



Extreme climate indices in Brazil: evaluation of downscaled earth system models at high horizontal resolution

Alvaro Avila-Diaz¹ · Gabriel Abrahão¹ · Flavio Justino¹ · Roger Torres² · Aaron Wilson³

Received: 18 July 2019 / Accepted: 24 April 2020 / Published online: 30 April 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

This study evaluated the performance of 25 earth system models (ESMs), statistically and dynamically downscaled to a high horizontal resolution (0.25° of latitude/longitude), in simulating extreme climate indices of temperature and precipitation for 1980–2005. Datasets analyzed include 21 statistically downscaled ESMs from the National Aeronautics and Space Administration (NASA) Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) and dynamically downscaled Eta Regional Climate Model simulations driven by 4 ESMs generated by the Brazilian National Institute for Space Research (INPE). Downscaled outputs were evaluated against observational gridded datasets at 0.25° resolution over Brazil, quantifying the skill in simulating the observed spatial patterns and trends of climate extremes. Results show that the downscaled products are generally able to reproduce the observed climate indices, although most of them have poorest performance over the Amazon basin for annual and seasonal indices. We find larger discrepancies in the warm spell duration index for almost all downscaled ESMs. The overall ranking shows that three downscaled models (CNRM-CM5, CCSM4, and MRI-CGCM3) perform distinctively better than others. In general, the ensemble mean of the statistically downscaled models achieves better results than any individual models at the annual and seasonal scales. This work provides the largest and most comprehensive intercomparison of statistically and dynamically downscaled extreme climate indices over Brazil and provides a useful guide for researchers and developers to select the models or downscaling techniques that may be most suitable to their applications of interest over a given region.

Keywords CMIP5 · Model evaluation · Climate extremes · Performance metrics · Trends

1 Introduction

The attention of global climate change impacts is progressively moving from the assessments of mean (or climatology) patterns to assessments of present and future trends of climate extreme events, such as the warmest day of the year,

heat/cold waves, heavy or very heavy precipitation events, consecutive dry or wet spells (Burger et al. 2011; Alexander and Arblaster 2017; Avila-Diaz et al. 2020). In this sense, the Expert Team on Climate Change Detection and Indices (ETCCDI) have been developing and publicizing a set of internationally-accepted indices based on daily measures of air temperature and precipitation (Alexander et al. 2006; Donat et al. 2013; Sillmann et al. 2013a, b).

Many studies around the world have applied the ETCCDI climate indices to analyze the risk of climate extremes to human and natural systems, for past events using observed historical data (e.g., Aguilar et al. 2005; Santos et al. 2017), and for future trends in extremes using climate models projections (e.g., Debortoli et al. 2017; Mysiak et al. 2018; Alexander and Arblaster 2017). Despite all efforts, investigations on future climate extremes have been frequently constrained by coarse resolutions in climate models, that lead to results that can not be assumed to reproduce local weather extremes. For instance, using General Circulation Models

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00382-020-05272-9>) contains supplementary material, which is available to authorized users.

✉ Alvaro Avila-Diaz
alvaroaviladiaz@outlook.com

¹ Department of Agricultural Engineering, Universidade Federal de Viçosa, Viçosa, MG, Brazil

² Natural Resources Institute, Universidade Federal de Itajubá, Itajubá, MG, Brazil

³ Byrd Polar and Climate Research Center, The Ohio State University, Columbus, OH, USA

(GCMs) from the Coupled Model Intercomparison Project Phase 3 (CMIP3), Rusticucci et al. (2010) and Marengo et al. (2010b) found that those models exhibit a higher frequency of some climate extremes compared to observations over South America. Additionally, Sillmann et al. (2013a) evaluated the CMIP3 and CMIP5 models over the South American region and noted that many models overestimate the total precipitation in wet days, underestimate the maximum consecutive dry days and generally overestimate temperature extremes.

In light of an increasing need for finer resolution information of climate change projections (horizontal resolution less than 100 km), statistical and dynamical downscaling techniques provide more details of climatic patterns over a particular region, improving the accuracy and relevance of simulations and projections for climate impact studies (Burger et al. 2011; Ambrizzi et al. 2019; Bozkurt et al. 2019). Although such research efforts are relatively rare in Brazil, several regionally/locally downscaled projections have been developed using various methodologies in recent years (Boulanger et al. 2006, 2007; Marengo et al. 2010a, 2012; Thrasher et al. 2012; Chou et al. 2014a; Valverde and Marengo 2014; Avila-Diaz et al. 2020). For example, Valverde and Marengo (2014) and Chou et al. (2014a) assessed regional climate simulations applying dynamical downscaling using the Eta model and noted that the model reproduced reasonably well the extreme climatic events; although, simulations contain more extreme values than the observations.

The Brazilian economy has been highly vulnerable to climatic variability, especially to climate extremes of air temperature and precipitation, that can lead to considerable losses in agricultural activities and problems in the management of water resources (Tomasella et al. 2013; Ray et al. 2015; Debortoli et al. 2017; Marengo et al. 2017). In this way, current climate change projections are likely to have negative socio-economic impacts on the country, increasing the number of natural disasters in regions where climate change will be more pronounced (Torres and Marengo 2014; Darela-Filho et al. 2016).

In recent years, impactful natural hazards related to climate extremes have affected Brazil, such as droughts in the Northeast from 2012 to 2016 (Marengo et al. 2017, 2018; Brito et al. 2018) or dry and warmer summers (December–March) in 2014 and 2015 in Southeast Brazil (Coeelho et al. 2016a, b). However, at the same time, unprecedented floods were reported in the summer of 2014 in the southwestern Amazon Basin (Espinoza et al. 2014). The frequency of such catastrophes spurs the need for reliable simulations of climate extremes on local to regional scales that can inform the development of public policies, proper management of hydrological resources, and the mitigation of their impacts on human activity and the environment (Marengo et al. 2009).

Therefore, current impact projections rely on climate models with coarse resolutions (> 100 km), thus lacking the detail needed for regionally relevant impact assessments. The main goal of this work is to evaluate how well the current climate model downscaling products can simulate variability and trends of climate extremes events in Brazil. We investigate the performance of 25 statistically and dynamically downscaled earth system models (ESMs) to a high horizontal resolution in capturing the observed behavior of extreme temperature and precipitation events over the major Brazilian watersheds. Two main downscaled ESMs data sources were analyzed. First, 21 statistically downscaled ESMs with a horizontal resolution of $0.25^\circ \times 0.25^\circ$ of latitude/longitude (approximately $25 \text{ km} \times 25 \text{ km}$) were taken from National Aeronautics and Space Administration (NASA) Earth Exchange Global Daily Downscaled Projections (NEX-GDDP). Second, 4 dynamically downscaled ESMs using the Eta model at a 20-km spatial resolution were provided by the Brazilian National Institute of Spatial Research (INPE). These two data sources, with relatively high spatial and temporal resolutions, have greatly captured the observed climatic patterns and have been used in studies of climate change impacts on a regional/local scale (Debortoli et al. 2017; Missirian and Schlenker 2017; Lyra et al. 2018; Raghavan et al. 2018; Liao et al. 2019).

The paper is organized as follows: Sect. 2 describes the different data sources and methodology used; results focused on observations and model evaluations are presented in Sect. 3; and finally, summaries, discussions and concluding remarks are presented in Sect. 4.

2 Data and methods

2.1 Extreme climate indices

Twenty-seven extreme climate indices are recommended by the ETCCDI and are calculated using daily maximum (TX) and minimum temperature (TN) and daily precipitation (PR). In this study, some ETCCDI indices are excluded, because their definitions are not appropriate across Brazil. For instance, the index was excluded if the study area has few records of extremely low temperatures as frost days (FD), ice days (ID), cold spell duration indicator (CSDI), and the common magnitude of growing season length is nearly 365 days (GSL). In addition, coldest day (TXn), warmest night (TNx), annual counts of daily minimum temperature greater than 20°C (tropical nights–TR), the maximum temperature greater than 25°C (summer days–SU), and days with rainfall greater than 1/10 mm (R1 mm/R10 mm) were excluded from this analysis, because these thresholds are not relevant to describe extreme climate events in Brazil.

We evaluate 16 extreme climate indices at the annual scale, 8 are associated with temperature and 8 with precipitation. Detailed descriptions are provided in Table 1, and further details may also be found in Zhang et al. (2004) and Zhang et al. (2011), or at https://etccdi.pacificclimate.org/list_27_indices.shtml.

The extreme climate indices chosen can be calculated seasonally or monthly, albeit most of the impactful extreme events mentioned in the previous section can be described by annual indices (Aerenson et al. 2018). However, the seasonal analysis was done for the two extreme climate seasons over Brazil: austral summer (December, January, and February—DJF) and winter (June, July, and August—JJA), representing the wet and dry seasons, respectively, for the most of the country (Marengo et al. 2010a; Torres and Marengo 2014; Rao et al. 2016; Lyra et al. 2018). For this purpose, seasonal analysis was carried out for selected warm extremes (TXx, TX90P), cold extremes (TNn, TN10P), wet extremes precipitation (PRCPTOT, RX1day, RX5day) and the maximum number of consecutive dry days (CDD), which is associated with dry conditions (Zhang et al. 2011) and also indicative of potential water stress (Aerenson et al. 2018).

The climate indices were chosen, because they allow the assessment of intensity, frequency, and duration of extreme climate events. Also, this set of indices has been used to describe hydrometeorological hazards such as droughts, floods, heavy rains, and heat waves in Brazilian climate conditions (Alexander et al. 2006; Sillmann et al. 2013a,

b; Skansi et al. 2013; Ávila et al. 2016; Avila-Diaz et al. 2020). Noteworthy, the ETCCDI indices are widely used to evaluate the capability of ESMs in simulating the observed climate extremes of temperature and precipitation (Marengo et al. 2010b; Rusticucci et al. 2010; Alexander and Arblaster 2017; Nguyen et al. 2017; Avila et al. 2019; Dosio et al. 2019; Loaiza et al. 2020).

All extreme indices were calculated using gridded datasets (observational and reanalysis) and 25 downscaled ESMs that are shown in Sects. 2.2 and 2.3, respectively. The calculations are performed with the `climindex.pcic.ncdf` package maintained by the Pacific Climate Impacts Consortium (PCIC), which runs on R software and is freely available at <https://github.com/pacificclimate/climindex.pcic.ncdf>.

