMs have been evaluated for the following target variables in a historical period over a reanalysis: temperature, precipitation, wind speed, relative humidity, and cloud cover. Following the VALUE framework established by Maraun et al. (2015), the first aspect evaluated is the correlation between the downscaled and observed daily series (Fig. 3). In general, Analog methods and Weather Generators present poor correlations, while the other families reach higher correlations. High correlations are a good indicator for the ability of SDMs to reproduce temporal aspects, such as warm or dry spells. RAW (interpolation) also shows high correlations, but they do not capture the variability at all (Fig. 4). MOS methods are, in general, the ones with lower biases in the variance, while other methods tend to a certain underestimation of it. Precipitation variability presents a special difficulty

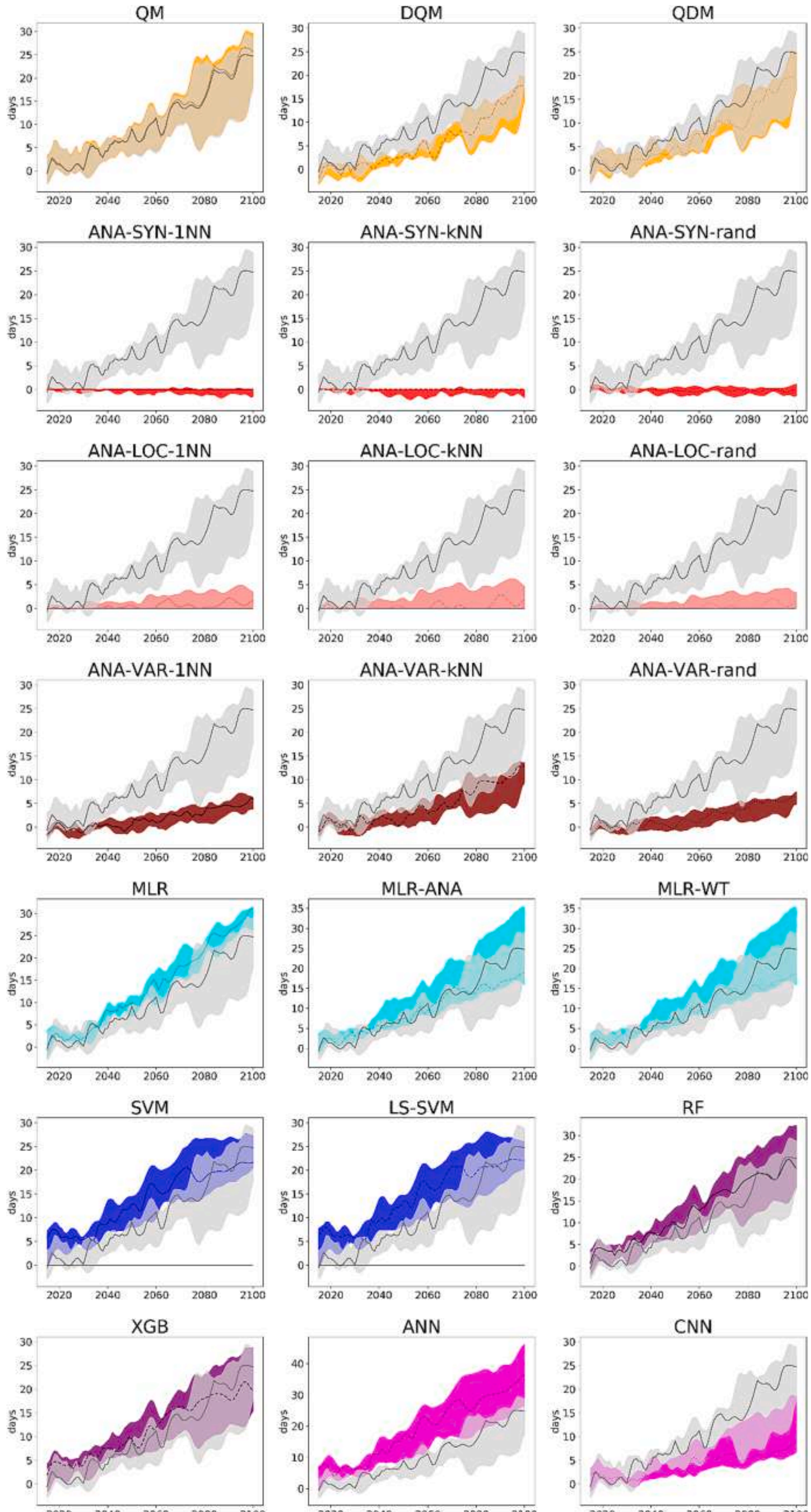

Fig. 7. Evolution of the change in the number of days over the 90th percentile (days), spatially averaged, by RAW (gray) and each SDM bias corrected using QM by season over GCMs in Table 1 and under SSP5-8.5 from 2015 to 2100. SDMs have been colored by families: MOS (orange), ANA / WT (red, light red and dark red), LIN (cyan) and ML (dark blue, purple and pink). The shaded area expands from the 25th to the 75th percentiles, and the line represents the median of the multi-model ensemble. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

