



OPEN

CLIMBra - Climate Change Dataset for Brazil

DATA DESCRIPTOR

André Simões Ballarin¹✉, Jullian Souza Sone¹, Gabriela Chiquito Gesualdo¹, Dimaghi Schwambach¹, Alan Reis¹, André Almagro² & Edson Cezar Wendland¹

General Circulation and Earth System Models are the most advanced tools for investigating climate responses to future scenarios of greenhouse gas emissions, playing the role of projecting the climate throughout the century. Nevertheless, climate projections are model-dependent and may show systematic biases, requiring a bias correction for any further application. Here, we provide a dataset based on an ensemble of 19 bias-corrected CMIP6 climate models projections for the Brazilian territory based on the SSP2-4.5 and SSP5-8.5 scenarios. We used the Quantile Delta Mapping approach to bias-correct daily time-series of precipitation, maximum and minimum temperature, solar net radiation, near-surface wind speed, and relative humidity. The bias-corrected dataset is available for both historical (1980–2013) and future (2015–2100) simulations at a $0.25^\circ \times 0.25^\circ$ spatial resolution. Besides the gridded product, we provide area-averaged projections for 735 catchments included in the Catchments Attributes for Brazil (CABra) dataset. The dataset provides important variables commonly used in environmental and hydroclimatological studies, paving the way for the development of high-quality research on climate change impacts in Brazil.

Background & Summary

General Circulation and Earth System Models (GCMs/ESMs) play an important role in simulating the physics and dynamics of the Earth system, as well as in assessing and understanding projected changes in the global climate^{1,2}. The employment of climate models along with observed meteorological data is key to inform policy and decision-making based on hydroclimatic modeling^{3–6}. Nevertheless, to provide a large volume of climatic data at a global scale, GCMs/ESMs often present (i) coarse spatial resolution (100–300 km), hampering the development of reliable and detailed studies at finer scales⁷ and (ii) systematic biases, leading to misrepresentation of different statistical properties of observed climate variables^{8,9}. Apart from these two limitations, these climate models have intrinsic uncertainties that undermine studies on climate change impacts¹⁰. In this context, statistical and dynamical downscaling approaches are used to bridge the inherent gap between projected and observed data by improving the spatial resolution of GCMs products to a finer scale. Dynamical downscaling is based on the integration of Regional Climate Models (RCMs) — which are able to capture local features and dynamics to better represent the climate of a specific and limited area — with the initial and lateral boundary conditions derived from GCMs^{11,12}. Despite the improvement in GCMs resolution and the incorporation of local scale-effects into their projections, RCMs may still contain systematic biases that are propagated through the future simulations. These remaining biases, derived from the used GCMs or incorporated during the dynamical downscaling procedure due to limited understanding of the processes, may lead to the need of further corrections of the dynamically downscaled product^{13–15} at a large computational effort¹⁶. On the other hand, the statistical downscaling is based on statistical transfer functions that adjust the probability distribution function of projections to resemble observed data at local/regional sites¹⁷. The statistical framework is able to correct systematic biases found in the original projections of both RCMs (generally referred to as “hybrid downscaling”^{18,19}) and GCMs using simple modelling structures that require less computational effort^{20,21}. Given these advantages, the statistical approach is generally preferred over the dynamical for climate-based studies^{3,22}.

Despite the importance of a dataset with historical and future data for meteorological and hydrological studies, to our knowledge there is no dataset of bias-corrected climate change data available for the Brazilian territory based on the recently released Sixth Assessment Report (AR6) of the Coupled Model Intercomparison Project

¹Department of Hydraulics and Sanitation, São Carlos School of Engineering, University of São Paulo, CxP. 359, São Carlos, São Paulo, 13566-590, Brazil. ²Faculty of Engineering, Architecture and Urbanism, and Geography, Federal University of Mato Grosso Do Sul, CxP 549, Campo Grande, Mato Grosso Do Sul, 79070-900, Brazil. ✉e-mail: asballarin@gmail.com; andre.ballarin@usp.br

Model	Country/Region	Resolution	Reference
MRI-ESM2*	Japan	1.12° × 1.12°	60
EC-EARTH3*	Europe	0.7° × 0.7°	61
CMCC-ESM2*	Europe	0.9° × 1.25°	62
INM-CM4-8*	Russia	1.5° × 2.0°	63
NorESM2-MM*	Norway	0.9° × 1.25°	64
MPI-ESM1.2-HR*	Germany	0.9° × 0.9°	65,66
INM-CM5*	Russia	1.5° × 2.0°	67
ACCESS-ESM1-5*	Australia	1.87° × 1.25°	68,69
TaiESM1	Taiwan	1.9° × 1.25°	70
NESM3	China	1.9° × 1.9°	71
KIOST-ESM	South Korea	1.87° × 1.87°	72
K-ACE	South Korea	1.87° × 1.25°	73
GFDL-CM4	USA	1.0° × 1.25°	74
GFDL-ESM4	USA	1.0° × 1.25°	75
ACCESS-CM2	Australia	1.87° × 1.25°	76
HadGEM3-GC31-LL	UK	1.87° × 1.25°	77
IPSL-CM6A*	France	2.5° × 1.3°	78
UKESM1.0	UK	1.87° × 1.25°	79
MIROC6*	Japan	1.4° × 1.4°	80

Table 1. CMIP6-GCMs\ESMs used in our dataset. *Simulations available for all evaluated variables.

phase 6 (CMIP6). Brazil is a continental country with diverse hydroclimatic conditions, and climate-related hazards have become more frequent, widespread, and interconnected^{23–28}. Thereby, here we developed the Climate Change Dataset for Brazil - CLIMBra, a bias-corrected dataset comprising six important meteorological variables used in hydro-climatic and economic studies related to climate change: precipitation (*pr*), maximum (*tasmax*) and minimum temperature (*tasmin*), net shortwave surface radiation (*rss*), near-surface wind speed (*sfcWind*) and relative humidity (*hur*). Besides the bias-corrected gridded daily data at a spatial resolution of 0.25°, we provide an area-averaged, point-based scale data for 735 catchments of the Catchments Attribute for Brazil (CABra) large-sample dataset²⁹. The developed dataset consists of bias-corrected historical (1980–2013) and future (2015–2100) simulations of 19 GCMs/ESMs, forced by the CMIP6 SSP2-4.5 and SSP5-8.5 scenarios. Our product is of paramount importance to support climate change assessment in the context of the water-energy-food nexus and hence inform policy and decision-making.

Methods

Datasets. To generate the bias-corrected product, we used observed data covering the historical period (1980–2013) and simulated data covering both historical and future (2015–2100) periods. As observed data, we adopted the meteorological dataset developed by Xavier *et al.*³⁰, which includes gridded daily series with 0.25° × 0.25° spatial resolution for the six meteorological variables evaluated in this study. This dataset uses data from 3,625 ground-based rain gauges and 735 weather stations provided by the National Institute of Meteorology (INMET), National Water and Sanitation Agency (ANA), and the Water and Electric Power Department of São Paulo (DAEE/SP) as input to produce interpolations. Six different interpolation techniques were evaluated: arithmetic averaging, thin plate spline, natural neighbor, inverse distance weighting, angular distance weighting, and ordinary point kriging. In general, the dataset shows a good performance in describing the weather station observations and is widely applied in hydrological and climatological studies in Brazil^{31,32}, and also as the ground truth to GCMs/RCMs projections assessment and impact studies^{33,34}.

For historical and future projections, we used daily data from 19 CMIP6 GCMs\ESMs (Table 1). The models were selected based on the following criteria: (1) availability of daily data for the evaluated climate variables under the r1i1p1f1 variant, and (2) nominal spatial resolution up to 250 km. For relative humidity (*hur*), wind speed (*sfcWind*), and net solar radiation (*rss*), only 10 GCMs were available at the stipulated conditions. For future projected changes, we considered two Shared Socioeconomic Pathways: the middle of the road (SSP2-4.5) and the fossil-fueled development (SSP5-8.5). The latter scenario represents the high end of future pathways with enough emissions to achieve a radiative forcing of 8.5 W.m⁻² by 2100, whilst SSP2-4.5 represents the medium part of the range of future pathways. The two scenarios are an update of the previous Representative Concentration Pathways (RCPs) from the CMIP5^{35,36}. The SSPs include mitigation and adaptation efforts based on economic and social changes, such as societal-economic development³⁷. There are two underlying reasons behind considering these scenarios: (1) since they represent the intermediate (SSP2-4.5) and extreme (SSP5-8.5) future climate change conditions, they are able to represent a wide range of expected changes in global climate dynamics, encompassing other scenarios, such as the SSP3-7.0, which were not available for all evaluated CMIP6 GCMs/ESMs at the time of the pre and post-processing tasks; (2) they are the most used future scenarios in Brazilian climate change studies^{12,33,34,38}. The main climatological institute of the country, the National Institute for Space Research (INPE), provides RCM-simulated future data for the country considering the two

forcing-equivalent CMIP5 RCPs scenarios (RCP4.5 and RCP8.5)¹¹. Therefore, the use of these scenarios here may enable future CMIP5-CMIP6 comparison studies.

Pre and post-processing. Given the coarse and different spatial resolutions of the CMIP6 GCMs/ESMs, we performed a bilinear interpolation following previous studies (e.g.^{3,6,10,39}) to regrid all models to a common 0.25° spatial grid. Moreover, to obtain a multi-model ensemble seeking to encompass different representations and uncertainties of all evaluated models, the GCMs outputs need to be in the same spatial grid resolution^{40,41}. It is worth mention that regridding may severely impact the statistical properties of meteorological variables, especially those linked to extreme events⁴². Hence, such limitation should be considered in future studies using our dataset.

Besides the coarse spatial resolution exhibited by GCM/ESMs, they also show an inherent inability to simulate the present-day climate conditions leading to systematic errors that are propagated for future simulations⁴³. Thus, climate products often require bias correction. Here, we used the Quantile Delta Mapping (QDM) approach²⁰ since it explicitly preserves relative or absolute changes in quantiles between historical and future simulations⁴⁴. This method is based on two widely used correction procedures: the quantile delta change and the detrend quantile mapping^{45,46}. According to Cannon *et al.*²⁰, the QDM can be performed in three steps. First, all the individual future projected quantiles are detrended. Then, the detrended quantiles are bias-corrected using the quantile mapping technique. Lastly, the projected changes are then superimposed on the bias-corrected outputs. Let denote o and p as observed and projected data, and h and f as historical and future periods, respectively. The definition of the non-exceedance probability of observed ($x_{h,o}$) and projected historical ($x_{h,p}$) data, and future data ($x_{f,p}$) are accounted as:

$$p_{f,p}(t) = F(x_{f,p}(t)) \quad (1)$$

$$p_{h,p}(t) = F(x_{h,p}(t))$$

$$p_{h,o}(t) = F(x_{h,o}(t))$$

where p and F respectively denote the non-exceedance probability associated with a specific value at time t and the empirical cumulative distribution function (ECDF). We adopted the non-parametric probability distributions as they showed better performance over the parametric ones when downscaling both RCMs and GCMs/ESMs outputs¹⁷. Also, it is easier to apply it to different meteorological variables, despite their different underlying distributions³. Then, we computed a change factor (Eq. 2), which associates the historical simulation output with that of the future period:

$$\Delta^M(t) = \frac{F_{f,p}^{-1}(p_{f,p}(t))}{F_{h,p}^{-1}(p_{f,p}(t))} = \frac{x_{f,p}(t)}{F_{h,p}^{-1}(p_{f,p}(t))} \quad (2)$$

$$\Delta^A(t) = F_{f,p}^{-1}(p_{f,p}(t)) - F_{h,p}^{-1}(p_{f,p}(t)) = x_{f,p}(t) - F_{h,p}^{-1}(p_{f,p}(t))$$

where F^{-1} indicates the inverse ECDF; and $\Delta^M(t)$ and $\Delta^A(t)$ are respectively the multiplicative and the additive change factor between simulated quantiles of the historical and future periods. The former is suitable for precipitation variables preserving relative changes between quantiles, whereas the latter preserves absolute changes in projected quantiles, suitable for temperature-derived variables^{20,44}. For precipitation, a frequency adaptation suggested by Themeßl *et al.*⁴⁷ was also applied in order to account for a methodological problem that arises when the frequency of modelled dry days is greater than the frequency of observed dry days, resulting in a systematic wet bias. Moreover, we used a wet-day threshold of 1 mm/day following previous studies^{3,17,48} to minimize drizzle effects.