2.2 Observed datasets

We examine the daily records from two gridded datasets. The first observational dataset (OBS-BR) contains daily fields of temperature and precipitation interpolated from 9259 rain gauges and 735 weather stations gridded to a regular grid of $0.25^\circ \times 0.25^\circ$ latitude/longitude covering all of Brazil territory over the period 1980–2015 (Xavier et al. 2015, 2017). The dataset is available at <https://utexas.app.box.com/v/Xavier-et-al-IJOC-DATA>. Noteworthy, daily fields of the observations and simulations covered 1980–2005, because the historical experiment for each downscaled model are only available through 2005 (see Sect. 2.3). The second

Table 1 The extreme temperature and precipitation indices used in this study recommended by the ETCCDI

Index	Indicator name	Indicator definitions	Units
TXx	Hottest day	Maximum value of daily maximum temp (TX)	°C
TNn	Coldest nights	Minimum value of daily minimum temp (TN)	°C
DTR	Diurnal temperature range	Mean difference between TX and TN	°C
TN10p	Cold nights	Percentage of days when TN < 10th percentile	%
TN90p	Warm nights	Percentage of days when TN > 90th percentile	%
TX10p	Cold days	Percentage of days when TX < 10th percentile	%
TX90p	Warm days	Percentage of days when TX > 90th percentile	%
WSDI	Warm spell duration indicator	Annual count of days with at least 6 consecutive days when TX > 90th percentile	Days
PRCPTOT	Annual total wet-day precipitation	Total precipitation (PR) in wet days ($PR \geq 1$ mm)	mm
RX1day	Max 1-day precipitation amount	Highest maximum 1-day precipitation	mm
RX5day	Max 5-day precipitation amount	Highest maximum 5-day precipitation	mm
R95p	Very wet days	Annual total precipitation from days > 95th percentile	mm
SDII	Simple daily intensity index	The ratio of annual total precipitation to the number of wet days ^a (≥ 1 mm)	mm/day
R20mm	Number of very heavy precipitation days	Number of days when $PR \geq 20$ mm	Days
CWD	Consecutive wet days	Maximum number of consecutive days with $PR \geq 1$ mm	Days
CDD ^a	Consecutive dry days	Maximum number of consecutive days with $PR < 1$ mm	Days

More details on definitions of the core indices given at https://etccdi.pacificclimate.org/list_27_indices.shtml

^aThe ETCCDI defined a wet (dry) day when precipitation ≥ 1 mm ($PR < 1$ mm)

dataset used is from Global Meteorological Forcing Dataset (GMFD) (Sheffield et al. 2006), which consists of 3-hourly, 0.25°-resolution fields of near-surface meteorological variables for global land areas for 1948–2016, available at the Terrestrial Hydrology Research Group website at Princeton University (<https://hydrology.princeton.edu/data.GMFD.php>). The GMFD is a merge of several datasets from the National Centers for Environmental Prediction—National Center for Atmospheric Research reanalysis (NCEP-NCAR), the satellite-based Global Precipitation Climatology Project (GPCP), Tropical Rainfall Monitoring Mission (TRMM) and interpolated ground observations from Climatic Research Unit (CRU).

The OBS-BR and GMFD employ different interpolation methods, quality control, and station networks in their development. GMFD was designed for pixel-scale hydrological consistency and has to rely on the NCEP reanalysis as the basis for daily weather variability (Sheffield et al. 2006). The use of a reanalysis product introduces an additional error source, and can lead to smoother meteorological series in comparison to what is observed in the weather stations, which can dampen the magnitude of extreme climate events (Zhang et al. 2011). OBS-BR, on the other hand, directly interpolates the daily observations of its larger weather station/rain gauge network. For these reasons, we considered OBS-BR as the reference daily gridded dataset for meteorological variables (TN, TX, and PR) in Brazil. Special attention was given to the comparison between GMFD and OBS-BR datasets, because the National Aeronautics and Space Administration (NASA) Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) dataset (Thrasher et al. 2012) used GMFD as the observational reference for its statistical downscaling technique. The following subsection describes the NEX-GDDP dataset.

2.3 Earth system model data

For each dynamical and statistical dataset (see Table 2 for the list of models), we used the daily output of maximum and minimum temperature, and daily precipitation to study extreme climate indices from 1980 to 2005.

The 21 statistically downscaled CMIP5 ESMs were obtained from the NEX-GDDP dataset (Thrasher et al. 2012). This dataset is available at <https://nex.nasa.gov/nex/projects/1356/>. It consists of the results of 21 CMIP5 models, bias-corrected and disaggregated to a grid of horizontal resolution of 0.25° of latitude/longitude using a spatially designed statistical technique that compares the model's historical runs (1950–2005) with the GMFD dataset.

On the other hand, the dynamically downscaled simulations employed in our study have been generated by the ETA regional climate model, provided by The Brazilian Center for Weather Forecasts and Climate Studies—CPTEC and

Brazilian National Institute for Space Research—INPE, available at <https://projeta.cptec.inpe.br> (Chou et al. 2014a, b; MCTI 2016; Lyra et al. 2018). The ETA simulations are based on 4 ESMs that have been downscaled to a 20-km resolution (Table 2). The model domain covers South America and most of Central America, available from 1960 to 2005. Regarding to Chou et al. (2014a) and Lyra et al. (2018), the ETA model largely improves the seasonal cycles and precipitation frequency distributions when compared to the driving ESM. However, they retain some of the distortions of trends in extreme indices present in the ESM simulations, such as the cooling trend in maximum and minimum temperatures in Eta-MIROC5, and different spatial patterns of extreme precipitation trends among the models (Chou et al. 2014a). More information on the simulations, including a detailed comparison between their results for some extreme indices, can be found in Chou et al. (2014a) and Lyra et al. (2018). For intercomparison purposes, the ETA 20 km grid here was interpolated to a common 0.25° × 0.25° grid, using a first-order conservative remapping technique (Jones 1999), as proposed in the literature (Giorgi 2006; Cheng and Knutson 2008; Sillmann et al. 2013a; Torres and Marengo 2014).

Besides analyzing each model separately, we test whether using Multi-Model Ensembles (MMEs) can improve the representation of climate extremes. Taking the mean of a model ensemble is a common technique for avoiding the large spread found in individual model results (Knutti et al. 2010; Sillmann et al. 2013a; Nguyen et al. 2017). For that end, the mean of each index among the statistical (MME-Sta) and dynamical (MME-Dyn) models was calculated, and we treated those as separate results.

2.4 Evaluation metrics and trend calculation

The metrics used for evaluating the simulated indices include Percent Bias (*PBIAS*), *RMSE*-observations standard deviation ratio (*RSR*), refined index of agreement (d_r) (Willmott et al. (2012) and the Pearson correlation coefficient (*CORR*). These statistical parameters are calculated as follows:

$$PBIAS = \frac{\sum_{i=1}^n (m_i - O_i) \times 100}{\sum_{i=1}^n O_i}, \quad (1)$$

$$RSR = \frac{RMSE}{STDEV_{Obs}} = \frac{\sqrt{(m_i - O_i)^2}}{\sqrt{(O_i - \bar{O})^2}}, \quad (2)$$

$$d_r = \begin{cases} 1 - \frac{\sum_{i=1}^n |m_i - O_i|}{2 \times \sum_{i=1}^n |O_i - \bar{O}|}, & \text{when } \sum_{i=1}^n |m_i - O_i| \leq 2 \times \sum_{i=1}^n |O_i - \bar{O}| \\ \frac{2 \times \sum_{i=1}^n |O_i - \bar{O}|}{\sum_{i=1}^n |m_i - O_i|} - 1, & \text{when } \sum_{i=1}^n |m_i - O_i| > 2 \times \sum_{i=1}^n |O_i - \bar{O}| \end{cases}, \quad (3)$$

Table 2 Information on the 25 general circulation models used in the present analysis

Model	Modeling center	Resolution (Lat°×Lon°)	VL ^a
(a) Statistical downscaled models			
1. ACCESS1-0	Commonwealth Scientific and Industrial Research Organization and Bureau of Meteorology, Australia	1.875°×1.25°	38
2. BCC-CSM1-1	Beijing Climate Center, China Meteorological Administration, China	2.8°×2.8°	26
3. BNU-ESM	Beijing Normal University, China	2.8°×2.8°	26
4. CanESM2	Canadian Centre for Climate Modeling and Analysis, Canada	2.8°×2.8°	35
5. CCSM4	National Center for Atmospheric Research (NCAR), USA	1.25°×0.9°	27
6. CESM1-BGC	National Science Foundation, Department of Energy and NCAR, USA	1.25°×0.9°	27
7. CNRM-CM5	Centre National de Recherches Météorologiques and Centre Européen de Recherche et Formation Avancée en Calcul Scientifique, France	1.4°×1.4	31
8. CSIRO-MK3-6-0	Commonwealth Scientific and Industrial Research Organization, Australia	1.875°×1.875°	18
9. GFDL-CM3	NOAA Geophysical Fluid Dynamics Laboratory, USA	2.5°×2.0°	48
10. GFDL-ESM2G	NOAA Geophysical Fluid Dynamics Laboratory, USA	2.5°×2.0°	24
11. GFDL-ESM2M	NOAA Geophysical Fluid Dynamics Laboratory, USA	2.5°×2.0°	24
12. INMCM4	Institute for Numerical Mathematics (INM), Russia	~1.5°×2.0°	21
13. IPSL-CM5A-LR	Institut Pierre-Simon Laplace, France	3.75°×1.895	39
14. IPSL-CM5A-MR	Institut Pierre-Simon Laplace, France	2.5°×1.27°	39
15. MIROC5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine- Earth Science and Technology	1.4°×1.4°	40
16. MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	2.8°×2.8°	80
17. MIROC-ESM-CHEM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	2.8°×2.8°	80
18. MPI-ESM-LR	Max Planck Institute for Meteorology, Germany	1.875°×1.875°	47
19. MPI-ESM-MR	Max Planck Institute for Meteorology, Germany	1.875°×1.875°	95
20. MRI-CGCM3	Meteorological Research Institute, Japan	1.125°×1.125°	48
21. NorESM1-M	Norwegian Climate Centre, Norway	2.5°×~1.89°	26
(b) Dynamical downscaled models			
22. BESM	Brazilian Earth System Model version (Version 2.3.1), Brazil	1.875°×1.875°	28
23. CanESM2	Canadian Centre for Climate Modeling and Analysis, Canada	2.8°×2.8°	35
24. HadGEM2-ES	Met Office Hadley Centre, UK	1.25°×1.875°	38
25. MIROC5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	1.4°×1.4°	40

(a) ESMs used for generating the NEX-GDDP dataset through statistical downscaling. (b) ESMs used as boundary conditions to generate the Eta-INPE dataset through dynamical downscaling

^aVertical layers

$$CORR = \frac{\sum_{i=1}^n (m_i - \bar{m}_i)(O_i - \bar{O})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})} \sqrt{\sum_{i=1}^n (m_i - \bar{m}_i)}}, \tag{4}$$

where O_i is the observed value, \bar{O}_i is the mean of observed data, m_i is the simulated value, \bar{m}_i is the mean of simulated data, and n is the total observation number.

PBIAS indicates the average tendency of the simulation to be larger or smaller than the observed data (Gupta et al.

1999); values close to 0 indicate an optimal performance in a given model; positive and negative values indicate a bias toward overestimation or underestimation, respectively. *RSR* is calculated as the ratio of the *RMSE* and standard deviation of observed data (Moriassi et al. 2007); values closer to 0 mean better performing simulations. The refined index of agreement (d_r) developed by Willmott et al. (2012) varies between -1 and 1 . A d_r of 1 indicates a perfect agreement and $d_r = -1$ indicates either a lack of agreement between observed and simulated values or a lack of variability in

the observed data (Willmott et al. 2015). Finally, *CORR* is used to describe the temporal association between observed data and model simulations. *CORR* is between -1 and 1 . A *CORR* of 1 (-1) shows complete positive (negative) linear relation. If the *CORR* is 0 , there is a lack of any linear relationship between observed (O_i) and simulated (m_i) data.

Individual extreme climate index scores allow us to rank models based on the performance metrics (*PBIAS*, *RSR*, *dr*, and *CORR*). To summarize all the ranking possibilities, the comprehensive model rank (M_R) has also been calculated (Jiang et al. 2015; You et al. 2017; Zhang et al. 2018). M_R is a measure of how consistently each model is classified among all the ranking possibilities (indices and metrics):

$$M_R = 1 - \frac{1}{n \times m} \sum_{i=1}^n (\text{Rank}_{i_{PBIAS}} + \text{Rank}_{i_{RSR}} + \text{Rank}_{i_{dr}} + \text{Rank}_{i_{CORR}}), \quad (5)$$

where n is the total number of indices, m is the number of models and the Rank_i indicates downscaled model's order on each index in a given performance metric. Note that we also rank all the downscaled ESMs along with the two MMEs (Table 2). Therefore, the maximum value of M_R is 1 , indicating that the model is the best in all indices and metrics (Jiang et al. 2015; You et al. 2017).