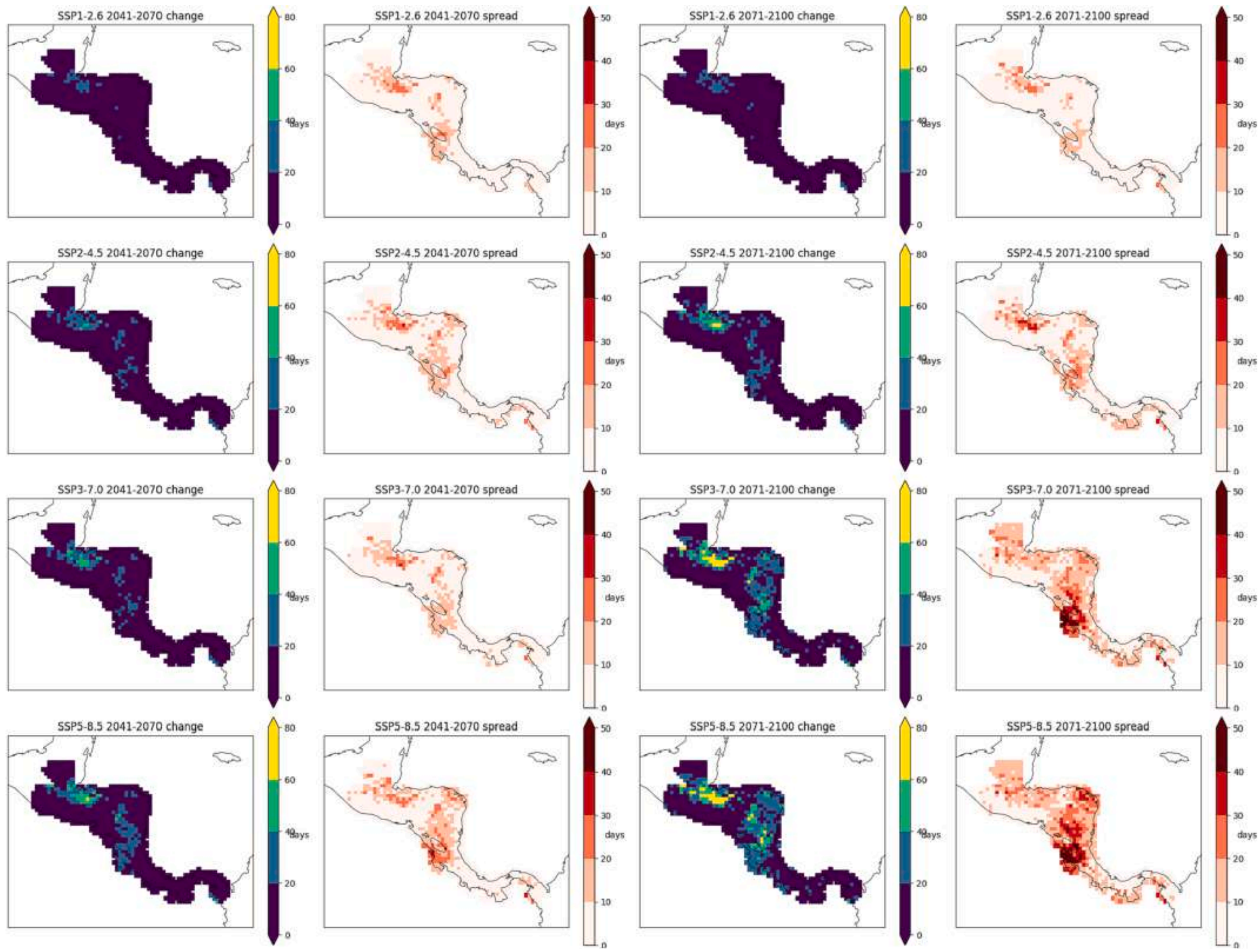

Fig. 8. Change in the number of days over the 90th percentile (days) in the dry season, given by the ensemble median (columns 1 and 3) of the GCMs from Table1, and ensemble spread (columns 2 and 4), given by the interquartile range. The downscaling method applied is XGB, bias corrected using QM by season. SSP1-2.6 (first row), SSP2-4.5 (second row), SSP3-7.0 (third row) and SSP5-8.5 (fourth row). Changes correspond to 2041–2070 (columns 1 and 2) and 2071–2100 (columns 3 and 4) with respect to the reference period (1979–2005).