The bias-corrected projected data are computed according to Eq. 3, following the (i) quantile-mapping technique, which statistically transforms the distribution of the projected data to resemble the distribution of the observed data; and (ii) the change factor approach, which superimposes relative (or absolute) changes between historical and future projected values:

$$\hat{x}_{f,p}(t) = \Delta^M(t) \cdot F_{h,o}^{-1}(p_{f,p}(t)) \quad (3)$$

$$\hat{x}_{f,p}(t) = \Delta^A(t) + F_{h,o}^{-1}(p_{f,p}(t))$$

We considered the entire observed period (1980–2013) to correct the future projections (2015–2100) with the QDM algorithm. Both pre and post-processing tasks, depicted in Fig. 1, were carried out using the *downscaleR* package, which is an R-based framework developed to address the needs of different climate impact studies within the *Climate4R* project^{18,49}. Despite the wide range of applications of QDM in climate change impact assessments, we must point out the main limitations and uncertainties related to its usage. In general, the QDM (1) shows high sensitivity to the historical reference data used for calibration and (2) may be affected by the downscaling process due to a resolution mismatch between the model simulations and observations⁵⁰. To reduce

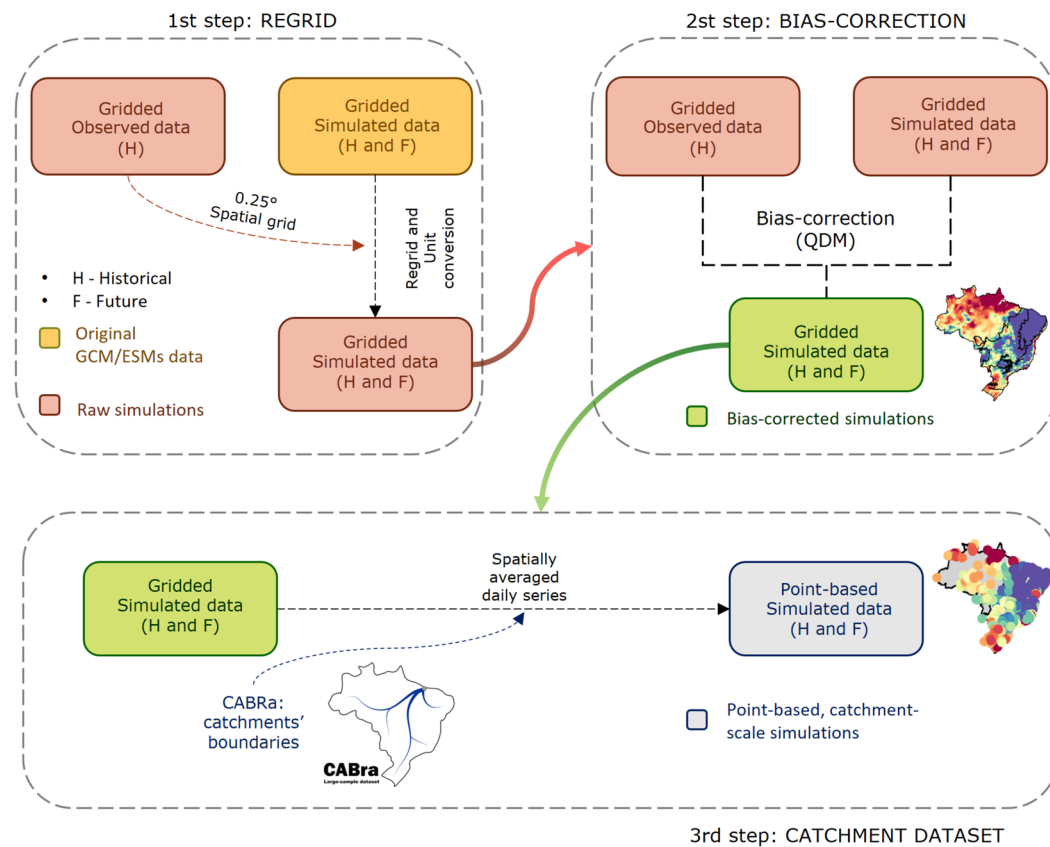


Fig. 1 Flowchart representing the core steps used to generate the CLIMBra's products. Step 1 and 2 represent the regrid and bias-correction tasks, respectively. Step 3 represents the framework required to rescaled the gridded dataset to the CABra's catchments.

this uncertainty, high-quality reference data are required. Therefore, we used a high-resolution meteorological gridded dataset³⁰ that comprises the largest number of ground-based observations in Brazil. We also highlight our effort to mitigate the introduction of more uncertainty and overcome some methodological limitations for the development of a reliable product for climate change assessment in this country. Nevertheless, it is known that the use of bias-correction methods is controversial⁵¹ as they are not able to retain spatial and intervariable dependencies⁵². Lastly, bias-corrections methods may produce physically unrealistic values and also hind some fundamental models deficiencies^{50,51}.

Catchment-scale dataset. The development of a regional gridded meteorological dataset aims at advancing hydro-climate studies in Brazil. Nevertheless, working with gridded climate variables is not a trivial task, often requiring high computation effort. Therefore, we also developed a catchment-scale version of the dataset in order to assist climate-change impact studies/applications (Fig. 1) for both scientific and technical fields. To this end, we rescaled our gridded dataset to match the catchments in the CABra large-sample dataset²⁹. CABra includes a set of more than 100 observed climate, hydrological, and physiographic attributes for 735 Brazilian catchments (Fig. 2).

For each CABra's catchment, we generated a gridded meteorological daily series comprising its extension. In sequence, we spatially averaged the grid-cell time series within the catchment boundaries. This process was conducted for each variable and climate change scenario. Given the continental extensions of Brazil, the country shows a large range of catchment areas. Thus, for some catchments, the daily time series were obtained using only one grid and, for others, averaging more than 50 grids. Moreover, our gridded dataset only comprises the Brazilian territory, and hence the averaged daily time-series only accounted for the catchments' area within the country boundaries. To compute the spatial averaged daily time series, we weighted the grids according to their latitude, as, for regular grids, the grid cell's area changes as you move towards the poles. It is also worth noting that the averaged time-series may hinder extreme events in large catchments since extreme high or low events recorded in a specific grid may be smoothed by non-extreme events recorded in neighboring grids when taking the average.

Data Records

The present study describes a gridded dataset ($0.25^\circ \times 0.25^\circ$) and a spatially averaged dataset at a catchment-scale. The first includes raw and bias-corrected gridded (*netCDF*) daily time-series (*pr*, *tasmax*, *tasmin*, *rss**, *sfcWind**, and *hur**) of 19 (10*) CMIP6 GCMs/ESMs (Table 1) for both the historical (1980–2013) and future (2015–2100) periods. For the future period, two CMIP6 SSP-scenarios were considered: SSP2-4.5 and

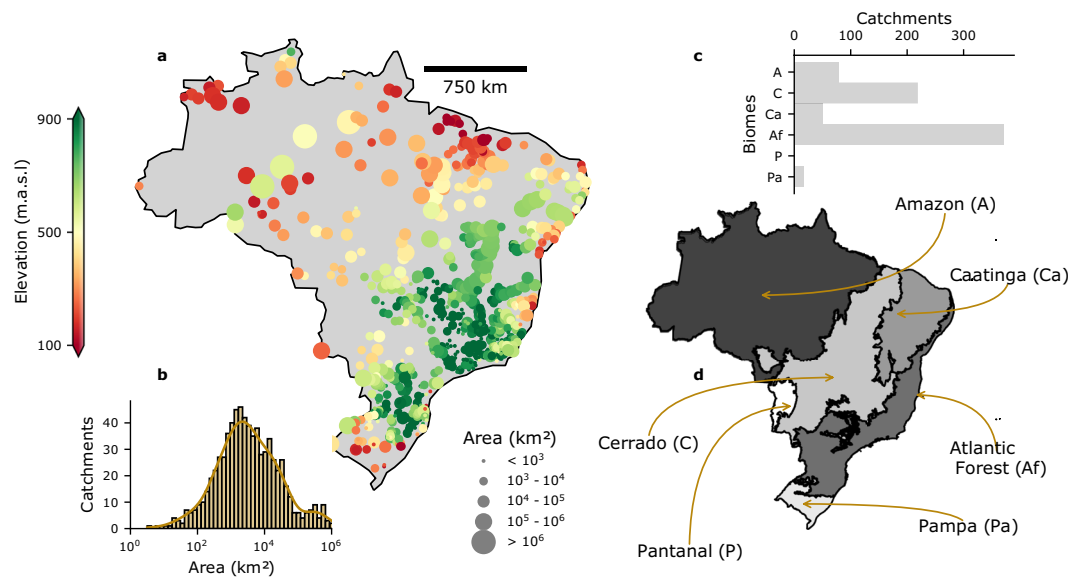


Fig. 2 (a) Streamflow gauge coordinates of CABra's catchments, colored according to their mean elevation and sized by their area. (b) Histogram of catchments' area. (c) Distribution of catchments per Brazilian biome. (d) Six main Brazilian biomes.

SSP5-8.5. The second part of the dataset consists of a point-based (.csv) daily time-series derived from the aforementioned bias-corrected gridded dataset for 735 Brazilian catchments in the CABra dataset²⁹. The time-series from this dataset were also generated considering the two SSP scenarios simulated by the 19 (10) CMIP6 GCMs\ ESMs. Both datasets are freely available (CC0 license) at Ballarin *et al.*⁵³.

Technical Validation

Bias-correction performance. *Mean and extreme values.* To get a preliminary overview of the bias-correction performance in describing the observed data, we computed the relative bias between the estimated and observed long-term means for the historical period (Fig. 3). We used absolute bias for the variables *tasmax* and *tasmin* since the denominator approaches to zero in many catchments, resulting in very high relative bias values. For both raw and bias corrected datasets, we estimated the long-term mean for each individual model and then computed the multi-model ensemble. We repeated this process for all validation analyses of the study. The bias correction significantly improved the performance of the historical simulations in describing observed long-term mean values for all the meteorological variables by reducing the bias to approximately zero.