The linear trends of extreme climate indices from downscaled ESMs are estimated and compared to two observed datasets using the Theil–Sen slope estimator (Sen 1968). The trend significance of the slope was evaluated through Mann–Kendall (Mann 1945; Kendall 1975) trend significance test at the 95% confidence level. These tests have been broadly used in hydrometeorological studies for detecting trends because of their non-parametric approach (Yue et al. 2002; Skansi et al. 2013; Avila et al. 2019).

The performance metrics and trends were evaluated for each grid point and averaged across each of the eight major hydrological basins (Fig. 1) used by the Brazilian National Water Agency (ANA). The basin acronyms in Fig. 1 refer to Amazon River (AMZ), Tocantins River (TOC), North Atlantic Region (NAR), São Francisco River (SFR), Central Atlantic Region (CAR), Parana River (PAR), Uruguay river (URU), and South Atlantic Region (SAR). The analysis in hydrological basins was done because the performance of each ESMs within a given hydrological basin is expected to be consistently representative throughout that specific region (Nguyen et al. 2017; Xu et al. 2019; Avila-Diaz et al. 2020).

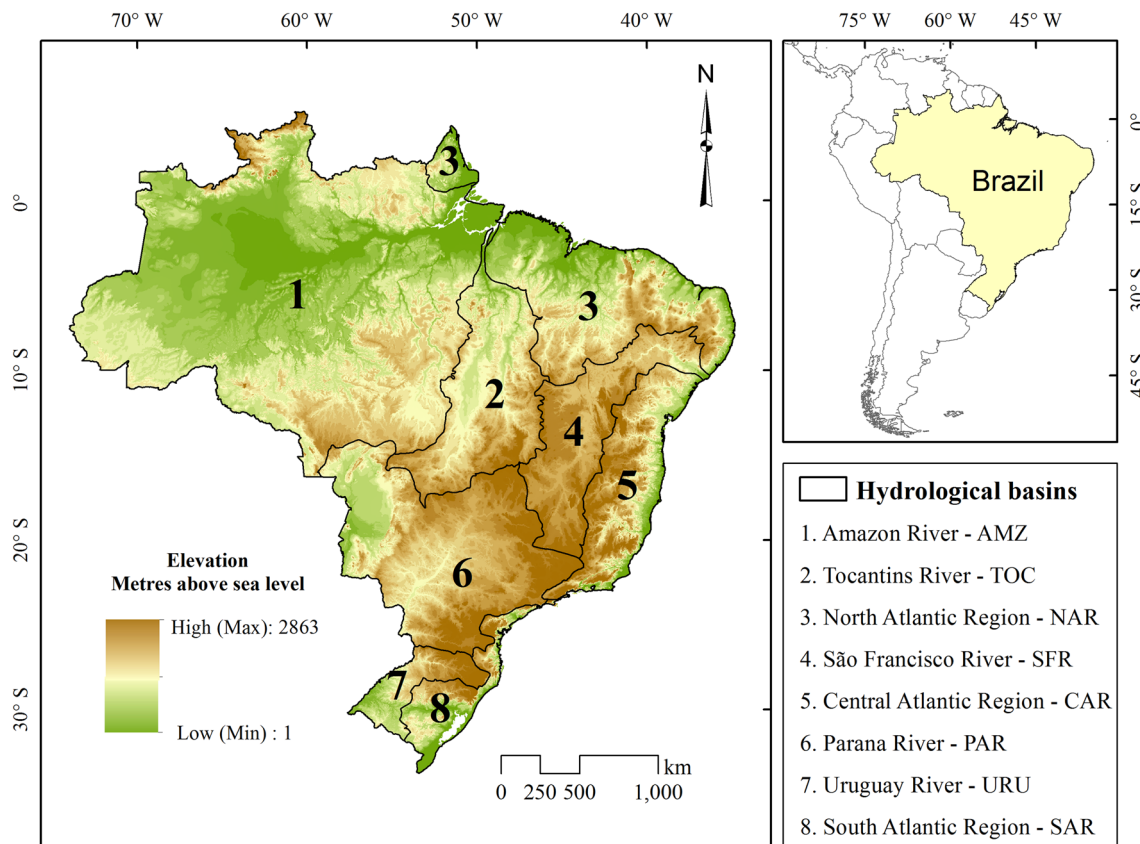


Fig. 1 Geographical location of the eight hydrological basins in Brazil according to the Brazilian National Water Authority (ANA) classification

3 Results and discussion

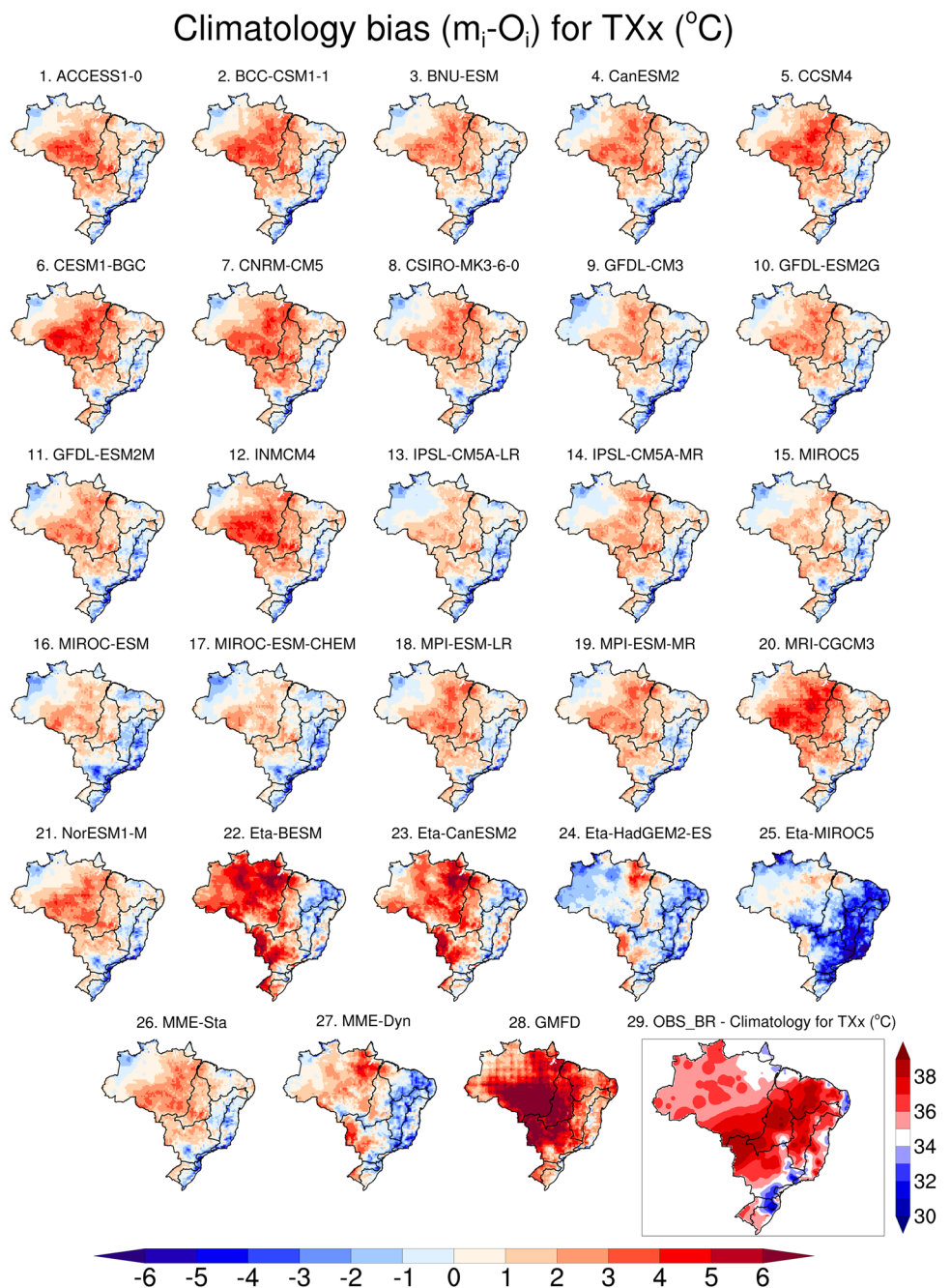
For sake of brevity, we discuss the results of climatology bias and spatial trend analysis for two indices, which represent extremes events of temperature (hottest days—TXx) and precipitation (annual total wet-day precipitation—PRCPTOT) as illustrative figures in the main text. Results for the other ETCCDI indices can be found in the supplementary material.

3.1 Temperature indices

3.1.1 Evaluation metrics

Figure 2 presents the climatology bias of the hottest day index (TXx). Almost all downscaled ESMs and the reanalysis (GMFD) captured the spatial pattern of the TXx relatively well. Furthermore, the evaluation metrics (*PBIAS*, *RSR*, *d_r*, and *CORR*) were calculated for each climate index, model, MMEs, and GMFD dataset in each hydrological basin and compared to observations during 1980–2005.

Fig. 2 Climatology bias for the annual maximum of daily maximum temperature—TXx (°C) for 21 statically (NEX-GDDP; models 1–21) and 4 dynamically (Eta-INPE; models 22–25) downscaled models, MME-Sta (26), MME-Dyn (27), and GMFD (28) from 1980 to 2005. Climatology for TXx in the observations dataset (OBS-BR; gray rectangle dataset 29) for 1980–2005



These evaluation metrics are summarized in the portrait diagrams shown for annual (Fig. 3) and seasonal (Fig. 4) results.

In general, the statistical downscaling approach captures well the general spatial patterns of the temperature indices at the annual scale. However, the statistically downscaled ESMs more frequently overestimate TXx above 2 °C over

the Amazon basin (Fig. 2). Also, there are generally low values of *PBIAS* for almost all temperature indices of GMFD (Fig. 3a) except for TXx, coldest night (TNn), and warm spell duration indicator (WSDI). The greatest *PBIAS* values for TXx are found over the Amazon, Tocantins, and Parana basins with values larger than 12%. GMFD underestimated observed values for TNn except for the São Francisco River

Fig. 3 Statistics of performance obtained for annual temperature indices for statically (NEX-GDDP) and dynamically (Eta-INPE) downscaled models, MME-Sta, MME-Dyn, and GMFD from 1980 to 2005 over eight hydrological basins (Fig. 1). **a** Percent Bias (*PBIAS*); **b** RMSE-observations standard deviation ratio (*RSR*); **c** a refined index of model performance (*dr*); **d** Correlation coefficients (*CORR*; the diagonal lines indicate significant correlations at 95% level). The horizontal purple lines refer to Eta-INPE datasets. For *PBIAS* and *RSR*, dark colors indicate models that perform worse than others, on average, and light colors indicate models that perform better than others, on average. Furthermore, for *dr* and *CORR*, dark (light) colors show models that have better (worse) statistical metrics than others, on average