for most methods, suggesting the need of a posterior bias correction for adjusting the variance. Finally, Fig. 5 shows the mean values bias over the whole testing period. Results achieved by most SDMs confirm the benefits of downscaling over performing a simple interpolation. pyClim-SDM allows for user-defined seasons in order to analyze this and other metrics by season. All evaluation figures shown in this study are produced by the tool, so the user can easily compare different metrics or aspects among the collection of included methods. Other important figures are also generated by pyClim-SDM, such as biases for other indexes, e.g. percentiles, extremes, spells, etc., or the spatial distributions of each variable and the correlation of those spatial patterns between observations and simulations. The analysis shown in this study is just a sample of what can be easily produced by the software, but in order to make decisions, users are encouraged to analyze more aspects. Take as an example the analysis on the spatial aspects, temporal aspects and extremes done in Widmann et al. (2019), Maraun et al. (2019) and Hertig et al. (2019), respectively.

3.3. Application

In this subsection, we show and explain how to use the software to downscale a user-defined target variable. Any variable can be potentially downscaled, from direct GCMs outputs such as soil moisture or solar radiation to secondary variables derived from them, as long as observations are available. In this case we have used a Fire Weather Index (FWI; Van Wagner, 1987), which depends on temperature, precipitation, wind speed and humidity. First, all SDMs have been applied and evaluated in a historical period over reanalysis, with and without a posterior bias correction. Then, trends of all SDMs have been compared with those given by raw GCMs. And finally, projections for several emission scenarios are shown for a selected SDM and different temporal horizons.

In order to downscale a user-defined target variable, observations of that variable are needed. In this case, we have generated the observed FWI from temperature, precipitation, wind speed, and relative humidity observations using the FireDanger package (https://github.com/steidani/FireDanger). For some methods, those which also make use of the coarse resolution version of the predictand as predictor, and for some purposes, the variable is also needed at low resolution. We have calculated the FWI for both reanalysis and GCMs and generated their netCDF files following the format explained in the input_data_template that accompanies the software.

In this example we will apply bias correction after downscaling. In order to apply both steps, long downscaled and observed series are needed. For this reason, instead of using the train/test split from the previous section (only 15 testing years), we will downscale 42 years

(1979–2020) using a k-fold approach. A k-fold technique consists on splitting the whole period in train/test several times (five in this case), so each round a different subset is downscaled, and the independency between training and testing datasets is preserved. This way, the preprocess, training, and process steps need to be run five times, selecting folds from 1 to 5. After the last fold has been processed, the bias correction can be done, selecting now 'all_testing', so the 42 years are used in this step.

When using new target variables, it is important to characterize them before applying any downscaling. Fig. 6 (top-left) shows the distribution (histogram) of the observed FWI. Most days present values close to zero, and the distribution is clearly non-gaussian. The interest of the FWI is in the high values, which indicate high risk of fire, so we will focus the study on the tail of the distribution, specifically on the number of days with FWI over the 90th percentile (FWI90p). In Fig. 6 (top-right) we also see the annual cycle of FWI, which is strongly related to the precipitation annual cycle (not shown). We will focus only on the Central American dry season, from January to April, when the highest FWIs are reached. As it can be seen in Fig. 6 (top-right), all SDMs accurately capture the annual cycle averaged over the whole study region. When analyzing how well they capture the tail of the distribution (extremes), given by FWI90p in the dry season, Fig. 6 (bottom-left) shows how all SDMs present a clear advantage over the simple interpolation. Nevertheless, after applying a bias correction by season using QM, Fig. 6 (bottom-right) biases both by RAW and SDMs are strongly reduced. This indicates that the simple interpolation can be a valid approach as long as a bias correction is applied over it (which is actually equivalent to the MOS methods here presented).

At this stage it is important to introduce some potential limitations of SDMs. One is the coherency of the spatial patterns, for methods that downscale each point independently. This limitation affects all methods included in pyClim-SDM except some versions of Analogs (ANA-SYN and ANA-VAR), and its impact can be analyzed with the evaluation metrics provided by the software, as it varies for each specific case (region, dataset, set of predictors, etc). Another one is the multivariate consistency, as methods downscale each target variable separately, which also affects all method included. As it has been mentioned, the software allows to downscale multivariate index directly in order to avoid this issue. And finally, different SDMs present different abilities to extrapolate to a future different climate. Although Fig. 6 (bottom-right) indicates that all SDMs can capture FWI extremes with good accuracy, that does not guarantee they will behave with the same accuracy in the future, so their future behaviors must be analyzed. An important analysis for SDMs is the test on their ability to preserve the signal of change given by GCMs averaged over large areas (because they operate in different spatial scales). Fig. 7 compares trends given by RAW with trends given by all bias corrected SDMs for FWI90p in the dry season under the SSP5-8.5 scenario. As it can be seen, many SDMs are incapable of reproducing those trends. The same behavior is frequently found in other variables with a strong signal of change, such as temperature (not shown). But for other variables or indexes it is possible that all SDMs preserve trends given by GCMs, so this analysis, which is integrated in pyClim-SDM must be carried out before any real application. It should be clarified that although differences between trends given by raw GCMs and downscaled simulations point to an incapability of the SDM to preserve trends, in some cases, a modified trend given by the SDM might be more realistic than the raw GCM (see Maraun et al., 2017). Thus, although pyClim-SDM produces figures allowing the comparison of trends for all the user selected indexes, the user needs to be critic on the reliability of raw GCMs trends.