We found the largest bias for *sfcWind* and *hur* when analyzing the ensemble data prior to bias correction. For *sfcWind*, overestimations of more than 30% were observed throughout the country. The same was noted in *hur*, but with the opposite signal. In general, the raw simulations were unable to capture long-term mean maximum and minimum temperatures. Similar to the findings of Mishra *et al.*³, the raw multi-model ensemble exhibited an overall cold bias (-1.70°C) for *tasmax* and a warm bias (0.87°C) for *tasmin*. Regarding the long-term mean precipitation, we observed an overall wet relative bias over the country (11.74%), except in the Amazon and Pampa biomes (northern and southern of Brazil, respectively) where we noted a relative dry bias. In general, the CMIP6 models exhibited good performance in estimating the long-term mean *rss*, showing a smaller mean bias (-4.34%) compared with the deviations found for the other variables.

In the remainder of this section, we discuss the dataset's performance considering our catchment-scale product, which is the main product of this study, since it provides ready-for-use meteorological time-series required for most hydrological studies and climate change impacts simulation. Nevertheless, similar conclusions can be drawn for the gridded dataset, which was the basis for the development of the catchment-scale product. To explore the models' performance beyond the characterization of mean values, we also computed the bias in the variables' long-term extreme properties: 90th (and/or 10th) percentiles and maximum (and/or minimum) records (Figs. 4, 5). It is worth mentioning that the spatial distribution of the bias in the long-term mean for the gridded (Fig. 3) and catchment-scale datasets is alike (the first column in Figs. 4, 5, plots a, d, and g), corroborating their equivalence in terms of performance. That is, although we spatially averaged the gridded dataset, the catchment-scale product is able to maintain the spatial distribution of relative bias.

In general, the CMIP6 models exhibited poor performance in simulating extreme values. The raw multi-model ensemble showed an overall warm bias (overestimation) for the long-term 10th percentile of *tasmin* and a cold bias (underestimation) for the long-term 90th percentile of *tasmax*. For the long-term maximum *tasmax* and minimum *tasmin*, we found an overall warm bias with a more heterogeneous spatial distribution than those observed in the long-term extreme percentiles. Interestingly, *tasmax* and *tasmin* showed a contrasting spatial distribution of bias respectively for the maximum and minimum values. This highlights that the raw output of the CMIP6 models fails to capture both extreme values and the temperature's amplitude.

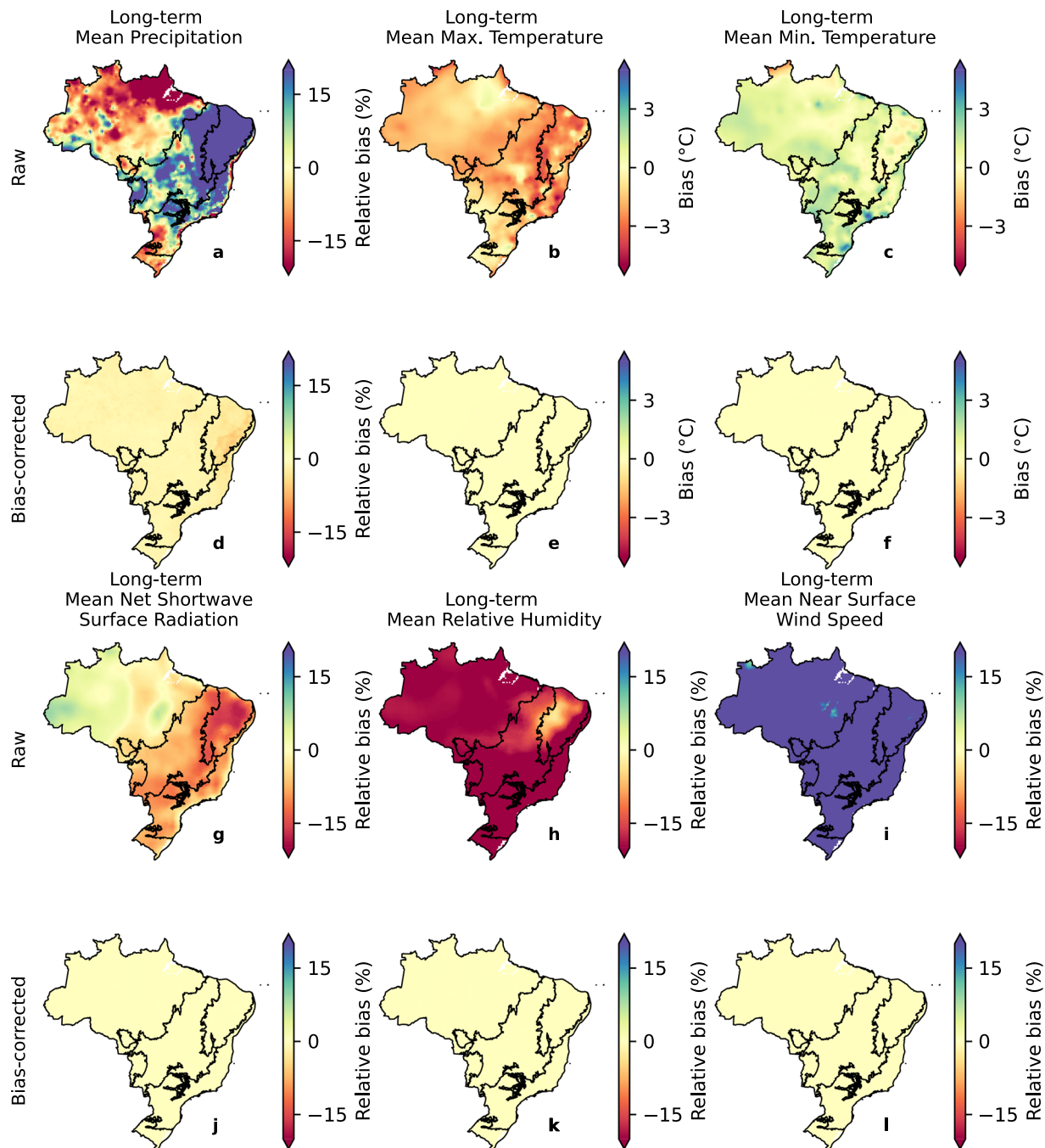


Fig. 3 Biases in long-term mean precipitation, maximum and minimum temperature, net shortwave surface radiation, relative humidity, and near surface wind speed considering the gridded dataset in both raw and bias-corrected conditions. The limits of Brazilian biomes are indicated in black borderlines.

Raw simulations of precipitation (*pr*) presented an absolute relative bias of up to 60% and 40% for the long-term 90th percentile and maximum values, respectively (Fig. 4), corroborating with the findings of Pereira *et al.*⁵⁴. Similar to *tasmax* and *tasmin*, the spatial distribution of the precipitation's relative biases in the long-term extreme percentile was alike to those found for the long-term mean: dry biases in the Amazon and Pampa biomes and an overall wet bias in the rest of the country. The long-term maximum precipitation did not show a clear spatial pattern. Such as observed in the mean long-term, the climate models were not able to characterize the extreme values of *hur* and *sfcWind*. For the former, underestimations of more than 20% and 50% were respectively found for the 90th and 10th percentiles. For the latter, the errors were even more significant, with relative biases of around 80% and 100% for the 90th and 10th percentiles, respectively. Again, the simulations of *rss* exhibited the best performance with relative biases of nearly 10% of magnitude and a tendency to slightly underestimate the observations.

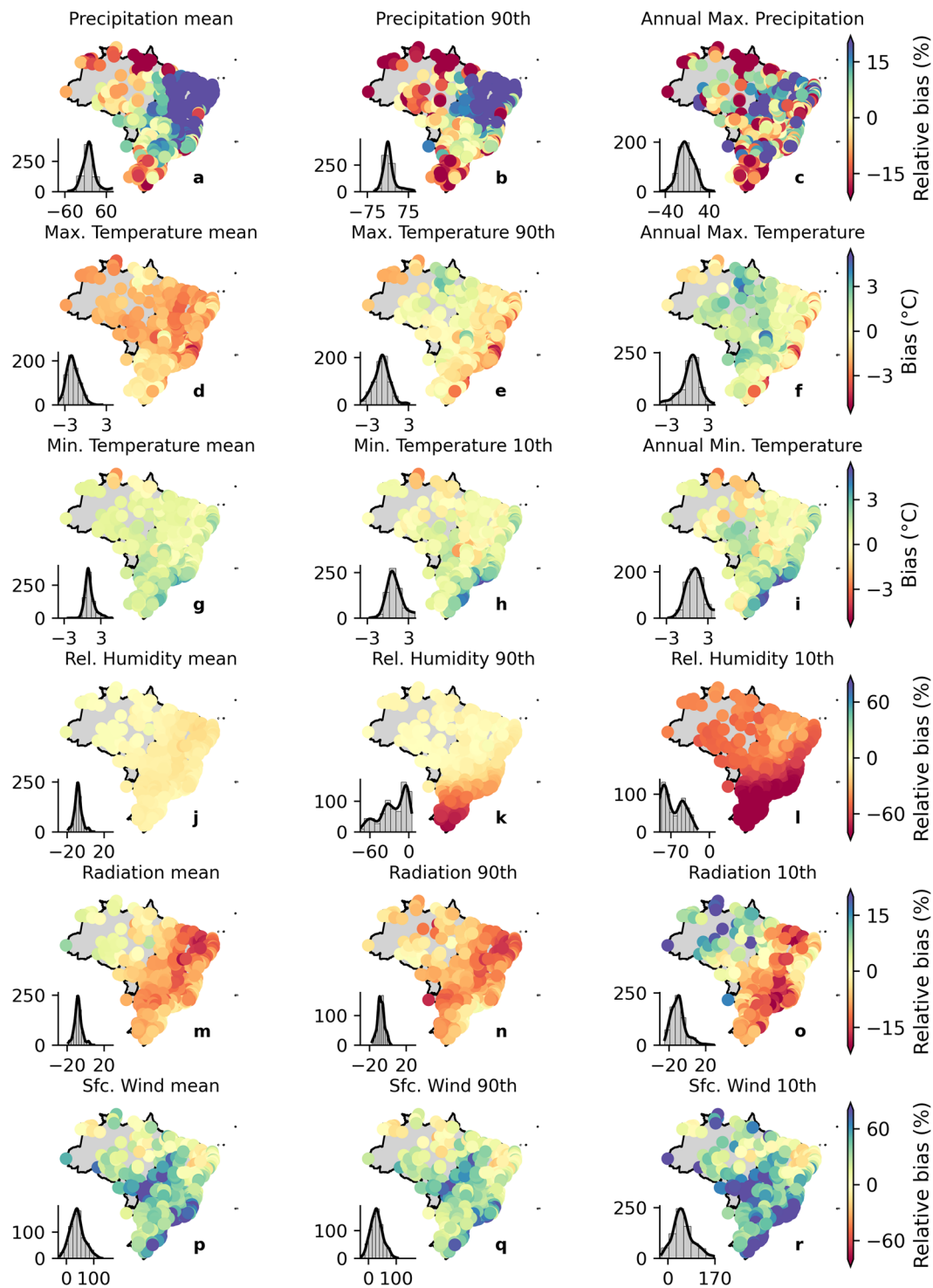


Fig. 4 Biases in the long-term mean and extreme values of precipitation, maximum and minimum temperature, net shortwave surface radiation, relative humidity, and near surface wind speed (catchment-scale dataset) for the raw simulations. Histograms in each of the panels indicate the frequency of occurrence of bias.