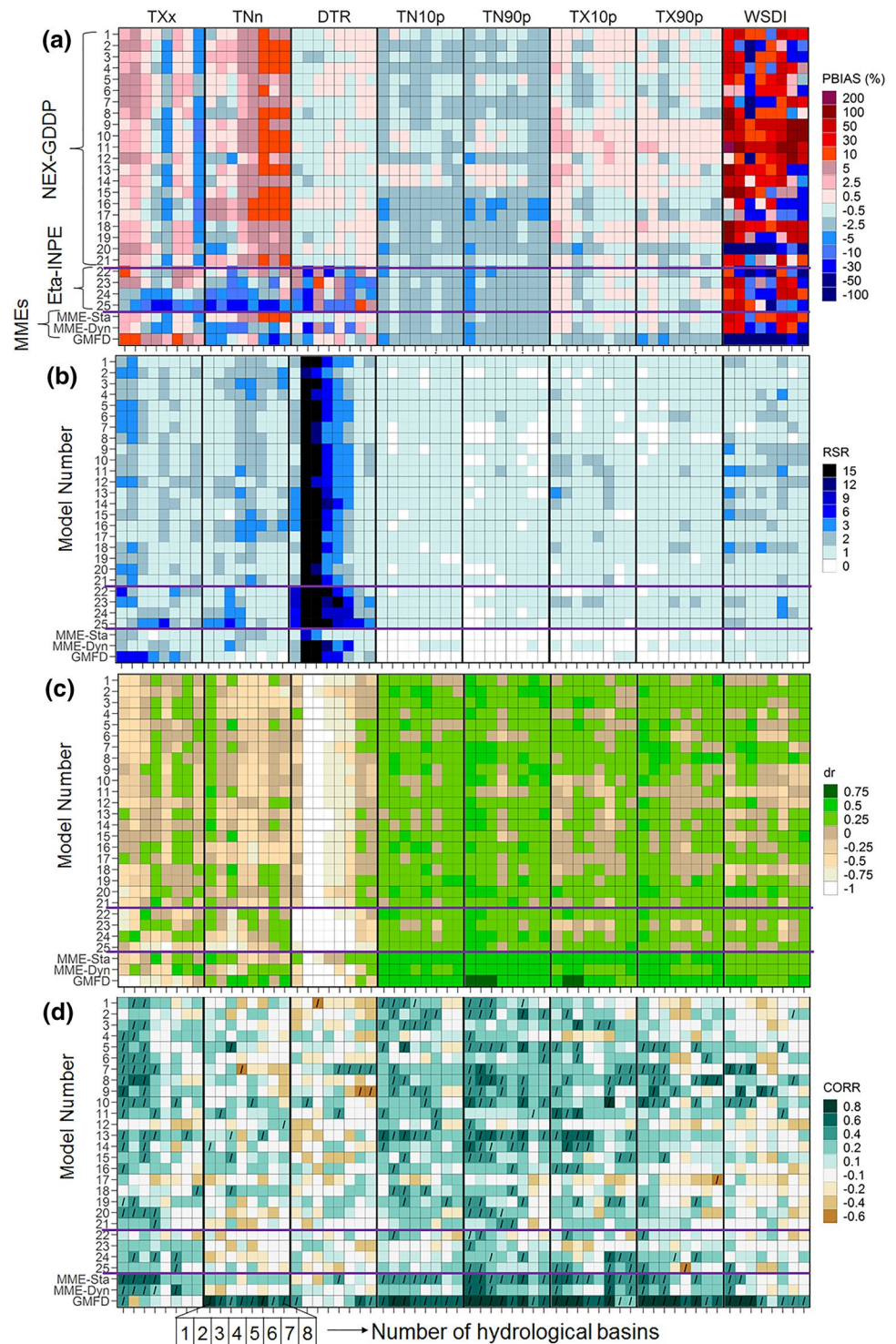
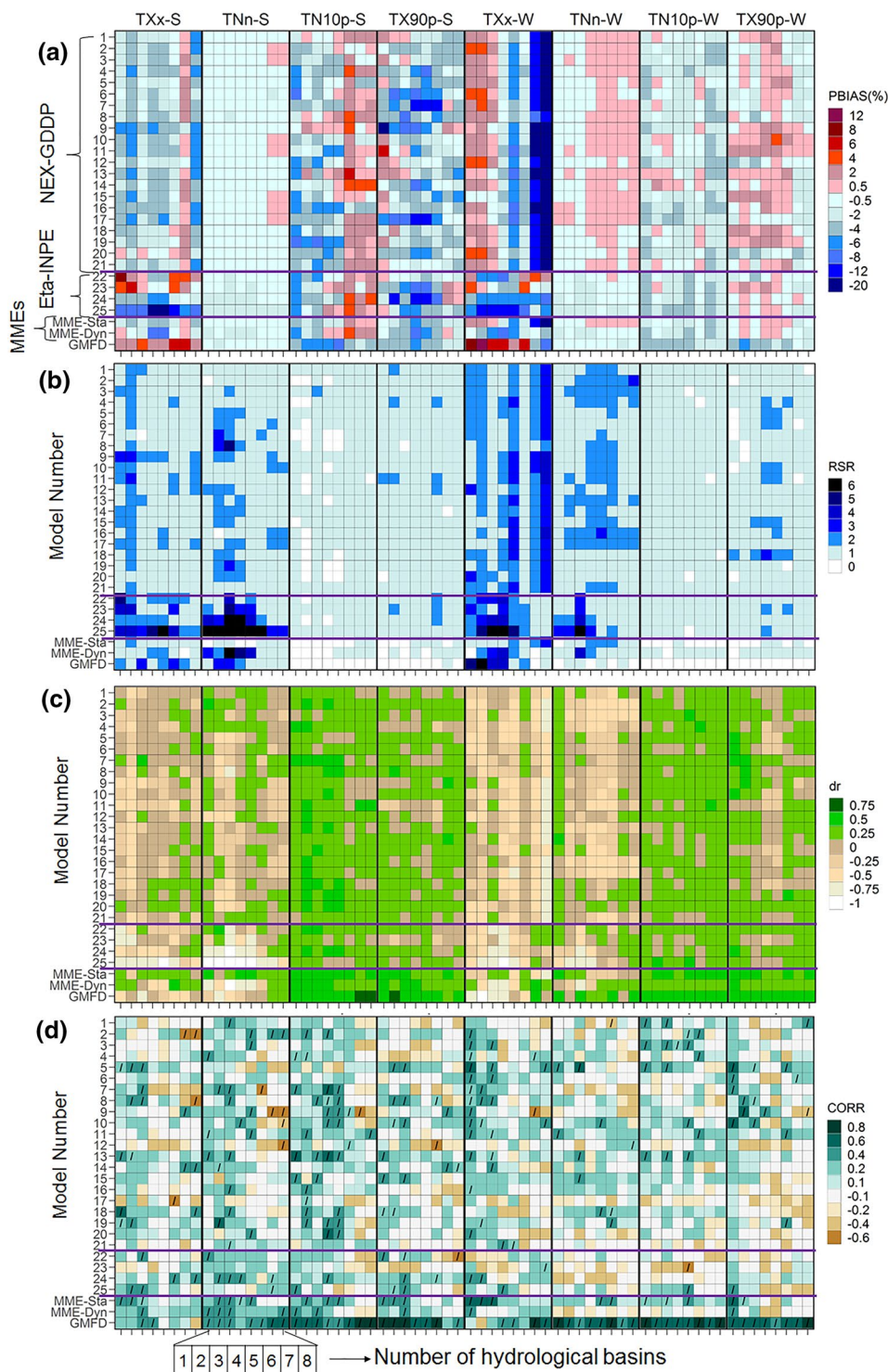


Fig. 4 As in Fig. 2, but for extreme temperature indices in summer (DJF) and winter (JJA). The nomenclature of the ETCCDI indices was adapted to Index—“S” for summer and Index—“W” for winter



and Central Atlantic basins (overestimated by slightly more than 1%). The worst values of *PBIAS* for TNn are identified over the Parana and Uruguay basins with -4 and -13% , respectively.

The *PBIAS* of TNn (Fig. 2a; second column) shows that some downscaled models are too cold [e.g.,

Eta-HadGEM2-ES (24) and Eta-MIROC5 (25)] or simulate higher values of minimum temperatures [e.g., MIROC-ESM (16) and MIROC-ESM-CHEM (17)] over Brazil. Furthermore, models 2, 3, 16, and 17 did not perform well for the TNn index over the Parana River, Uruguay River, and South Atlantic basins with values above 19% (Fig. 3a). In terms of

the RSR , d_r , and correlation, higher limitations have been found for the majority of the 25 downscaled ESMs for TXx and DTR, especially for basins located in the North and Northeast of Brazil. It is important to note that TXx and TNn indices are generally underestimated in the Eta simulations over Brazil (e.g., models 24 and 25), in agreement with the results of Chou et al. (2014a).

Figure 4 displays the performance obtained for temperature indices in summer and winter. In general, the downscaled models underestimate the observations of the TXx and TNn indices in summer in almost all basins except for the Uruguay basin, which shows a warm bias only for TXx. Two of the ETA-INPE models (BESM and CanESM2) have the reverse behavior, tending to overestimate Txx in the summer in most basins, which offsets the strong underestimating bias of ETA MIROC5 in MME-Dyn. The results for winter show that the downscaled ESMs overestimate the TXx over Amazon, Tocantins, and North Atlantic basins, but strongly underestimate ($PBIAS > 30\%$) over Uruguay, Parana, and South Atlantic basins. The TNn index, for the winter, shows $PBIAS$ values lesser than 2%; however, poor performance in RSR , d_r , and $CORR$. The TXx index (TNn) shows that GMFD has a warm (cold) bias in both seasons (bottom of Fig. 4a). For the rest of the metrics (RSR , d_r , and $CORR$), GMFD shows better performance to reproduce TNn than the TXx index over the majority of basins.

The discrepancy of the majority of downscaled models is more evident for WSDI, which simulates higher values than the observations, especially over the Amazon, Tocantins, and Parana basins (Fig. 3b). The WSDI underestimates the observed values by more than 25% and 30% for the Uruguay River and South Atlantic basins, respectively, and for the other basins by more than 66%. Moreover, eight temperature indices have values of RSR close to zero (Fig. 3b), except for the TXx index. Additionally, the redefined index values ($d_r < 0.5$; Fig. 3c) and correlation coefficients ($CORR < 0.5$; Fig. 2d) show the poor performance of the downscaled models to reproduce the TXx and DTR (diurnal temperature range), especially over the Amazon and Tocantins basins.

Evaluation metrics of the GMFD dataset demonstrate reasonable skill in the representation of the temperature-based percentile indices at the annual scale (e.g., TN10p, TX10p, TN90p, and TX90p) across all hydrological basins in Brazil (bottom of Fig. 3a–d). For these four extreme climate indices in almost all basins, $PBIAS$ is within $\pm 3\%$, $RSR < 1$, $d_r \geq 0.50$, and $CORR \geq 0.54$. However, it has been found that TX10p has particularly poor performance over the Uruguay River and South Atlantic basins. Similar to the annual scale, GMFD performs well in reproducing the summer and winter patterns (Fig. 4a–d) of the selected percentile indices (TN10p, TX90p).

The evaluation metrics display good performance of the downscaled models to reproduce TN10p, TX10p, TN90p,

and TX90p at the annual scale (see four to seven columns of Fig. 3). For these indices, the $PBIAS$ varies between -5 to 4%, and the RSR values are close to zero. According to d_r and correlation, models CSIRO-MK3-6-0 (8), CNRM-CM5 (7), and MRI-CGCM3 (20) show consistent performance over all eight basins shown in Fig. 1. The seasonal patterns of TN10p and TX90p are reproduced reasonably well by downscaled models (Fig. 4a–d). For TN10p and TX90p, the $PBIAS$ varies between -8 to 6%, and the RSR and d_r show good accuracy with values close to 0–1 over the majority of basins.