Finally, we will focus on the method XGB, which has shown a very good accuracy in the historical evaluation and preserves trends given by GCMs. A method based on decision trees such as XGB cannot produce values out of the observed range. And GCMs project a marked increase for FWI in the future. Thus, it is unexpected for XGB to preserve trends given by GCMs for FWI. Nevertheless, it should be noticed that the trend has been analyzed only for FWI90p, *i.e.,* the count of days over a threshold which lies inside the observed range. For this application example we will show the projected change in FWI90p in the dry season, using XGB bias corrected by season with QM, in two temporal horizons (2041–2070 and 2071–2100) under SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5 scenarios. Fig. 8 shows the change in FWI90p with respect to the reference period (do not mistake it with the absolute number of days). Let us bear in mind that, being the dry season formed by four months (around 120 days), the mean number of days over the 90th percentile in the reference period should be around 12 days. The signal of change is consistent over the different SSPs and temporal horizons, with a more intense increase by the end of the century and by the most emissive scenarios (Fig. 8.) While some regions will not see FWI90p increased even under the most emissive scenario, other regions will experience a significant increase no matter which scenario we end up in. The most remarkable result is the one reached for a region in the North West, between Honduras and Guatemala, where even for the most conservative scenario there is a projected change of more than 20 days, and for the other SSPs of 60 to 80 days. This means that the threshold that used to be surpassed by 12 out of 120 days in the reference period will be surpassed by 72 to 92 days under those scenarios. Such a result must be analyzed in terms of robustness too, of course. Thus, Fig. 8 shows not only the projected change but also the uncertainty given by the multi-model ensemble spread. Some regions show higher uncertainties than others. Particularly, the mentioned region presents uncertainties of the same order as the mean signal of change for the less emissive scenarios, while for the most emissive ones it does not appear to present significant uncertainties. Thus, for the less emissive scenarios, although there is an important level of uncertainty associated with the models, the mean signal of change appears to be consistent with the spatial pattern projected under the other scenarios. And for the most emissive scenarios, the spread among models is quite low, so the extreme signal of change in that region appears to be very robust. It should be noticed that, although the increase on the number of days over the 90th percentile is very high in some regions, the 90th percentile is not really high itself in those areas (not shown). Thus, those increments should not necessarily represent an actual high fire danger. This application example using the FWI should be understood as an illustrative example on how the software can be used and the possibilities it offers, not as a comprehensive risk analysis.

4. Conclusions

In this paper we have presented a service based on a new software for statistical downscaling of climate change projections, pyClim-SDM. This service provides the user with a very simple graphical interface and it is ready to be used by only defining a few basic settings and preparing input data in a standard specific format. The software includes a set of *state-of-the-art* methods belonging to different families, which have been evaluated in the test example, showing clear benefits compared to simple interpolations. pyClim-SDM is prepared to downscale surface daily temperature, precipitation, wind, humidity, and cloud cover. Additionally, the software can be applied to any user-defined target variable, as it has been shown for a Fire Weather Index. This way, multivariable indexes can be tackled as target variables themselves instead of being calculated from the downscaled fundamental variables, avoiding thus potential intervariable incoherencies. Due to the high computational cost of some methods, this software can be run both in serial processing and in parallel in a HPC cluster. Although this service was originally developed within the Spanish national climate change adaptation plan for evaluating and generating downscaled climate change projections, it can be easily applied for similar purposes both in academic or operational frameworks. Results shown in this paper do not intend to conform an exhaustive comparison between the available methods or to outline one specific method over the others. Such a comparison and evaluation must be made by the user, focusing on his/her specific purposes, and bearing in mind that results can be different for different regions and/or datasets. Although pyClim-SDM can be easily used without any specific knowledge in climate or downscaling, its use is recommended only for meteorologist/climatologist who can assess the choice of methods and predictors, and properly interpret the evaluation results. Otherwise, projections generated with this tool could be completely wrong.

In conclusion, pyClim-SDM is a user-friendly software which will allow users to evaluate and generate their own data with minimum set up, and which has already been used in different projects.