It is clear that the bias correction procedure significantly improved the estimations (Fig. 5), successfully removing most of the bias in both long-term mean and extreme values for almost all variables, except for the long-term maximum precipitation, where the improvement was not so perceptible. After the correction, the bias reduced to nearly 0% in all catchments (see histograms in Fig. 5). This is true even for *hur* and *sfcWind*, which were significantly misrepresented by the raw GCMs/ESMs. These results indicate that the QDM was able to overcome one of the main limitations of commonly used bias-correction methods: correcting systematic errors

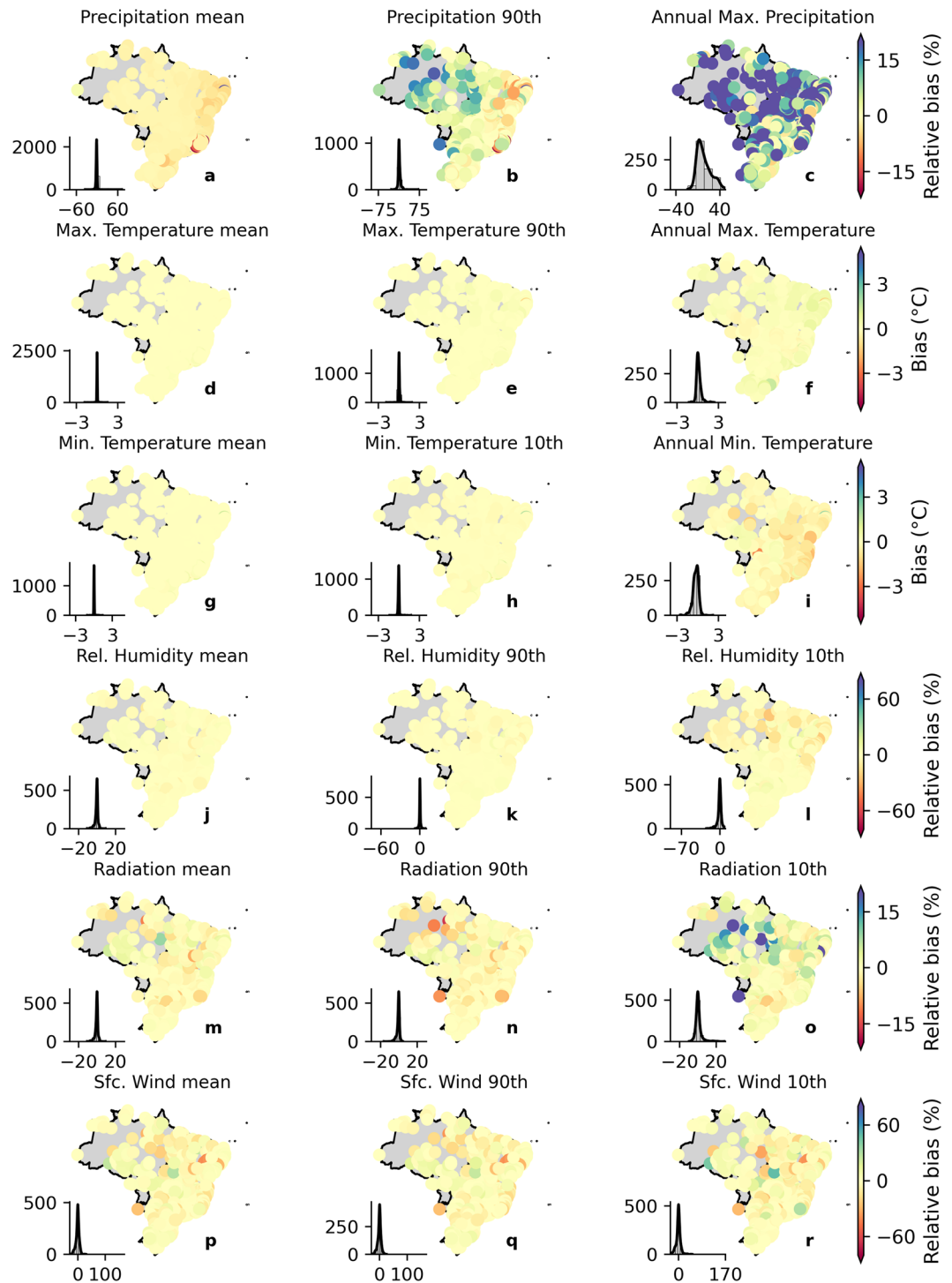


Fig. 5 Biases in the long-term mean and extreme values of precipitation, maximum and minimum temperature, net shortwave surface radiation, relative humidity, and near surface wind speed (catchment-scale dataset) for the bias-corrected simulations. Histograms in each of the panels indicate the frequency of occurrence of bias.

in different quantiles of the probability distributions of raw simulations, such as biases present in the mean and in the tail of the GCMs probability distribution⁵².

Seasonality. We assessed the intra-annual variability of *pr*, *tasmax*, and *tasmin* in each Brazilian biome (Figs. 6, 7 for raw and bias-corrected datasets, respectively). The intra-annual performance of the other variables is available in the accompanying Supplementary Material (Figures S1, S2 – Supplementary File 1) since these variables are not available in all 19 CMIP6 climate models. This investigation seeks to confirm if the simulations are able

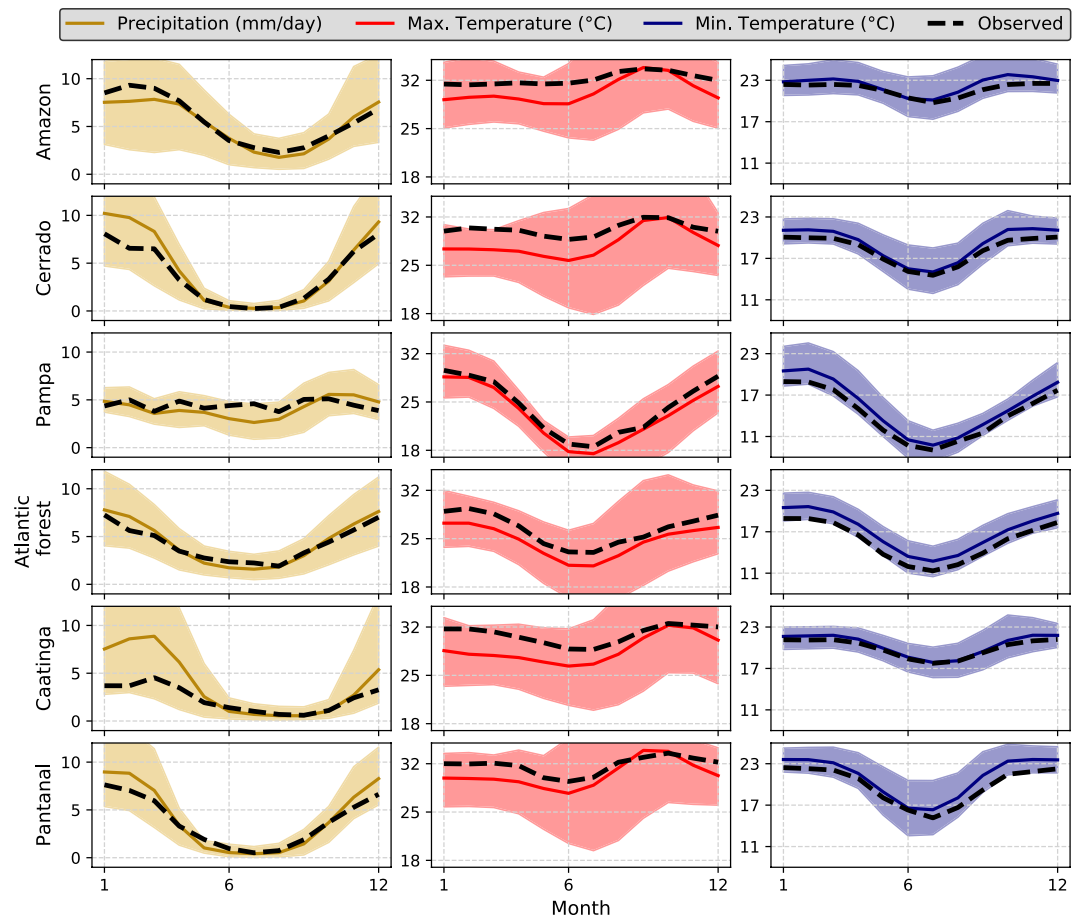


Fig. 6 Long-term (1980–2013) monthly mean of precipitation and maximum and minimum temperature in each Brazilian biome. Highlighted lines represent the intra-annual cycle simulated by the raw *multi-model ensemble*. Dashed lines indicate the observed mean intra-annual cycle. Confidence intervals represent the maximum and minimum values simulated by the raw 19 CMIP6 GCMs/ESMs.

to reproduce the intra-annual cycle of the evaluated variables, an important aspect to assess especially for hydrological modelling purposes. Overall, both the raw and bias products exhibited good performance in reproducing the seasonal cycle of *pr*, *tasmax*, and *tasmin* (Figs. 6, 7). The confidence intervals, defined by the minimum and maximum values found in the 19 GCMs/ESMs, encompassed the observed pattern of the three variables in all biomes.

The smallest uncertainties, indicated by the confidence intervals, were found for *tasmin* followed by *pr* and *tasmax*. The models showed excellent performance in reproducing the intra-annual variability of *tasmin*, with a narrowed confidence interval and an ensemble mean close to the observations. Despite the large hydroclimatic variability in the country, we noted a clear seasonal pattern in *pr*. Both the raw and bias-corrected datasets exhibited a larger uncertainty in the characterization of the rainy season (October to March), as indicated by wider confidence intervals and a large difference between the ensemble mean and observations. Similar results were also found by Almazroui *et al.*⁵⁵ when evaluating the performance of CMIP6 models in characterizing mean properties of rainfall and temperature in South America. In fact, more extreme rainfall events, which were significantly misrepresented by the models (Figs. 3, 4), are more likely to be experienced in the rainy season. In contrast, in the dry period, the ensemble mean approaches the observation showing also narrower confidence intervals. The exception here is the Pampa, where we did not find a clear pattern in both rainy and dry seasons since they are not well-defined in this biome³³. An opposite situation was found for *tasmax*: larger confidence intervals in the dry period when the highest temperature amplitude and, consequently, climatic variability occur in Brazil⁵⁶. In all biomes, raw models showed a slight underestimation of *tasmax*.