The low bias found in percentile indices is similar to previous studies that used raw ESMs (Marengo et al. 2010b; Rusticucci et al. 2010; Sillmann et al. 2013a) and regional climate model results over South America (Marengo et al. 2009; Dereczynski et al. 2013). The good performance for percentile indices is likely a consequence of their construction, which includes exceedance rates (in percentage) of temperatures colder than the 10th percentile or warmer than the 90th percentile with respect to a base period, potentially minimizing model characteristics (Zhang et al. 2011). Moreover, the percentile indices have less extreme features of climate variability than absolute indices (e.g., TXx and TNn) (Sillmann et al. 2013a). Finally, the $PBIAS$ magnitudes of WSDI are within $\pm 200\%$. The worst performance (based on RSR , d_r , and $CORR$) across almost all the 25 downscaled ESMs is found in the basins located over the North, Northeast, and Central-West regions of Brazil (see Figs. 2, 3, 4).

In general, both MMEs over or underestimate the majority of temperature indices by less than 10%, except the WSDI, with $PBIAS$ varying between -11 to 71% and -32 to 53% for MME-Sta and MME-Dyn, respectively. The MME-Sta display lower $PBIAS$ and RSR and higher correlations and d_r than MME-Dyn for nearly all temperature indices. Our results suggest that MMEs-Sta can better reproduce the interannual variability of temperature extremes in Brazil than MME-Dyn. Some of the downscaled models show better DTR (models 2, 8, 7, 20, and MME-Sta) and WSDI (models 6 and 16) than the raw models analyzed by Sillmann et al. (2013a). This may be related to the quantile mapping applied to the statistical downscaling, which makes the probability distribution of the downscaled data more narrowed. As discussed by Tang et al. (2016), statistical downscaling is based on linear regression with fewer degrees of freedom with respect to the dynamical counterpart (Wilby and Dawson 2013). In terms of precipitation, the complexity in the latter approach is even higher due to the non-linear interaction between clouds, atmospheric circulation, meso-scale processes, and land–atmosphere interaction.

3.1.2 Trend analysis in temperature indices

Trends are calculated for the OBS-BR and GMFD observations and for each downscaled ESMs for temperature indices at the annual scale for 1980–2005 (Figs. 5, 6). The OBS-BR dataset shows warming trends for most of the temperature indices in all hydrological basins of Brazil, most of which are significant at the 95% confidence level. The interested reader is referred to Fig. S8 to follow the seasonal results for warm extremes (TXx, TX90p) and cold extremes (TNn, TN10p), that also shows warming trends. The warming is generally larger in indices related to the warmest days (TXx) than in the coldest days (TNn) (Fig. 5 and see the spatial trends of TXx in Fig. 6).

The trend signal of percentile indices (TN10p, T90p, TX10p, and TX90p) is in line with observational analyses from Vincent et al. (2005), Skansi et al. (2013), Donat et al. (2013) and Avila-Diaz et al. (2020), indicating warmer conditions over Brazil at annual (Fig. 5) and seasonal scales. Furthermore, the positive trends in TXx (Figs. 5a, 6) and a

narrowing tendency of DTR (Fig. 5c) over southern Brazil by the OBS-BR are consistent with the results observed by Marengo and Camargo (2008) and Rosso et al. (2015) during 1960–2002 and 1961–2011 periods, respectively. However, they found positive trends for the TNn index, but our results indicate negative trends in southern Brazil. These authors employed different periods and more years than the ones used in this study, with low-frequency features of the time series potentially changing the evaluated trends in TNn.

The magnitude of warming trends in cold nights, warm nights, cold days, warm days, and warm spell duration indices is relatively coherent across all the downscaled ESMs datasets and model ensembles (Fig. 5). Also, seasonal trend patterns in summer and winter for percentile indices (e.g., TN10p and TN90p) are well captured in almost all downscaled models. The upward trends found for most indices are a common feature delivered on the models evaluated except for MIROC-ESM-CHEM (17). This model shows negative (cooling) trends in several indices at annual and seasonal scales that are positive (warming) according

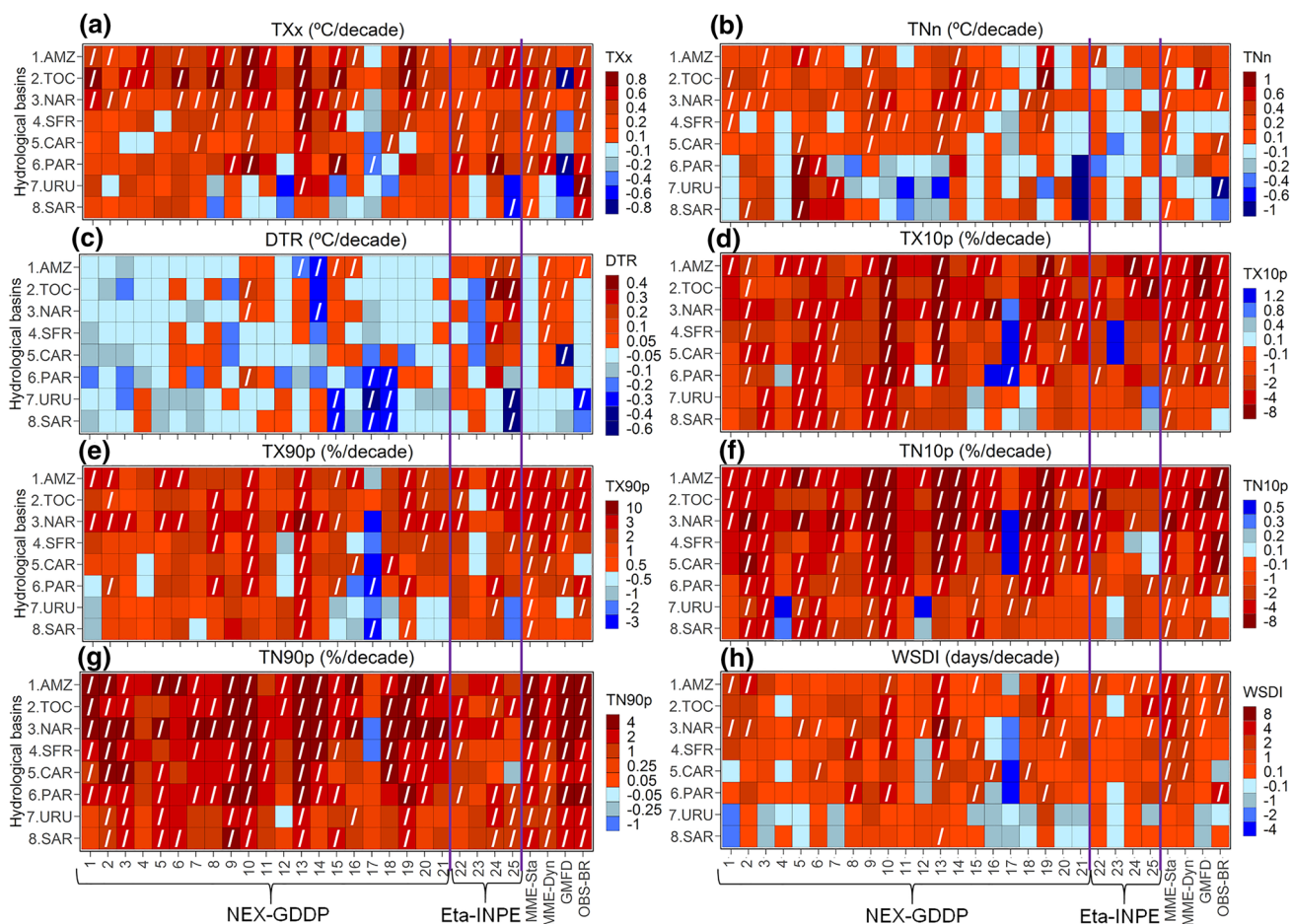
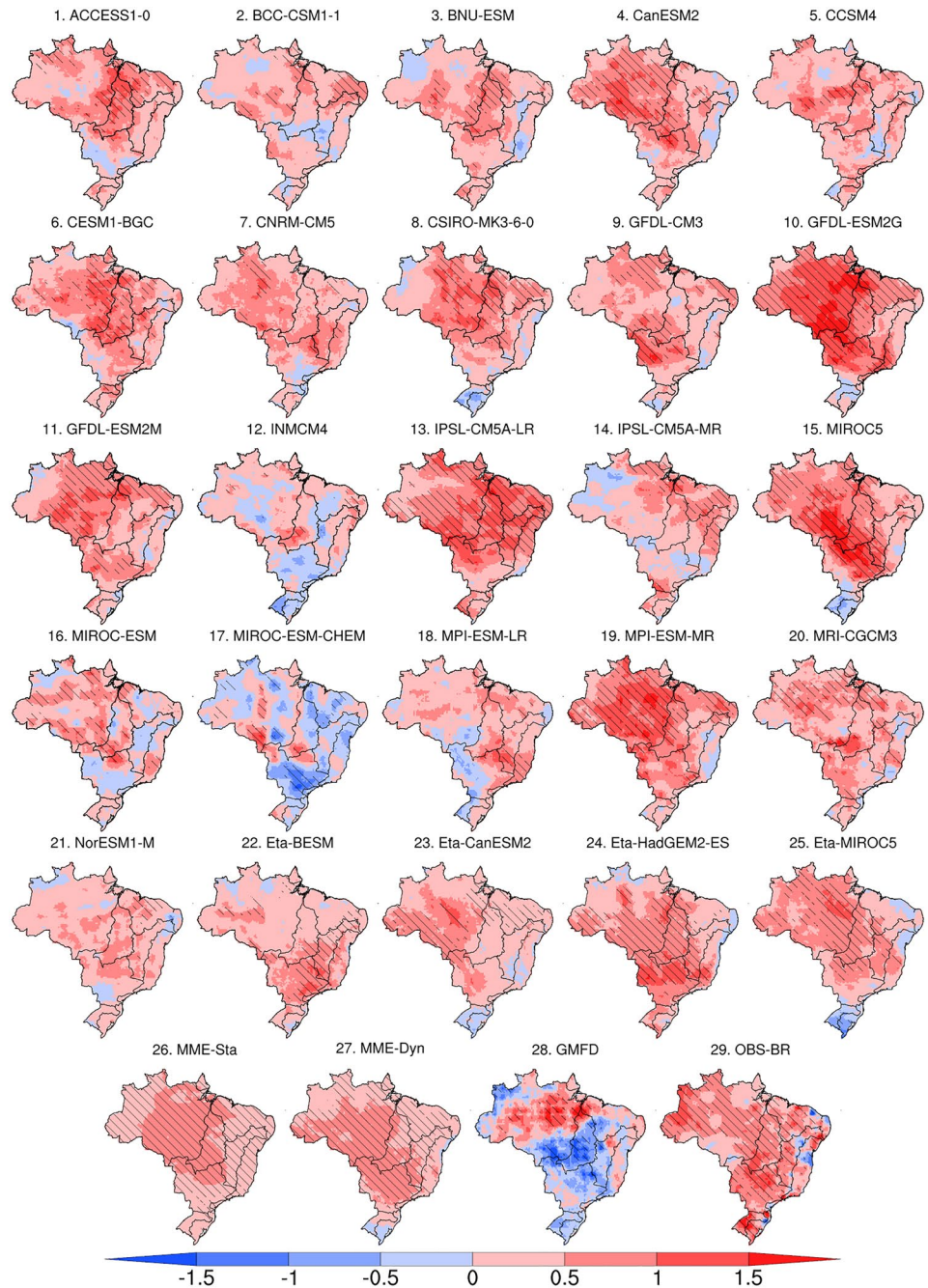


Fig. 5 Trends per decade from 1980 to 2005 for temperature indices at the annual scale **a–h** for 21 NEX-GDDP climate models (1–21), 4 Eta-INPE climate models (22–25), MME-Sta (26), MME-Dyn (27),

GMFD (28) and OBS-BR over eight hydrological regions in Brazil (Fig. 1). Diagonal lines indicate significant trends at 95% level. The vertical purple lines refer to ESMs from Eta-INPE datasets

Fig. 6 Trends ($^{\circ}\text{C}/\text{decade}$) in hottest days (TXx) at the annual scale for 21 NEX-GDDP climate models (1–21), 4 Eta-INPE Models (22–25), MME-Sta (26), MME-Dyn (27), GMFD (28) and OBS-BR (29; gray rectangle) from 1980 to 2005. Hatching indicates where trends are significant at the 95% level



to the observations. MIROC family of models generally has contradictory trends in many indices (e.g., TXx, DTR, and TX90p) compared to observations, particularly MIROC-ESM-CHEM.