Data and code availability

CMIP6 GCMs are available at Earth System Grid Federation nodes (https://esgf-node.llnl.gov/search/cmip6/) and at Copernicus Climate Data Store (CDS; https://cds.climate.copernicus.eu/). ERA5 reanalysis is available at the Meteorological Archival and Retrieval System of the ECMWF (MARS; https://confluence.ecmwf.int/display/UDOC/MARS + user + documentation) and at Copernicus CDS The high resolution observational grid is available at AEMET website (https://www.aemet.es/es/serviciosclimaticos/cambio_climat/datos_diarios?w = 2). And pyClim-SDM is available at https://github.com/ahernanzl/pyClim-SDM under GNU General Public License v3.0.

CRediT authorship contribution statement

Alfonso Hernanz: Software, Data curation, Visualization, Writing – original draft. **Carlos Correa:** Software, Data curation, Writing – review & editing. **Juan Andrés García-Valero:** Writing – review & editing. **Marta Domínguez:** Data curation, Writing – review & editing. **Esteban Rodríguez-Guisado:** Writing – review & editing. **Ernesto Rodríguez-Camino:** Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data used for this study is freely available (see the Data and Code availability section)

Acknowledgements

We acknowledge the World Climate Research Programme which, through its Working Group on Coupled Modeling, coordinated and promoted CMIP6. We thank the climate modeling groups (listed in Table 1 of this paper) for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. A special gratitude goes to Eduardo Petisco de Lara, Pilar Amblar Francés, María Asunción Pastor Saavedra and Petra Ramos Calzado for their previous developments in the analog and regression downscaling methods. Finally, we would like to thank Carlos Santos Burguete, Xavier Calbet and José Luis Casado Rubio for their efforts to strengthen the developers community inside AEMET. And we also acknowledge two anonymous reviewers, whose constructive comments have helped improve the quality and clarity of the manuscript.

References

Amblar-Francés, P., Casado-Calle, M.J., Pastor-Saavedra, M.A., Ramos-Calzado, P., Rodríguez-Camino, E., 2017. Guía de escenarios regionalizados de cambio climático sobre España a partir de los resultados del IPCC-AR5. Available at. https://www.aemet.es/documentos/es/conocermas/recursos_en_linea/publicaciones_y_estudios/publicaciones/Guia_escenarios_AR5/Guia_escenarios_AR5.pdf.

Amblar-Francés, M.P., Ramos-Calzado, P., Sanchis-Lladó, J., Hernanz-Lázaro, A., Peral-García, M.C., Navascués, B., Dominguez-Alonso, M., Pastor-Saavedra, M.A., Rodríguez-Camino, E., 2020. High resolution climate change projections for the Pyrenees region. Adv. Sci. Res. 17, 191–208. https://doi.org/10.5194/asr-17-191-2020.

Bedia, J., Baño-Medina, J., Legasa, M.N., Iturbide, M., Manzanas, R., Herrera, S., Casanueva, A., San-Martín, D., Cofiño, A.S., Gutiérrez, J.M., 2020. Statistical downscaling with the downscaleR package (v3.1.0): contribution to the VALUE intercomparison experiment. Geosci. Model Dev. 13, 1711–1735. https://doi.org/10.5194/gmd-13-1711-2020.

Benestad, R.E., 2021. A Norwegian approach to downscaling. Geoscientific Model Development Discussion (in review). https://doi.org/10.5194/gmd-2021-176.

Benestad, R. E., A. Mezghani, and K. M. Parding, (2007). ESD: Climate analysis and empirical-statistical downscaling (ESD) package for monthly and daily data. Tech. Rep., 272 pp., https://github.com/metno/esd.

Boser, B., Guyon, I., Vapnik, V., 1992. A Training Algorithm for Optimal Margin Classifier. In: Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory 5. https://doi.org/10.1145/130385.130401.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Cannon, A.J., Sobie, S.R., Murdock, T.Q., 2015. Bias correction of GCM precipitation by quantile mapping: how well do methods preserve changes in quantiles and extremes? J. Clim. 28, 6938–6959. https://doi.org/10.1175/JCLI-D-14-00754.1.

Charles, S.P., Bates, B.C., Smith, I.N., Hughes, J.P., 2004. Statistical downscaling of daily precipitation from observed and modelled atmospheric fields. Hydrological Processes, Special Issue: Scale and Scaling in Hydrology 18, 1373–1394. https://doi.org/10.1002/hyp.1418.

Chen, T., Guestrin, C., 2016. In: XGBoost: A Scalable Tree Boosting System. ACM, New York, NY, USA, pp. 785–794. https://doi.org/10.1145/2939672.2939785.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20, 273–297. https://doi.org/10.1007/BF00994018.