The bias-corrected simulations exhibited, again, excellent performance in characterizing the observed data (Fig. 7). They were able to significantly reduce the average biases approaching the observed and simulated monthly cycles, in addition to reducing the uncertainties expressed by narrower confidence intervals. Despite this, the limitations present in the raw simulations remained (Fig. 6), albeit with a less clear pattern: greater uncertainties in the simulation of *pr* and *tasmax* in the rainy and dry seasons, respectively. Similar conclusion can be drawn for *rss*, *hur*, and *sfcWind* (Figures S1, S2 - Supplementary File 1). Although not as significantly as observed for the *tasmax*, *tasmin* and *pr*, the bias correction improved the performance of the raw simulations in

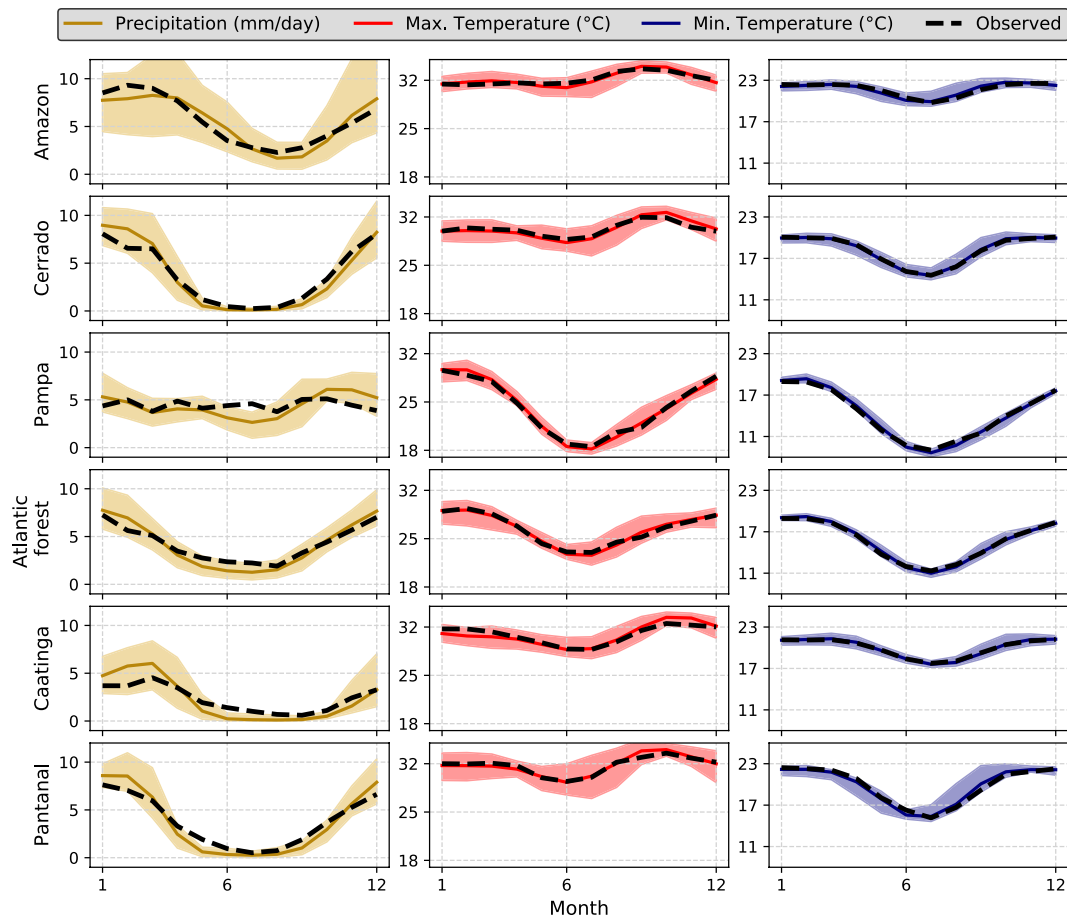


Fig. 7 Long-term (1980–2013) monthly mean of precipitation and maximum and minimum temperature in each Brazilian biome. Highlighted lines represent the intra-annual cycle simulated by the bias-corrected *multi-model ensemble*. Dashed lines indicate the observed mean intra-annual cycle. Confidence intervals represent the maximum and minimum values simulated by the bias-corrected 19 CMIP6 GCMs/ESMs.

describing the monthly cycles of these variables, which showed smaller uncertainties and better accuracy. This can be explained by the fact that the GCMs/ESMs showed a lower performance in the characterization of the seasonality of these three variables (Figure S1 – Supplementary File 1), especially *sfcWind* and *hur*.

Projected changes. Mean and extreme values. To analyze projected changes simulated by the bias-corrected CMIP6 models, we computed relative changes between the historical (1980–2013) and distant future (2070–2100) periods considering the long-term mean and extremes properties for both SSP2-4.5 and SSP5-8.5 scenarios (Figs. 8, 9, respectively). For *tasmax* and *tasmin*, we computed absolute changes (°C) to avoid extremely high values of relative changes due to denominator values close to zero. Projected changes simulated by the raw models are shown in the Supplementary Material (Figures S3, S4, respectively). It is worth noting that the projected changes were similar for both raw and corrected simulations (Fig. 8 and S3 for SSP2-4.5 and Fig. 9 and S4 for SSP5-8.5), indicating that the QDM method was capable to overcome another limitation of commonly used bias correction methods: it did not deteriorate trends and/or relative changes projected by the models, which may hamper the fully understanding of climate change effects^{57,58}.

As expected, the changes simulated by the SSP5-8.5 scenario exhibited greater magnitude than those simulated by the SSP2-4.5. Among the six evaluated variables, *tasmax*, *tasmin*, *pr*, and *sfcWind* presented significant changes (>10%) between historical and future simulations in most parts of the Brazilian territory considering both scenarios. For *hur* and *rss*, the projected changes showed a smaller magnitude (<10%). Long-term mean precipitation and the 90th percentile exhibited similar spatial patterns of change over the country. A reduction in projected long-term mean *pr* was observed in the Amazon, Caatinga, and part of the Cerrado biomes. This expected reduction also corroborates the findings of Du *et al.*⁵⁹. In the Pampa and Atlantic Forest biomes, a slight increase is expected. Similar conclusions can be drawn for the 90th percentile of precipitation. This pattern of change is more evident considering the SSP5-8.5 scenario (Fig. 9). Regarding the long-term maximum *pr*, a significant increase was observed (>10% for the SSP2-4.5 scenario and >20% for the SSP5-8.5 scenario) throughout the country, even in the biomes where a reduction in mean precipitation was projected (Figs. 8, 9).

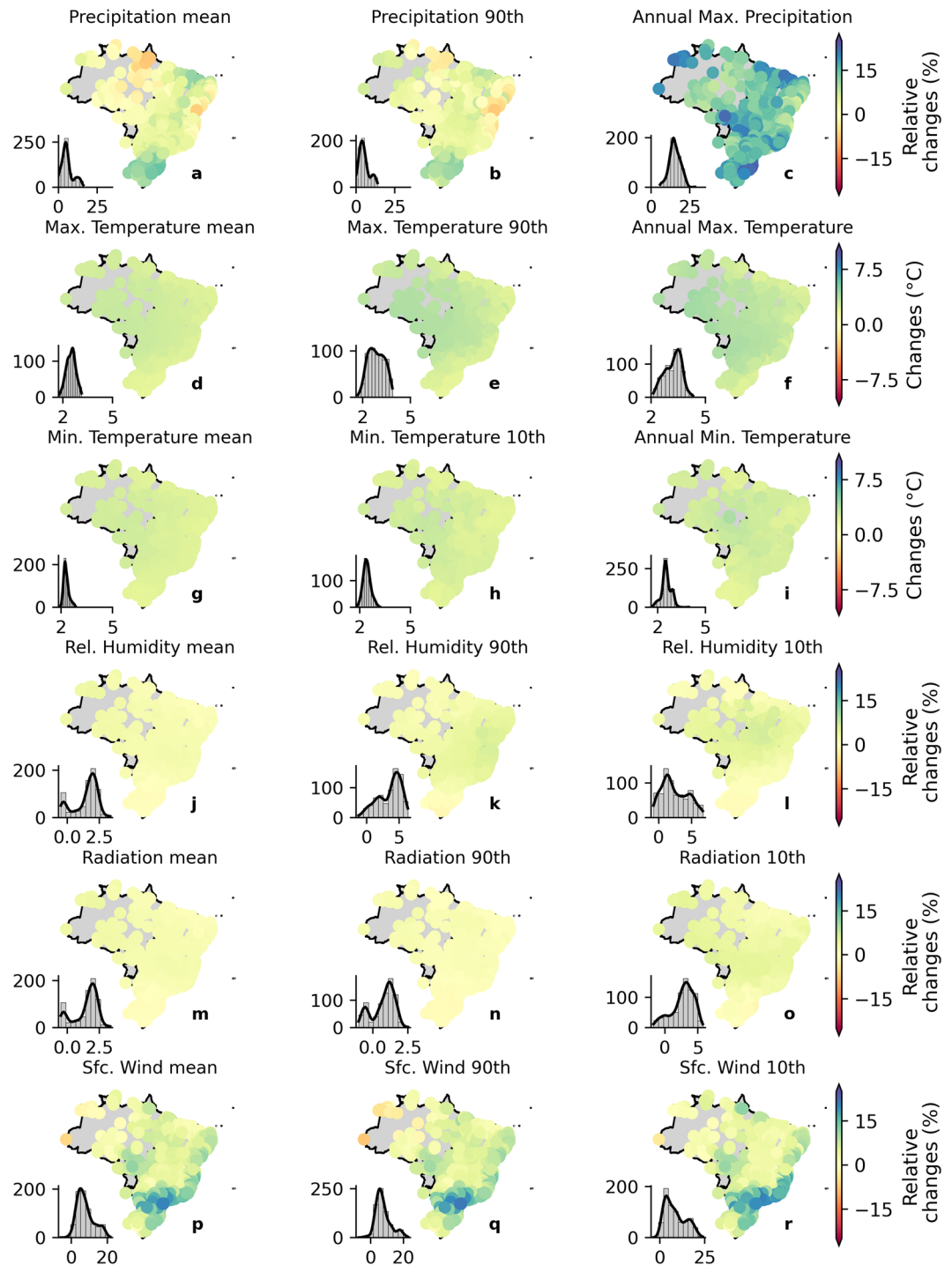


Fig. 8 Relative changes in the long-term mean and extreme values of precipitation, maximum and minimum temperature, net shortwave solar radiation, relative humidity, and near surface wind speed between the historical (1980–2013) and distant future (2070–2100; SSP2-4.5) periods (bias-corrected catchment-scale dataset). Histograms in each panel indicate the frequency of occurrence of relative changes.

In general, an increase in both *tasmax* and *tasmin* was observed in the two scenarios in Brazil. The projected increase in the long-term maximum temperature is slightly higher in the Amazon, Pantanal, and part of the Cerrado biomes than in the others. Furthermore, the positive magnitude of change in extreme values is larger than that projected for the long-term mean values. That is, both scenarios projected changes with greater magnitude in the extreme characteristics of *tasmax* (maximum and the 90th percentile) than in its mean values. This is also valid for *tasmin*, but with a lower difference between magnitudes of change.