Few downscaled models capture even moderately well the diurnal temperature range (DTR; Fig. 5c models 1, 2, 3, and 5) trends in most hydrological basins. In the case of the Eta-INPE models, none are able to replicate even the sign of the trend in all basins. In fact, the GMFD dataset also shows DTR trends slightly different from OBS-BR. This is possibly because DTR is highly affected by land surface

characteristics, which are both transient in time and very heterogeneous inside the grid cells of climate models for both the GCMs (> 100 km of horizontal resolution) and Eta (20 km). This affects both the Eta-INPE models, which contain raw GCM output, and the NEX-GDDP models, for which the downscaling procedure explicitly attempts to conserve the GCM modeled trends (Thrasher et al. 2012). Maximum and minimum temperatures in GMFD are affected by both the underlying NCEP-NCAR reanalysis model and the monthly average DTR of the CRU dataset, which uses fewer meteorological stations in the region than OBS-BR.

Finally, our results suggest that the better alternative for estimating the sign and magnitude of the temperature indices at the annual and seasonal scales is the use of the downscaled model ensembles (MME-Sta and MME-Dyn).

3.2 Precipitation indices

3.2.1 Evaluation metrics

For the annual total wet-day precipitation index (PRCP-TOT; Fig. 7), all statistically downscaled models show low bias (close to zero), especially for ACCESS1-0 (1),

CESM1-BGC (6), and NorESM1-M (21). The dynamically downscaled models show less precipitation in the North region and slightly higher in the South region respect to OBS-BR (Fig. 8a; first column).

Most of the downscaled models underestimate the observed values for intensity indices such as the annual maximum 1-day (RX1 day) and the maximum 5-day precipitation amount (RX5 day), especially in the North Atlantic basin (Fig. 8a; second and third column). Besides, models from statistically downscaled models (NEX-GDDP) overestimate the OBS-BR values, especially for the Tocantins River basin. Moreover, basins located in the South and Southern regions

Fig. 7 Climatology bias for the annual total wet-day precipitation—PRCPTOT (mm) for 21 statically (NEX-GDDP; models 1–21) and 4 dynamically (Eta-INPE; models 22–25) downscaled models, MME-Sta (26), MME-Dyn (27), and GMFD (28) from 1980 to 2005. Climatology for PRCPTOT in the observations dataset (OBS-BR; gray rectangle dataset 29) for 1980–2005

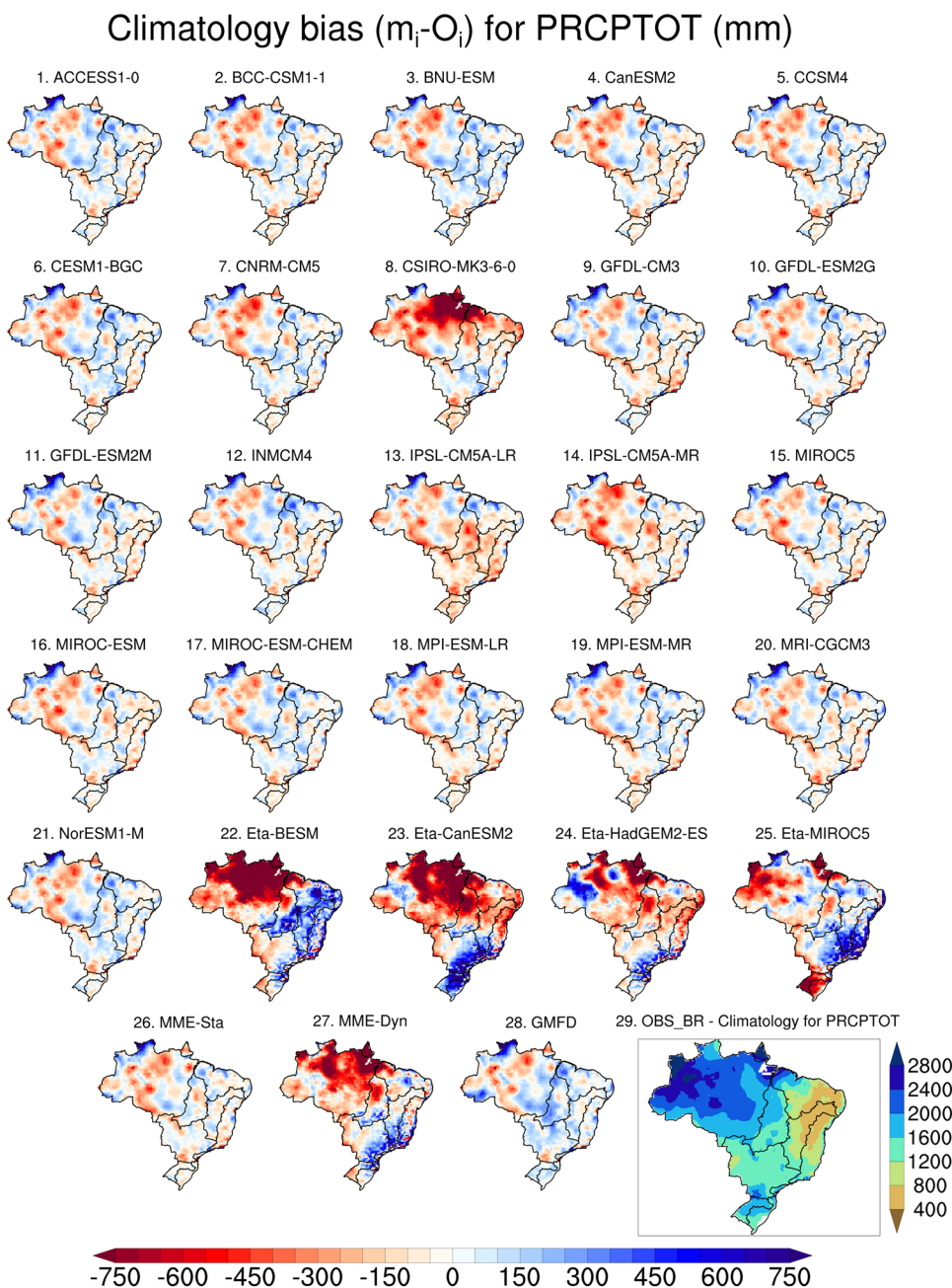
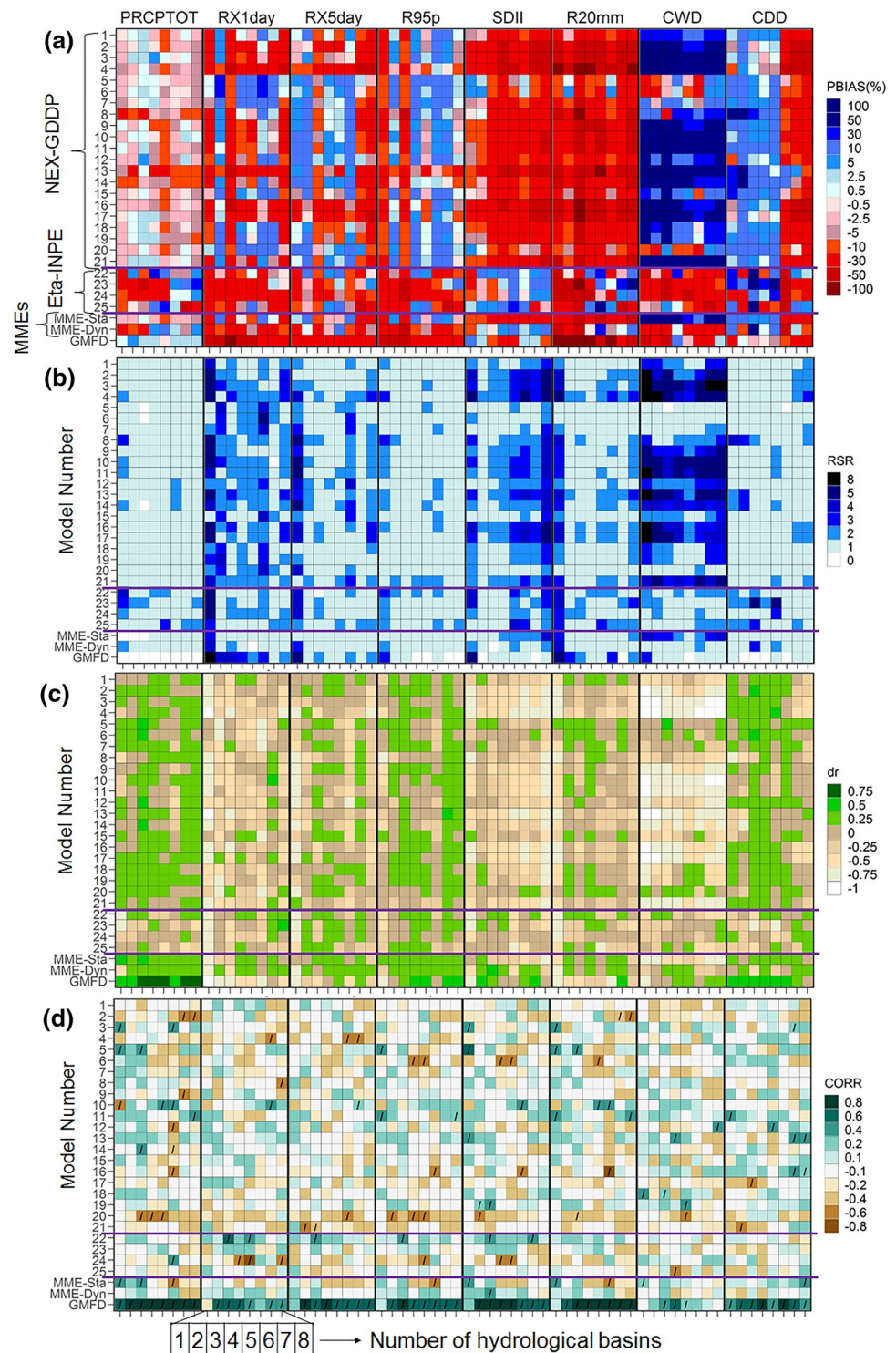