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arneth, A., Arsouze, T., Bergmann, T., Bernadello, R., Bousetta, S., Caron, L.-P., Carver, G., Castrillo, M., Catalano, F., Cvijanovic, I., Davini, P., Dekker, E., Doblas-Reyes, F. J., Docquier, D., Echevarria, P., Fladrich, U., Fuentes-Franco, R., Gröger, M., v. Hardenberg, J., Hieronymus, J., Karami, M. P., Keskinen, J.-P., Koenigk, T., Makkonen, R., Massonnet, F., Ménégoz, M., Miller, P. A., Moreno-Chamarro, E., Nieradzik, L., van Noije, T., Nolan, P., O'Donnell, D., Ollinaho, P., van den Oord, G., Ortega, P., Prims, O. T., Ramos, A., Reerink, T., Rousset, C., Ruprich-Robert, Y., Le Sager, P., Schmith, T., Schrödner, R., Serva, F., Sicardi, V., Sloth Madsen, M., Smith, B., Tian, T., Tourigny, E., Uotila, P., Vancoppenolle, M., Wang, S., Wårlind, D., Willén, U., Wyser, K., Yang, S., Yepes-Arbós, X., and Zhang, Q. (2021). The EC-Earth3 Earth System Model for the Climate Model Intercomparison Project 6, Geosci. Model Dev. Discuss. [preprint], doi: 10.5194/gmd-2020-446, in review.

Drucker, H., Burges, C.J., Kaufman, L., Smola, A., Vapnik, V., 1997. Support vector regression machines. Adv. Neural Inf. Proces. Syst. 9, 155–161.

Erlandsen, H.B., Parding, K.M., Benestad, R., Mezghani, A., Pontoppidan, M., 2020. A hybrid downscaling approach for future temperature and precipitation change. J. Appl. Meteorol. Climatol. 59 (11), 1793–1807. https://doi.org/10.1175/JAMC-D-20-0013.1.

Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J., Taylor, K.E., 2016. Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. Geosci. Model Dev. 9, 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016.

García-Valero, J.A., 2021. Redes neuronales artificiales. Aplicación a la regionalización de la precipitación y temperaturas diarias. AEMET Nota Técnica, Área de Evaluación y Modelización del Cambio Climático. AEMET, Spain.

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., Chen, T., 2018. Recent advances in convolutional neural networks. Pattern Recogn. 77, 354–377.

Gutiérrez, J.M., Maraun, D., Widman, M., Huth, R., Hertig, E., Benestad, R., Roessler, O., Wibig, J., Wilcke, R., Kotlarski, S., San Martin, D., Herrera, S., Bedia, J., Casanueva, A., Manzanas, R., Iturbide, M., Vrac, M., Dubrovsky, M., Ribalaygua, J., Portoles, J., Raty, O., Raisanen, J., Hingray, B., Raynaud, D., Casado, M.J., Ramos, P., Zerenner, T., Turco, M., Bosshard, T., Stepanek, P., Bartholy, J., Pongracz, R., Keller, D.E., Fischer, A.M., Cardoso, R.M., Soares, P.M.M., Czernecki, B., Page, C., 2019. 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.

Hernanz, A., García-Valero, J.A., Domínguez, M., Ramos-Calzado, P., Pastor-Saavedra, M.A., Rodríguez-Camino, E., 2021. Evaluation of statistical downscaling methods for climate change projections over Spain: present conditions with perfect predictors. Int. J. Climatol. 42 (2), 762–776. https://doi.org/10.1002/joc.7271.

Hernanz, A., García-Valero, J.A., Domínguez, M., Rodríguez-Camino, E., 2022a. Evaluation of statistical downscaling methods for climate change projections over Spain: future conditions with pseudo reality (transferability experiment). Int. J. Climatol. 42 (7), 3987–4000.

Hernanz, A., García-Valero, J.A., Domínguez, M., Rodríguez-Camino, E., 2022b. Evaluation of statistical downscaling methods for climate change projections over Spain: present conditions with imperfect predictors (global climate model experiment). Int. J. Climatol. 42 (13), 6793–6806.

Hernanz, A., García-Valero, J.A., Domínguez, M., Rodríguez-Camino, E., 2022c. 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 (6) https://doi.org/10.1002/asl.1087.

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., Thépaut, J.-N., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146 (730), 1999–2049.

Hertig, E., Maraun, D., Bartholy, J., Pongracz, R., Vrac, M., Mares, I., Gutiérrez, J.M., Wibig, J., Casanueva, A., Soares, P.M.M., 2019. 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 (9), 3846–3867.

Lorenz, E.N., 1969. Atmospheric predictability as revealed by naturally occurring analogues. J. Atmos. Sci. 26 (4), 636–646.

Maraun, D., 2016. Bias correcting climate change simulations - a critical review. Curr. Clim. Change. Rep. 2, 211–220. https://doi.org/10.1007/s40641-016-0050-x.

Maraun, D. and Widmann, M. (2018). Statistical Downscaling and Bias Correction for Climate Research. Cambridge: Cambridge University Press. doi: 10.1017/9781107588783.