The projected changes in *rss* and *hur* showed smaller magnitude when compared with the other variables. For the former, an average increase in the long-term mean of about 2% and 4% is projected in the SSP2-4.5 and

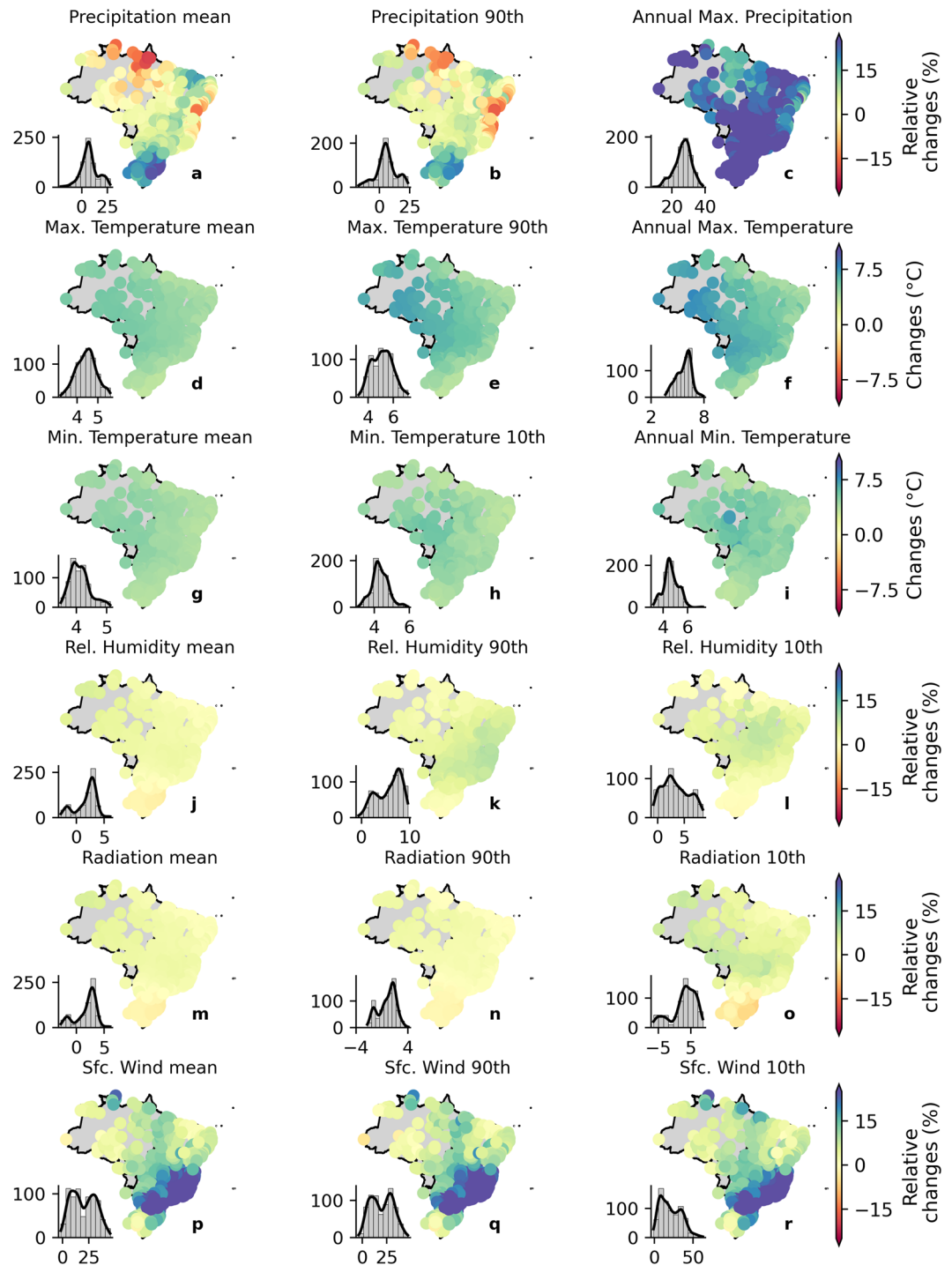


Fig. 9 Relative changes in the long-term mean and extreme values of precipitation, maximum and minimum temperature, net shortwave solar radiation, relative humidity, and near surface wind speed between the historical period (1980–2013) and the distant future (2070–2100; SSP5-8.5) (bias-corrected catchment-scale dataset). Histograms in each of the panels indicate the frequency of occurrence of relative changes.

SSP5-8.5, respectively. Unlike *pr*, *tasmax*, and *tasmin*, there is no clear difference between the expected changes in mean and extreme values of *rss*, except the 10th percentile in the SSP5-8.5 scenario that shows a negative variation in the Pampa biome and in part of the Atlantic Forest biome and a positive variation in the central region of the country. The *hur* variable also showed a similar pattern of projected changes in the long-term mean and extreme values across the country, indicating a slight increase of <2.5% in the SSP2-4.5. For the SSP5-8.5 scenario, we noted slightly larger changes (<5%) in both the mean and extreme values. Interestingly, the 10th and 90th percentiles showed an opposite spatial distribution in the SSP5-8.5 scenario, in which we found larger increases in the 90th percentile and smaller increases in the 10th percentile of *hur*. Lastly, large increases in

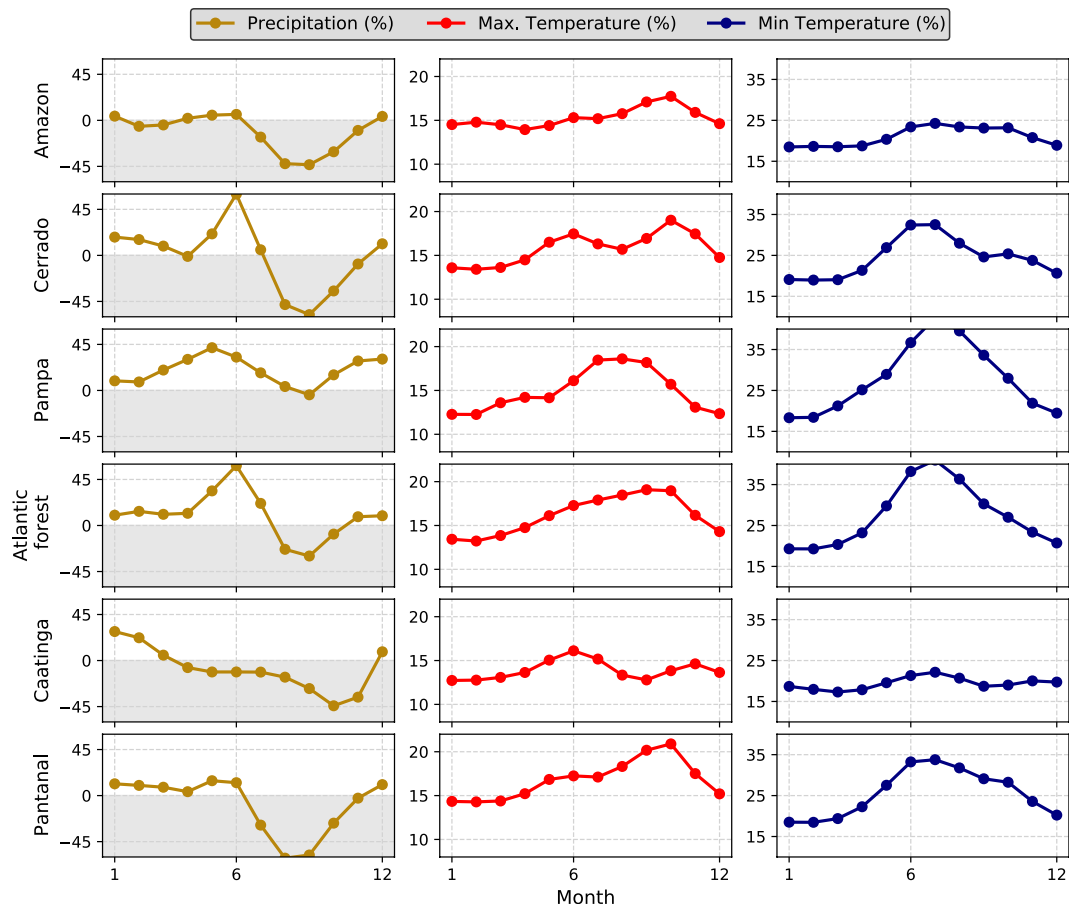


Fig. 10 Relative changes in the long-term mean intra-annual cycles of precipitation and maximum and minimum temperatures between the historical (1980–2013) and distant future (2070–2100, SSP5-8.5) periods. Highlighted lines represent the changes in the intra-annual cycle simulated by the bias-corrected *multi-model ensemble*.

the *sfcWind* of 15% in SSP2-4.5 and 25% in SSP5-8.5 are expected especially in the Pampa (South region) and Atlantic Forest biomes (Southeast region). In the other regions, smaller increases were noted. It was also not possible to distinguish different spatial patterns between the changes in mean and extreme events.

Seasonality. Significant changes in intra-annual cycles are also expected (Fig. 10). For better visualization, we only detailed the relative changes (*multi-model ensemble*) expected in the monthly averages of *pr*, *tasmax*, and *tasmin*. The results for the other three variables are provided in the Supplementary File 1 (Figure S5). We considered the bias-corrected simulations of the historical (1980–2013) and ‘distant’ future (SSP5-8.5; 2070–2100) to compute the relative changes. However, very similar changes were observed for the simulations without bias correction.

Changes in projected precipitation seasonal cycles varied among the biomes (Fig. 10). For the Atlantic Forest and Cerrado, a shift in seasonality is projected, indicated by a large increase in precipitation in the first months of the dry period (April to July) and a reduction in the last months (August and September). For the Amazon, Caatinga, and Pantanal biomes, an overall decrease in precipitation is expected, but with the maintenance of seasonal cycles. Finally, a general increase in precipitation is projected in the Pampas biome, however, with a heterogeneous distribution throughout the year: larger variations were observed at the beginning of both dry (April to June) and rainy seasons (October to December). We also found positive relative changes for *tasmin* and *tasmax*. For the former, a more significant increase is projected between May and September in all biomes, especially in the Pampas, Pantanal and Atlantic Forest biomes where the increase was more significant (up to 35%). The projected increase in *tasmax* is expected to be more uniform throughout the year than those projected for *tasmin*, being concentrated in the second half of the year, except for the Caatinga and Amazon biomes, which presented a similar rate of increase throughout the year for both *tasmin* and *tasmax*.

Usage Notes

Here, we described the CLIMBra - Climate Change Dataset for Brazil, which provides raw and bias-corrected daily time series of six meteorological variables at both gridded and catchment scales using simulations of up to 19 CMIP6 GCM/ESMs. The simulations were provided for both historical (1980–2013) and future periods (2015–2100) forced by two different emission scenarios: SSP2-4.5 and SSP5-8.5. CLIMBra products may be

useful for different hydroclimatic purposes such as hydrological modeling and climate change impact assessment. Moreover, CLIMBra may also be of interest to users not only in the hydrometeorological field but also in others such as agriculture, public health, and ecology. The high-resolution of the gridded data ($0.25^\circ \times 0.25^\circ$) is key to developing regional assessments, providing information to decision and policy-making not only in Brazil but also in South America given the Brazil's continental proportion and its role in global climate dynamics.