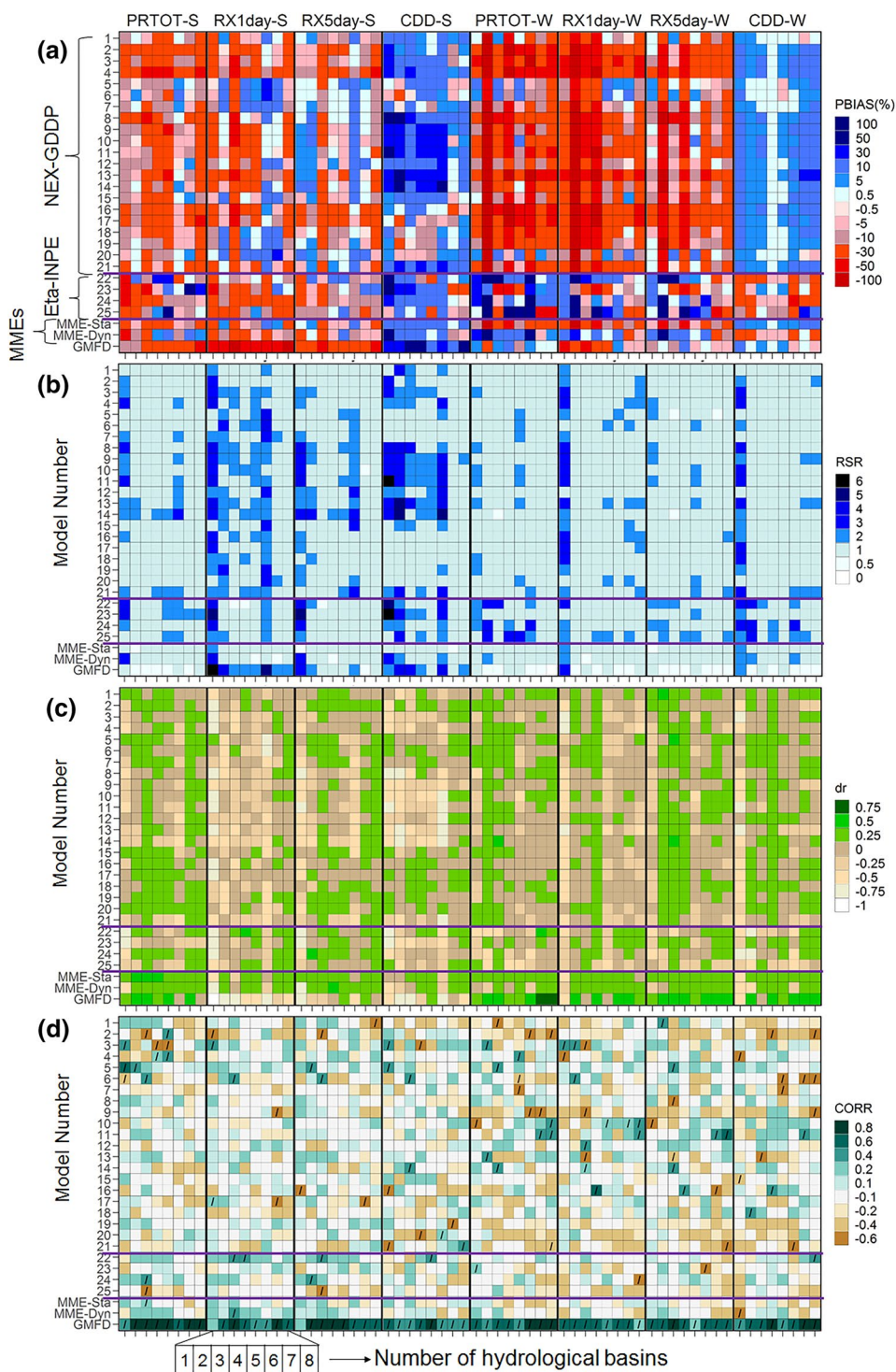
Fig. 8 Statistics of performance obtained for annual precipitation indices for statically (NEX-GDDP) and dynamically (Eta-INPE) downscaled models, MME-Sta, MME-Dyn, and GMFD from 1980 to 2005 over eight hydrological basins (Fig. 1). **a** Percent Bias (PBIAS); **b** RMSE-observations standard deviation ratio (RSR); **c** a refined index of model performance (d_r); **d** Correlation coefficients (CORR; the diagonal lines indicate significant correlations at 95% level). The horizontal purple lines refer to Eta-INPE datasets. For PBIAS and RSR, dark colors indicate models that perform worse than others, on average, and light colors indicate models that perform better than others, on average. Furthermore, for d_r and CORR, dark (light) colors show models that have better (worse) statistical metrics than others, on average



of Brazil show good performance according to RSR and d_r . Additionally, for the very wet days (R95p) index, all evaluation metrics show poor performance for all models over the Amazon River basin (Fig. 8; fourth column). On the other hand, the dynamically downscaled models from the Eta-INPE dataset tend to underestimate the R95p index for

almost all basins. We note that summer and winter indices (e.g., PRCPTOT, RX1day, and RX5day) are generally underestimated across all Brazil for almost all downscaled models except for Eta-INPE models (models 22, 23, 24, and 25) that show wet bias in winter over most of the watersheds (Fig. 9a–d). Similar to the annual scale (Fig. 8a), the weak

Fig. 9 As in Fig. 8, but for extreme precipitation indices in summer (DJF) and winter (JJA). The nomenclature of the ETCCDI indices was adapted to Index—“S” for summer and Index—“W” for winter



performance of downscaled ESMs is more evident for the Amazon basin.

In almost all basins, the statistically downscaled ESMs models underestimate the simple daily intensity index (SDII) and the number of very heavy precipitation day (R 20 mm) indices (see fifth and sixth columns of Fig. 8). For these indices, the performance of the Eta-INPE dataset is better

than NEX-GDDP. The *PBIAS* shows that the simulations underestimate the observed values for the Amazon River and overestimate in Uruguay River and South Atlantic basins. In general, for all downscaled ESMs, the poorest performance (*RSR*, d_r and *CORR*) is found over the Amazon River basin.

For the duration indices like consecutive dry days (CDD) and consecutive wet days (CWD) (see last two

columns of Fig. 8), some models show the largest disagreement when compared with the observed dataset, and thus indicate considerable uncertainty. For instance, models 8, 13, 14, and 23 are generally too dry while others too wet (models 2, 3, 4, 11, 13, 16, and 17) over the North and Northeast of Brazil. The statistically downscaled ESMs show worse performance over the Amazon River, Tocantins Rivers, and the North Atlantic basin (Fig. 8). On the other hand, some models such as CCSM4 (5) and CESM1-BGC (6) have relatively good performance in the Central-West, Southeast, and South of Brazil. Downscaled NEX-GDDP models show better skill in simulating the CDD index at seasonal scale than ETA-INPE models. Noteworthy, statistically downscaled ESMs have better scores (Fig. 9) in simulating CDD index in winter than summer (see models 8, 10, 13, 14 in Fig. 9).

Comparison between observations (OBS-BR) and the reanalyses shows that the GMFD dataset underestimates approximately all precipitation indices at the annual scale (see dataset 28 of Fig. 7), except for PRCPTOT as the *PBIAS* varies between 0 and 6% (see bottom of Fig. 8a–d). However, in general, the RX1day, RX5day, and R95p indices are overestimated for all basins (Fig. 8a). The results do not indicate a dominant positive or negative pattern of *PBIAS* for SDII, R20mm, CDD, and CWD. It should be noted that the worst performance is found over the Amazon River, Tocantins Rivers, and North Atlantic basins (Fig. 8). In this sense, the main discrepancies between OBS-BR and GMFD are found for several indices such as RX1day, RX5day, R20mm, and CWD (Fig. 8). Figure 9a shows a consistently dry bias in PRCPTOT, RX1day, RX5day, and CDD indices during the summer and low skill in reproducing intensity indices (RX1day, RX5day). GMFD also shows a better winter precipitation indices estimation over most parts of Brazil, according to *RSR*, *dr*, and *CORR* (Fig. 9b–d). Furthermore, the Amazon River basin is poorly represented in GMFD for the annual, summer, and winter for almost all precipitation indices, except for PRCPTOT and CDD (see Figs. 8, 9).

The overall performance assessment (see bottom of Fig. 8) shows that the models from NEX-GDDP and Eta-INPE underestimate precipitation intensity (RX1day and R95p) and frequency (R20mm) over the Amazon River basin. However, the statistically downscaled models perform better for the PRCPTOT and CDD indices on the both annual and seasonal scale (Figs. 8, 9). The relative errors could be because PRCPTOT and CDD are less dependent on fine scale phenomena than the indices that represent extreme precipitation events (e.g., RX1day and RX5day). Besides, the coarse resolution of the underlying ESMs makes them have special difficulties in representing the spatial and temporal heterogeneity of precipitation over tropical regions (Marengo et al. 2010b; Rusticucci et al. 2010; Sillmann et al. 2013a).

Of particular importance is the fact that for several models and regions, the sign of the bias in CDD is different in the annual and seasonal scales. For example, most models show a negative *PBIAS* (fewer dry days) at the annual scale, but a positive (more dry days) *PBIAS* in both summer and winter seasons in regions more to the south (e.g., models 6, 7 and 8). Transition seasons (spring, autumn) have a larger influence on the overall annual number of precipitation days across these higher-latitude regions of Brazil (Rao et al. 2016). The opposite is true for some statistically downscaled models in other regions, and the sign of the CDD bias is also reversed between summer and winter in the dynamically downscaled models. Since some activities such as agricultural production are particularly sensitive to dry spells in specific seasons (e.g., da Silva et al. 2013) special care should be taken when selecting downscaled models for this kind of application.

The MMEs have weakest representation of intensity indices principally over the Amazon basin at the annual and seasonal levels. Multi-model ensembles generally have a better performance than most individual models, but not all. Our results show that MME-Sta might be a better approach in precipitation indices (e.g., PRCPTOT and RX5day) over the Amazon River, where most models show poor performance (Figs. 8, 9). On the other hand, the SDII, R20mm, and CWD values from MMEs-Dyn generally agree more with the observations than MME-Sta over most hydrological basins. The MMEs-Sta and MME-Dyn overestimate and underestimate CWD and CDD, respectively, particularly over the Amazon, Tocantins, and North Atlantic basin. It should be highlighted that the bias is significantly smaller in the Eta simulations.

In general, the dynamically downscaled models simulate less total precipitation than OBS-BR, even for the NCEP-NCAR reanalysis used in GMFD. This underestimation by the Eta-simulations discussed here is consistent with the results obtained by Chou et al. (2014a) and Valverde and Marengo (2014), especially in northern Brazil. The agreement is generally much better for the statistically downscaled models, although the sign of the errors has a similar spatial pattern, with modest underestimation of total precipitation in northern Brazil. All downscaled models capture the main spatial features of extreme precipitation indices climatology, but significant biases were found, particularly in the Amazon River basin (Figs. 8, 9). The systematic rainfall underestimation by the models can be related to many factors, such as the poor representation of cumulus convection, the biosphere–atmosphere interactions in the rainforest, soil moisture, and land surface processes (Torres and Marengo 2013; Yin et al. 2013). For example, representation of aerosol-related processes is a major source of uncertainty on climate models (Seinfeld et al. 2016), and precipitation extremes are particularly affected by it (Lin et al. 2018).

On the other hand, there is poor data observation coverage in some portions of South America, mainly in the Amazon Basin, in which few meteorological stations are available. This influences the magnitude and location of the bias patterns, mainly for precipitation (Torres and Marengo 2013).