Maraun, D., Widmann, M., Gutiérrez, J.M., Kotlarski, S., Chandler, R.E., Hertig, E., Wibig, J., Huth, R. and Wilcke, R.A. (2015) Value: a framework to validate downscaling approaches for climate change studies. Earth's Future, 3(1), 1– 14. doi: 10.1002/2014EF000259.

Maraun, D., Shepherd, T.G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J.M., Hagemann, S., Richter, I., Soares, P.M.M., Hall, A., Mearns, L.O., 2017. Towards process-informed bias correction of climate change simulations. Nature Clim Change 7 (11), 764–773.

Maraun, D., Huth, R., Gutiérrez, J.M., Martín, D.S., Dubrovsky, M., Fischer, A., Hertig, E., Soares, P.M.M., Bartholy, J., Pongrácz, R., Widmann, M., Casado, M.J., Ramos, P., Bedia, J., 2019. The VALUE perfect predictor experiment: Evaluation of temporal variability. Int. J. Climatol. 39 (9), 3786–3818.

McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. 5, 115–133. https://doi.org/10.1007/BF02478259.

Müller, W.A., Jungclaus, J.H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., Ilyina, T., Kleine, T., Kornblueh, L., Li, H., Modali, K., Notz, D., Pohlmann, H., Roeckner, E., Stemmler, I., Tian, F., Marotzke, J., 2018. A higher-resolution version of the max planck institute earth system model (MPI-ESM1.2-HR). J. Adv. Model. Earth Syst. 10 (7), 1383–1413.

Petisco de Lara, S.E., 2008a. Método de regionalización de precipitación basado en análogos. Explicación y Validación. AEMET Nota Técnica 3A, Área de Evaluación y Modelización del Cambio Climático. AEMET, Spain.

Petisco de Lara, S.E., 2008b. Método de regionalización de temperatura basado en análogos. Explicación y Validación. AEMET Nota Técnica 3B, Área de Evaluación y Modelización del Cambio Climático. AEMET, Spain.

Ribalaygua, J., Gaitán, E., Pórtoles, J., Monjo, R., 2018. Climatic change on the Gulf of Fonseca (Central America) using two-step statistical downscaling of CMIP5 model outputs. Theor. Appl. Climatol. 132 (3-4), 867–883.

Richardson, C.W., 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resour. Res. 17 (1), 182–190. https://doi.org/10.1029/WR017i001p0018.

Rosenblatt, F., 1958. The perceptron: a probabilistic model for information storage and organization in the brain. Psychol. Rev. 65, 386–408.

Sailor, D.J., Li, X., 1999. A semiempirical downscaling approach for predicting regional temperature impacts associated with climatic change. J. Clim. 12 (1), 103–114. https://doi.org/10.1175/1520-0442-12.1.103.

Schoof, J.T., 2013. Statistical downscaling in climatology. Geography. Compass 7 (4), 249–265. https://doi.org/10.1111/gec3.12036.

Suykens, J., Vandewalle, J., 1999. Least squares support vector machine classifiers. Neural Process. Lett. 9, 293–300. https://doi.org/10.1023/A:1018628609742.

Swart, N.C., Cole, J.N.S., Kharin, V.V., Lazare, M., Scinocca, J.F., Gillett, N.P., Anstey, J., Arora, V., Christian, J.R., Hanna, S., Jiao, Y., Lee, W.G., Majaess, F., Saenko, O.A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., Winter, B., 2019. The Canadian Earth System Model version 5 (CanESM5.0.3). Geosci. Model Dev. 12, 4823–4873. https://doi.org/10.5194/gmd-12-4823-2019.

Themeßl, M.J., Gobiet, A., Leuprecht, A., 2011. Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. Int. J. Climatol. 31, 1530–1544. https://doi.org/10.1002/joc.2168.

Trzaska, S., Schnarr, E., 2014. A review of downscaling methods for climate change projections: African and Latin American Resilience to Climate Change (ARCC). Available at. http://www.ciesin.org/documents/Downscaling_CLEARED_000.pdf.

UNFCCC (2015). Paris Agreement. In UNFCCC, COP Report No. 21, Addendum, at 21, U. N. Doc. FCCC/CP/2015/10/Add, 1 (Jan. 29, 2016). https://unfccc.int/sites/default/files/english_paris_agreement.pdf.

Van Wagner, C.E., 1987: Development and Structure of the Canadian Forest Fire Weather Index System. Forestry Technical Report 35. Canadian Forest Service, Ottawa, Canada. Available at https://cfs.nrcan.gc.ca/publications?id=19927.

Vapnik, V.N. (Ed.), 1995. The Nature of Statistical Learning Theory. Springer New York, New York, NY.