Our main product, the catchment-scale dataset, is provided as comma-separated values format (.csv), which is easier to handle and download in comparison with *netCDF* files. Both the gridded and catchment-scale datasets are freely available at the Science Data Bank (<https://doi.org/10.57760/sciencedb.02316>). Despite the importance of the developed product, it is important to highlight some of its limitations. (1) Data users should be aware of significant bias when using our raw database due to its weak performance in representing observations, depending on its application. Thus, a performance analysis should be conducted to investigate whether the raw historical data are able to simulate observations. (2) Even exhibiting a better performance than the raw simulations (mainly in reproducing seasonal variability and extreme properties of the evaluated variables), the bias-corrected products may present inherent uncertainties, physically unrealistic values, and hide some fundamental deficiencies presented by the climate models. Finally, (3) data users should consider that the area-averaging process used to develop the catchment-scale dataset may hinder or smooth extreme events and misrepresent transboundary catchments.

Code availability

Pre- and post-processing tasks were carried out using the R-packages of the *Climate4R* project, extensively described in Bedia *et al.*¹⁸ and Iturbide *et al.*⁴⁹. This framework was developed to address the needs of different climate-impact studies and includes a roll of R-packages to access, pre- and post-process, and visualize climate data. All the packages and documentation, including tutorials and example-notebooks, are available through the following Github link: <https://github.com/SantanderMetGroup/climate4R>.

Received: 15 September 2022; Accepted: 10 January 2023;

Published online: 20 January 2023

References

- Ukkola, A. M. *et al.* Evaluating CMIP5 Model Agreement for Multiple Drought Metrics. *J. Hydrometeorol.* **19**, 969–988, <https://doi.org/10.1175/JHM-D-17-0099.1> (2018).
- Werner, A. T. & Cannon, A. J. Hydrologic extremes - An intercomparison of multiple gridded statistical downscaling methods. *Hydrol. Earth Syst. Sci.* **20**, 1483–1508, <https://doi.org/10.5194/HESS-20-1483-2016> (2016).
- Mishra, V., Bhatia, U. & Tiwari, A. D. Bias-corrected climate projections for South Asia from Coupled Model Intercomparison Project-6. *Sci. Data* **2020 71 7**, 1–13, <https://doi.org/10.1038/s41597-020-00681-1> (2020).
- Moustakis, Y., Papalexiou, S. M., Onof, C. J. & Paschalis, A. Seasonality, Intensity, and Duration of Rainfall Extremes Change in a Warmer Climate. *Earth's Futur.* **9**, e2020EF001824, <https://doi.org/10.1029/2020EF001824> (2021).
- Ombadi, M., Nguyen, P., Sorooshian, S. & Hsu, K. lin. Developing Intensity-Duration-Frequency (IDF) Curves From Satellite-Based Precipitation: Methodology and Evaluation. *Water Resour. Res.* **54**, 7752–7766, <https://doi.org/10.1029/2018WR022929> (2018).
- Tang, G., Clark, M. P. & Papalexiou, S. M. EM-Earth: The Ensemble Meteorological Dataset for Planet Earth. *Bull. Am. Meteorol. Soc.* **103**, E996–E1018, <https://doi.org/10.1175/BAMS-D-21-0106.1> (2022).
- Gao, S., Huang, D., Du, N., Ren, C. & Yu, H. WRF ensemble dynamical downscaling of precipitation over China using different cumulus convective schemes. *AtmRe* **271**, 106116, <https://doi.org/10.1016/J.ATMOSRES.2022.106116> (2022).
- Abdelmoaty, H. M., Papalexiou, S. M., Rajulapati, C. R. & AghaKouchak, A. Biases Beyond the Mean in CMIP6 Extreme Precipitation: A Global Investigation. *Earth's Futur.* **9**, e2021EF002196, <https://doi.org/10.1029/2021EF002196> (2021).
- Papalexiou, S. M. *et al.* Probabilistic Evaluation of Drought in CMIP6 Simulations. *Earth's Futur.* **9**, e2021EF002150, <https://doi.org/10.1029/2021EF002150> (2021).
- Xu, Z., Han, Y., Tam, C. Y., Yang, Z. L. & Fu, C. Bias-corrected CMIP6 global dataset for dynamical downscaling of the historical and future climate (1979–2100). *Sci. Data* **2021 81 8**, 1–11, <https://doi.org/10.1038/s41597-021-01079-3> (2021).
- Chou, S. C. *et al.* Evaluation of the Eta Simulations Nested in Three Global Climate Models. *Am. J. Clim. Chang.* **3**, 438–454, <https://doi.org/10.4236/AJCC.2014.35039> (2014).
- Lyra, A. *et al.* Climate change projections over three metropolitan regions in Southeast Brazil using the non-hydrostatic Eta regional climate model at 5-km resolution. *Theor. Appl. Climatol.* **132**, 663–682, <https://doi.org/10.1007/S00704-017-2067-Z/TABLES/4> (2018).
- Ballarin, A. S., Barros, G. L., Cabrera, M. C. M. & Wendland, E. C. A copula-based drought assessment framework considering global simulation models. *J. Hydrol. Reg. Stud.* **38**, 100970, <https://doi.org/10.1016/J.EJRH.2021.100970> (2021).
- Kotlarski, S. *et al.* Regional climate modeling on European scales: A joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geosci. Model Dev.* **7**, 1297–1333, <https://doi.org/10.5194/GMD-7-1297-2014> (2014).
- Xu, Z. & Yang, Z. L. An Improved Dynamical Downscaling Method with GCM Bias Corrections and Its Validation with 30 Years of Climate Simulations. *J. Clim.* **25**, 6271–6286, <https://doi.org/10.1175/JCLI-D-12-00005.1> (2012).
- Simonovic, S. P., Schardong, A., Sandink, D. & Srivastav, R. A web-based tool for the development of Intensity Duration Frequency curves under changing climate. *Environ. Model. Softw.* **81**, 136–153, <https://doi.org/10.1016/J.ENVSOFT.2016.03.016> (2016).
- Gudmundsson, L., Bremnes, J. B., Haugen, J. E. & Engen-Skaugen, T. Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – A comparison of methods. *Hydrol. Earth Syst. Sci.* **16**, 3383–3390, <https://doi.org/10.5194/HESS-16-3383-2012> (2012).
- Bedia, J. *et al.* 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> (2020).
- Turco, M., Quintana-Seguí, P., Llasat, M. C., Herrera, S. & Gutiérrez, J. M. Testing MOS precipitation downscaling for ENSEMBLES regional climate models over Spain. *J. Geophys. Res. Atmos.* **116**, 18109, <https://doi.org/10.1029/2011JD016166> (2011).
- Cannon, A. J., Sobie, S. R. & Murdock, T. Q. 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> (2015).
- Mandal, S., Srivastav, R. K. & Simonovic, S. P. Use of beta regression for statistical downscaling of precipitation in the Campbell River basin, British Columbia, Canada. *J. Hydrol.* **538**, 49–62, <https://doi.org/10.1016/J.JHYDROL.2016.04.009> (2016).
- Gutmann, E. *et al.* An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resour. Res.* **50**, 7167–7186, <https://doi.org/10.1002/2014WR015559> (2014).
- Dalagnol, R. *et al.* Extreme rainfall and its impacts in the Brazilian Minas Gerais state in January 2020: Can we blame climate change? *Clim. Resil. Sustain.* **1**, e15, <https://doi.org/10.1002/CLI2.15> (2022).

24. Feng, X. *et al.* How deregulation, drought and increasing fire impact Amazonian biodiversity. *Nat.* 2021 5977877 **597**, 516–521, <https://doi.org/10.1038/s41586-021-03876-7> (2021).
25. Filho, J. D. P., de Assis Souza Filho, F., Martins, E. S. P. R. & de Carvalho Studart, T. M. Copula-Based Multivariate Frequency Analysis of the 2012–2018 Drought in Northeast Brazil. *Water* 2020, Vol. 12, Page 834 **12**, 834, <https://doi.org/10.3390/W12030834> (2020).
26. Marengo, J. A., Torres, R. R. & Alves, L. M. Drought in Northeast Brazil—past, present, and future. *Theor. Appl. Climatol.* **129**, 1189–1200, <https://doi.org/10.1007/S00704-016-1840-8/FIGURES/9> (2017).
27. Nobre, P. *et al.* Climate Simulation and Change in the Brazilian Climate Model. *J. Clim.* **26**, 6716–6732, <https://doi.org/10.1175/JCLI-D-12-00580.1> (2013).
28. Nobre, C. A. *et al.* Some Characteristics and Impacts of the Drought and Water Crisis in Southeastern Brazil during 2014 and 2015. *J. Water Resour. Prot.* **8**, 252–262, <https://doi.org/10.4236/JWARP.2016.82022> (2016).
29. Almagro, A., Oliveira, P. T. S., Meira Neto, A. A., Roy, T. & Trocho, P. CABra: A novel large-sample dataset for Brazilian catchments. *Hydrol. Earth Syst. Sci.* **25**, 3105–3135, <https://doi.org/10.5194/HESS-25-3105-2021> (2021).
30. Xavier, A. C., King, C. W. & Scanlon, B. R. Daily gridded meteorological variables in Brazil (1980–2013). *Int. J. Climatol.* **36**, 2644–2659, <https://doi.org/10.1002/JOC.4518> (2016).
31. David, P. C. *et al.* Correspondence Between Model Structures and Hydrological Signatures: A Large-Sample Case Study Using 508 Brazilian Catchments. *Water Resour. Res.* **58**, e2021WR030619, <https://doi.org/10.1029/2021WR030619> (2022).
32. Cortez, B. N., Pires, G. F., Avila-Diaz, A., Fonseca, H. P. & Oliveira, L. R. Nonstationary extreme precipitation in Brazil. *Hydrol. Sci. J.* **67**, 1372–1383, https://doi.org/10.1080/02626667.2022.2075267/SUPPL_FILE/THSJ_A_2075267_SM5035.DOCX (2022).
33. Almagro, A., Oliveira, P. T. S., Nearing, M. A. & Hagemann, S. Projected climate change impacts in rainfall erosivity over Brazil. *Sci. Reports* 2017 717, 1–12, <https://doi.org/10.1038/s41598-017-08298-y> (2017).
34. Almagro, A., Oliveira, P. T. S., Rosolem, R., Hagemann, S. & Nobre, C. A. Performance evaluation of Eta/HadGEM2-ES and Eta/MIROC5 precipitation simulations over Brazil. *Atmos. Res.* **244**, 105053, <https://doi.org/10.1016/J.ATMOSRES.2020.105053> (2020).
35. O'Neill, B. C. *et al.* The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geosci. Model Dev.* **9**, 3461–3482, <https://doi.org/10.5194/GMD-9-3461-2016> (2016).
36. Tebaldi, C. *et al.* Climate model projections from the Scenario Model Intercomparison Project (ScenarioMIP) of CMIP6. *Earth Syst. Dyn.* **12**, 253–293, <https://doi.org/10.5194/ESD-12-253-2021> (2021).
37. Song, Y. H., Nashwan, M. S., Chung, E. S. & Shahid, S. Advances in CMIP6 INM-CM5 over CMIP5 INM-CM4 for precipitation simulation in South Korea. *Atmos. Res.* **247**, 105261, <https://doi.org/10.1016/J.ATMOSRES.2020.105261> (2021).
38. Dereczynski, C. *et al.* Downscaling of climate extremes over South America – Part I: Model evaluation in the reference climate. *Weather Clim. Extrem.* **29**, 100273, <https://doi.org/10.1016/J.WACE.2020.100273> (2020).
39. Mukherjee, S., Aadhar, S., Stone, D. & Mishra, V. Increase in extreme precipitation events under anthropogenic warming in India. *Weather Clim. Extrem.* **20**, 45–53, <https://doi.org/10.1016/J.WACE.2018.03.005> (2018).
40. Clark, M. P. *et al.* Characterizing Uncertainty of the Hydrologic Impacts of Climate Change. *Curr. Clim. Chang. Reports* **2**, 55–64, <https://doi.org/10.1007/S40641-016-0034-X/FIGURES/1> (2016).
41. Gleckler, P. J., Taylor, K. E. & Doutriaux, C. Performance metrics for climate models. *J. Geophys. Res. Atmos.* **113**, 6104, <https://doi.org/10.1029/2007JD008972> (2008).
42. Rajulapati, C. R., Papalexiou, S. M., Clark, M. P. & Pomeroy, J. W. The Perils of Regridding: Examples Using a Global Precipitation Dataset. *J. Appl. Meteorol. Climatol.* **60**, 1561–1573, <https://doi.org/10.1175/JAMC-D-20-0259.1> (2021).
43. Christensen, J. H., Boberg, F., Christensen, O. B. & Lucas-Picher, P. On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophys. Res. Lett.* **35**, <https://doi.org/10.1029/2008GL035694> (2008).
44. Tang, G., Clark, M. P. & Papalexiou, S. M. SC-Earth: A Station-Based Serially Complete Earth Dataset from 1950 to 2019. *J. Clim.* **34**, 6493–6511, <https://doi.org/10.1175/JCLI-D-21-0067.1> (2021).
45. Bürger, G., Sobie, S. R., Cannon, A. J., Werner, A. T. & Murdock, T. Q. Downscaling Extremes: An Intercomparison of Multiple Methods for Future Climate. *J. Clim.* **26**, 3429–3449, <https://doi.org/10.1175/JCLI-D-12-00249.1> (2013).
46. Sone, J. S. *et al.* Water Security in an Uncertain Future: Contrasting Realities from an Availability-Demand Perspective. *Water Resour. Manag.* **36**, 2571–2587, <https://doi.org/10.1007/S11269-022-03160-X/FIGURES/4> (2022).
47. Themeßl, M. J., Gobiet, A. & Heinrich, G. Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal. *Clim. Change* **112**, 449–468, <https://doi.org/10.1007/S10584-011-0224-4/FIGURES/14> (2012).
48. Piani, C., Haerter, J. O. & Coppola, E. Statistical bias correction for daily precipitation in regional climate models over Europe. *Theor. Appl. Climatol.* **99**, 187–192, <https://doi.org/10.1007/S00704-009-0134-9/FIGURES/4> (2010).
49. Iturbide, M. *et al.* The R-based climate4R open framework for reproducible climate data access and post-processing. *Environ. Model. Softw.* **111**, 42–54, <https://doi.org/10.1016/J.ENVSOF.2018.09.009> (2019).
50. Casanueva, A. *et al.* Testing bias adjustment methods for regional climate change applications under observational uncertainty and resolution mismatch. *Atmos. Sci. Lett.* **21**, e978, <https://doi.org/10.1002/ASL.978> (2020).
51. Switanek, B. M. *et al.* 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> (2017).
52. White, R. H. & Toumi, R. The limitations of bias correcting regional climate model inputs. *Geophys. Res. Lett.* **40**, 2907–2912, <https://doi.org/10.1002/GRL.50612> (2013).
53. Ballarin, A. S. *et al.* CLIMBra - Climate Change Dataset for Brazil. *Science Data Bank* <https://doi.org/10.57760/sciencedb.02316> (2022).
54. Pereira, M. F. R., Chaffe, P. L. B., de Amorim, P. B. & Rodrigues, R. R. A systematic analysis of climate model precipitation in southern Brazil. *Int. J. Climatol.* **42**, 4240–4257, <https://doi.org/10.1002/JOC.7460> (2022).
55. Almazroui, M. *et al.* Assessment of CMIP6 Performance and Projected Temperature and Precipitation Changes Over South America. *Earth Syst. Environ.* 2021 5255, 155–183, <https://doi.org/10.1007/S41748-021-00233-6> (2021).
56. Landau, E. C., Tavares, R. B. T., Guimarães, D. P. & Hirsch, A. Geoespacialização da Amplitude Térmica no Brasil. In *III Simpósio Internacional de Climatologia* (2009).
57. Hagemann, S. *et al.* Impact of a Statistical Bias Correction on the Projected Hydrological Changes Obtained from Three GCMs and Two Hydrology Models. *J. Hydrometeorol.* **12**, 556–578, <https://doi.org/10.1175/2011JHM1336.1> (2011).
58. Maurer, E. P. & Pierce, D. W. Bias correction can modify climate model simulated precipitation changes without adverse effect on the ensemble mean. *Hydrol. Earth Syst. Sci.* **18**, 915–925, <https://doi.org/10.5194/HESS-18-915-2014> (2014).
59. Du, Y. *et al.* Comprehensive assessment of CMIP5 and CMIP6 models in simulating and projecting precipitation over the global land. *Int. J. Climatol.* <https://doi.org/10.1002/JOC.7616> (2022).
60. Yukimoto, S. *et al.* The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component. *J. Meteorol. Soc. Japan. Ser. II* **97**, 2019–051, <https://doi.org/10.2151/JMSJ.2019-051> (2019).
61. Döscher, R. *et al.* The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6. *Geosci. Model Dev.* **15**, 2973–3020, <https://doi.org/10.5194/GMD-15-2973-2022> (2022).
62. Cherchi, A. *et al.* Global Mean Climate and Main Patterns of Variability in the CMCC-CM2 Coupled Model. *J. Adv. Model. Earth Syst.* **11**, 185–209, <https://doi.org/10.1029/2018MS001369> (2019).
63. Volodin, E. M. *et al.* Simulation of the modern climate using the INM-CM48 climate model. *Russ. J. Numer. Anal. Math. Model.* **33**, 367–374, <https://doi.org/10.1515/RNAM-2018-0032/MACHINEREADEABLECITATION/RIS> (2018).
64. Seland, Ø. *et al.* 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).

65. Gutzjahr, O. *et al.* 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).
66. Müller, W. A. *et al.* A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR). *J. Adv. Model. Earth Syst.* **10**, 1383–1413, <https://doi.org/10.1029/2017MS001217> (2018).
67. Volodin, E. M. *et al.* Simulation of the present-day climate with the climate model INMCM5. *Clim. Dyn.* **49**, 3715–3734, <https://doi.org/10.1007/S00382-017-3539-7/FIGURES/18> (2017).
68. Law, R. M. *et al.* The carbon cycle in the Australian Community Climate and Earth System Simulator (ACCESS-ESM1) - Part 1: Model description and pre-industrial simulation. *Geosci. Model Dev.* **10**, 2567–2590, <https://doi.org/10.5194/GMD-10-2567-2017> (2017).
69. Ziehn, T. *et al.* The Australian Earth System Model: ACCESS-ESM1.5. *J. South. Hemisph. Earth Syst. Sci.* **70**, 193–214, <https://doi.org/10.1071/ES19035> (2020).
70. Wang, Y. C. *et al.* Performance of the Taiwan Earth System Model in Simulating Climate Variability Compared With Observations and CMIP6 Model Simulations. *J. Adv. Model. Earth Syst.* **13**, e2020MS002353, <https://doi.org/10.1029/2020MS002353> (2021).
71. Cao, J. *et al.* The NUIST Earth System Model (NESM) version 3: Description and preliminary evaluation. *Geosci. Model Dev.* **11**, 2975–2993, <https://doi.org/10.5194/GMD-11-2975-2018> (2018).
72. Pak, G. *et al.* Korea Institute of Ocean Science and Technology Earth System Model and Its Simulation Characteristics. *Ocean Sci. J.* **56**, 18–45, <https://doi.org/10.1007/S12601-021-00001-7/FIGURES/7> (2021).
73. Lee, J. *et al.* Evaluation of the Korea Meteorological Administration Advanced Community Earth-System model (K-ACE). *Asia-Pacific J. Atmos. Sci.* **56**, 381–395, <https://doi.org/10.1007/S13143-019-00144-7/FIGURES/9> (2020).
74. Held, I. M. *et al.* Structure and Performance of GFDL's CM4.0 Climate Model. *J. Adv. Model. Earth Syst.* **11**, 3691–3727, <https://doi.org/10.1029/2019MS001829> (2019).
75. Dunne, J. P. *et al.* The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall Coupled Model Description and Simulation Characteristics. *J. Adv. Model. Earth Syst.* **12**, e2019MS002015, <https://doi.org/10.1029/2019MS002015> (2020).
76. Bi, D. *et al.* Configuration and spin-up of ACCESS-CM2, the new generation Australian Community Climate and Earth System Simulator Coupled Model. *J. South. Hemisph. Earth Syst. Sci.* **70**, 225–251, <https://doi.org/10.1071/ES19040> (2020).
77. Williams, K. D. *et al.* The Met Office Global Coupled Model 3.0 and 3.1 (GC3.0 and GC3.1) Configurations. *J. Adv. Model. Earth Syst.* **10**, 357–380, <https://doi.org/10.1002/2017MS001115> (2018).
78. Boucher, O. *et al.* Presentation and Evaluation of the IPSL-CM6A-LR Climate Model. *J. Adv. Model. Earth Syst.* **12**, e2019MS002010, <https://doi.org/10.1029/2019MS002010> (2020).
79. Sellar, A. A. *et al.* UKESM1: Description and Evaluation of the U.K. Earth System Model. *J. Adv. Model. Earth Syst.* **11**, 4513–4558, <https://doi.org/10.1029/2019MS001739> (2019).
80. Tatebe, H. *et al.* 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).

Acknowledgements

The authors acknowledge the Graduate Program in Hydraulic Engineering and Sanitation – PPGSHS at the University of São Paulo (USP-EESC) for the scientific support. This study was financially supported in part by the Coordination for the Improvement of Higher Education Personnel (CAPES, Finance Code 001), in part by the Brazilian National Council for Scientific and Technological Development — CNPq/Ministry of Science, Technology and Innovation — MCTI and in part by the São Paulo Research Foundation — FAPESP (grant numbers 2015/03806–1, 2020/08140–0, 2019/24292–7, 2021/14016–2, and 2022/06017–1).

Author contributions

All co-authors contributed to the design of the study and to the writing of the manuscript. J.S., G.G. and A.A. downloaded and pre-processed the CMIP6 data. A.B. pre- and post-processed the simulations and performed the validation analysis. D.S. and A.R. generated and standardized the final files. E.W. supervised the project and provided institutional support and financial acquisition.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-023-01956-z>.

Correspondence and requests for materials should be addressed to A.S.B.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023