Volodin, E.M., Mortikov, E.V., Kostrykin, S.V., Galin, V.Y., Lykosov, V.N., Gritsun, A.S., Diansky, N.A., Gusev, A.V., Yakovlev, N.G., 2017. Simulation of modern climate with the new version of the INM RAS climate model. Izv. Atmos. Ocean. Phys. 53 (2), 142–155.

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., Huth, R., 2019. Validation of spatial variability in downscaling results from the VALUE perfect predictor experiment. Int. J. Climatol. 39 (9), 3819–3845.

Wilby, R., Dawson, C., Barrow, E.M., 2002. SDSM—a decision support tool for the assessment of regional climate change impacts. Environ. Model. Softw. 17, 145–157. https://doi.org/10.1016/S1364-8152(01)00060-3.

Wilks, D.S., Wilby, R.L., 1999. The weather generation game: A review of stochastic weather models. Prog. Phys. Geogr. 23, 329–357. https://doi.org/10.1177/030913339902300302.

Yoo A.B., Jette M.A., Grondona M. (2003). SLURM: Simple Linux Utility for Resource Management. In: Feitelson D., Rudolph L., Schwiegelshohn U. (eds) Job Scheduling Strategies for Parallel Processing. JSSPP 2003. Lecture Notes in Computer Science, vol 2862. Springer, Berlin, Heidelberg. doi: 10.1007/10968987_3.

Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., Yabu, S., Yoshimura, H., Shindo, E., Mizuta, R., Obata, A., Adachi, Y., Ishii, M., 2019. The meteorological research institute earth system model version 2.0, MRI-ESM2. 0: description and basic evaluation of the physical component. Journal of the Meteorological Society of Japan. Ser. 97 (5), 931–965.

Zorita, E., von Storch, H., 1999. The analog method as a simple statistical downscaling technique: comparison with more complicated methods. J. Clim. 12, 2474–2489. https://doi.org/10.1175/1520-0442(1999)012<2474:TAMAAS>2.0.CO;2.

Further reading

Daohua, B.i., Martin, D., Simon, M., Siobhan, O., Arnold, S., Roger, B., Rachel, L., Ian, H., Jhan, S., Rashid, H.A., Peter, D., Chloe, M., Hailin, Y., Anthony, H., Abhishek, S., Boeira, D.F., Matthew, W., Russell, F., Aidan, H., 2020. Configuration and spin-up of ACCESS-CM2, the new generation Australian Community Climate and Earth System Simulator Coupled Model. Journal of Southern Hemisphere Earth Systems Science 70, 225–251. https://doi.org/10.1071/ES19040.

Dufresne, JL., Foujols, MA., Denvil, S. et al. Climate change projections using the IPSL-CM5 Earth System Model(2013). From CMIP3 to CMIP5. Clim Dyn **40,** 2123–2165. doi: 10.1007/s00382-012-1636-1.

Giorgetta, M.A., Jungclaus, J., Reick, C.H., Legutke, S., Bader, J., Böttinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H.-D., Ilyina, T., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Mueller, W., Notz, D., Pithan, F., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K.D., Stockhause, M., Timmreck, C., Wegner, J., Widmann, H., Wieners, K.-H., Claussen, M., Marotzke, J., Stevens, B., 2013. Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project Phase 5. J. Adv. Model. Earth Syst. 5 (3), 572–597.

Hernanz, A., Correa, C., García-Valero, J. A., Domínguez, M., Rodríguez-Guisado, E., & Rodríguez-Camino, E. (2022c). Statistical downscaling in the Tropics and Mid-latitudes: a comparative assessment for generating regional information on climate change. International Journal of Climatology. Submitted.

Sánchez-García, E., Abia, I., Domínguez, M., Voces, J., Sánchez, J.C., Navascués, B., Rodríguez-Camino, E., Garrido, M.N., García, M.C., Pastor, F., Dimas, M., Barranco, L., Ruiz del Portal, C., 2022. Upgrade of a climate service tailored to water reservoirs management. Clim. Serv. 25, 100281 https://doi.org/10.1016/j.cliser.2021.10028.

Switanek, M.B., Troch, P.A., Castro, C.L., Leuprecht, A., Chang, H.-I., Mukherjee, R., Demaria, E.M.C., 2017. Scaled distribution mapping: a bias correction method that preserves raw climate model projected changes. Hydrol. Earth Syst. Sci. 21, 2649–2666. https://doi.org/10.5194/hess-21-2649-2017.

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, doi: 10.5194/gmd-12-2727-2019.

Vovk, V., 2013. Kernel ridge regression. In: Schölkopf, B., Luo, Z., Vovk, V. (Eds.), Empirical Inference. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 105–116.

Wilby, R., Charles, S., Zorita, E., Timbal, B., Whetton, P. and Mearns, L. (2004). Guidelines for Use of Climate Scenarios Developed from Statistical Downscaling Methods. Supporting material of the Intergovernmental Panel on Climate Change.