3.2.2 Trend analysis in precipitation indices

Most of the climate trend analysis in precipitation extremes in Brazil have focused on specific basins in southern (Donat et al. 2013; Skansi et al. 2013; Carvalho et al. 2014; Ávila et al. 2016; Murara et al. 2018) or northern and northeastern Brazil (Oliveira et al. 2014, 2017; Valverde and Marengo 2014; Bezerra et al. 2019). It is quite challenging to compare these studies with ours since they included small areas and many factors can influence trends (e.g., study period, weather stations, data quality control, homogeneity and trend estimation methods). However, our findings are in line with the results of the prevalence of regions with an upward trend in the annual (Fig. 10) and summer maximum

daily rainfall. The interested reader should refer to Fig. S16 to the trends for the selected indices (PRCPTOT, RX1day, RX5day, and CDD) at the seasonal scale. Also, the positive trends in consecutive dry days are generally in line with those of Valverde and Marengo (2014) for southern Amazon, Upper São Francisco, Tocantins, and northern Paraná basins (Fig. 10).

Brazil-wide trends in precipitation indices are generally not significant for OBS-BR and GMFD (Fig. 10). Some hydrological basins have the same patterns, mainly showing decreases in PRCPTOT and CWD and some increases in CDD (Figs. 10, 11), especially in northeastern, southeastern, and southern Brazil. Also, results for the CDD index in winter and summer indicate dry trends in many downscaled models across the southern watersheds (e.g., PAR, URU, and SAR). The extreme precipitation indices display mixed signal trends and show less agreement between the different datasets than the temperature indices in both annual and seasonal scales. The precipitation trends in GFDL-ESM2G (10) and Eta-HadGEM2-ES (24) are especially troublesome

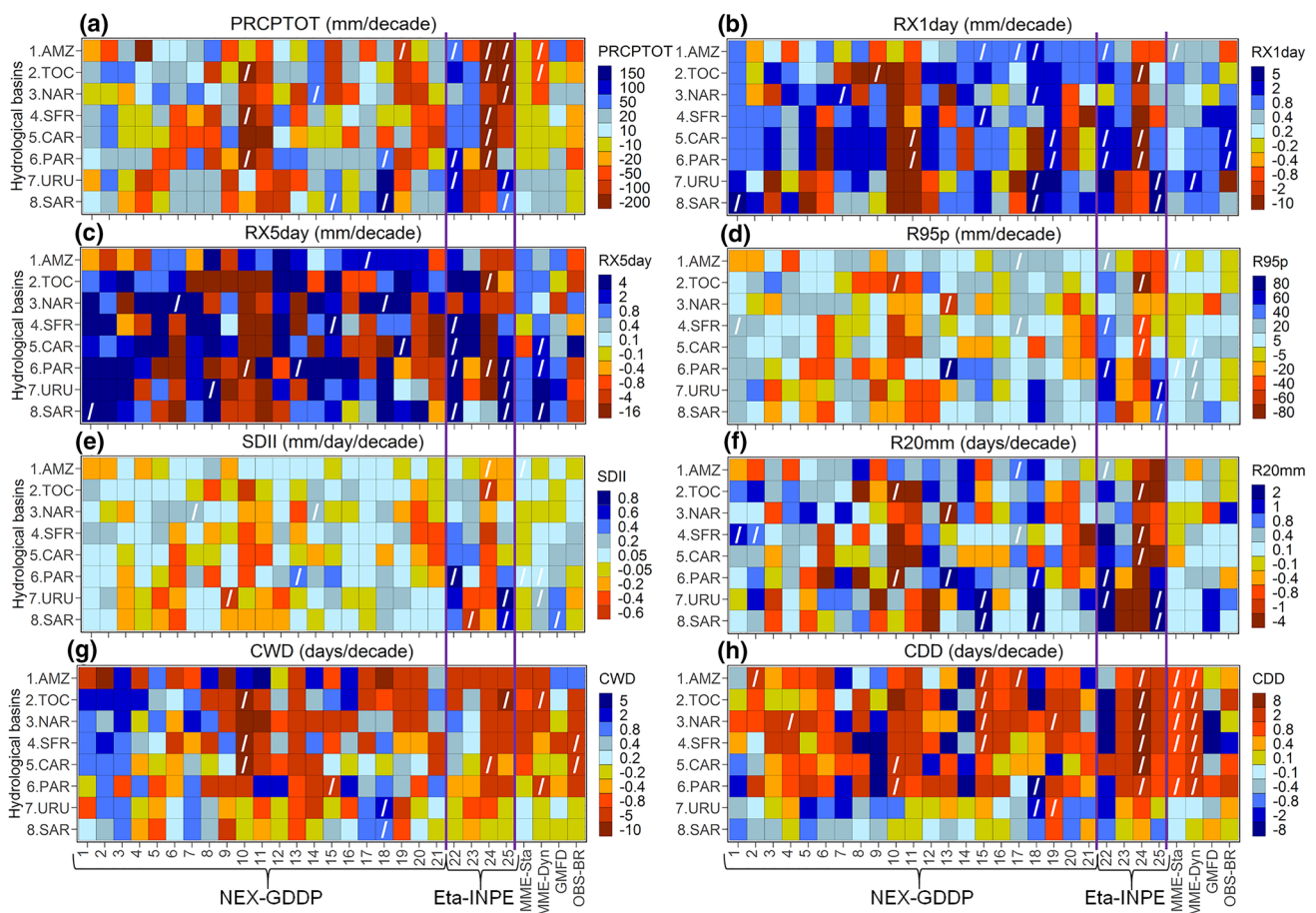
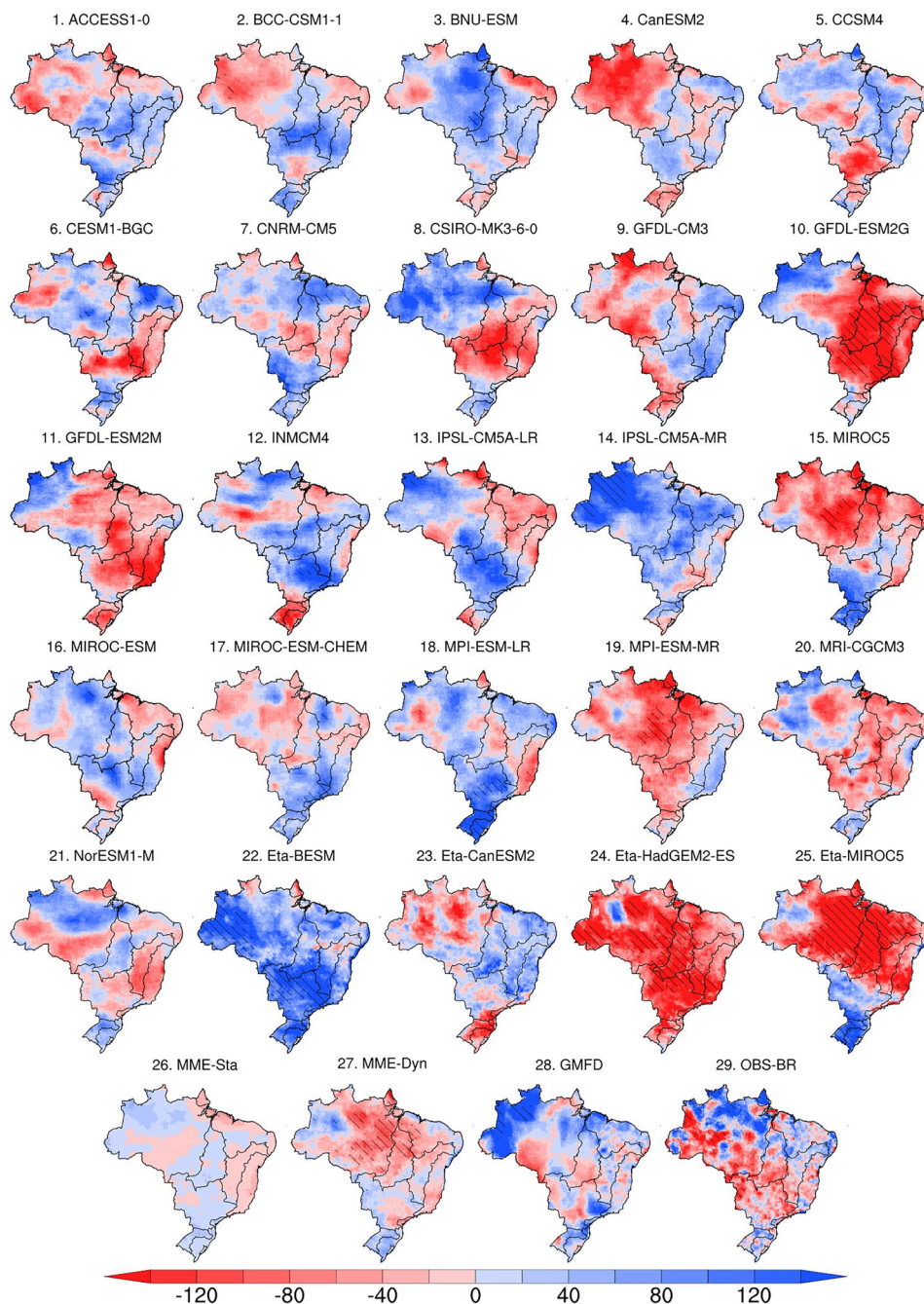


Fig. 10 Trends per decade from 1980 to 2005 for precipitation indices at the annual scale **a–h** for 21 NEX-GDDP climate models (1–21), 4 Eta-INPE climate models (22–25), MME-Sta (26), MME-

Dyn (27), GMFD (28) and OBS-BR over eight hydrological regions in Brazil (Fig. 1). Diagonal lines indicate significant trends at 95% level. The vertical purple lines refer to ESMs from Eta-INPE datasets

Fig. 11 Trends (mm/decade) in annual total wet-day precipitation (PRCPTOT) for 21 NEX-GDDP climate models (1–21), 4 Eta-INPE climate models (22–25), MME-Sta (26), MME-Dyn (27), GMFD (28) and OBS-BR (29; gray rectangle) from 1980 to 2005. Hatching indicates where trends are significant at the 95% level



(see Figs. 10, 11) in annual and seasonal trends, suggesting a much stronger drying trend than OBS-BR and other downscaled ESMs. Moreover, MMEs appear to agree better with trends in OBS-BR than trends in GMFD precipitation indices.

In general, there is not a single model that is the most appropriate to represent the observed trends for each index over the basins in both annual and seasonal temporal scales for the period 1980–2005. Although trend patterns vary widely across datasets (21 NEX-GDDP climate models, 4 Eta-INPE models, MMEs, GMFD, and OBS-BR), especially

for precipitation indices, the multi-model ensembles are a good alternative to better capture observed trends.

3.3 The comprehensive model rank (M_R)

Table 3 provides the ranking for all models analyzed using 16 climate indices at the annual scale over eight hydrological basins throughout Brazil. In terms of temperature indices, the best models for the whole domain are, in order, CSIRO-MK3-6-0 (8) and CNRM-CM5 (7); these are the only models with $M_R \geq 0.85$. The models with the lowest

Table 3 Ranking of downscaled ESMs and MMEs for temperature and precipitation indices at the annual scale over Brazil

Models and MMEs (ID)	Overall ranking (average of $M_{R\text{-temperature}}$ and $M_{R\text{-precipitation}}$)		Temperature indices		Precipitation indices	
	Rank	Skill score (M_R)	Rank	M_R	Rank	M_R
MME-Sta (26)	1	0.927	1	0.956	1	0.899
CNRM-CM5 (7)	2	0.877	3	0.883	4	0.870
MME-Dyn (27)	3	0.872	4	0.877	5	0.868
CCSM4 (5)	4	0.863	8	0.827	2	0.898
MRI-CGCM3 (20)	5	0.859	5	0.844	3	0.875
MPI-ESM-MR (19)	6	0.835	6	0.832	9	0.838
MIROC5 (15)	7	0.830	10	0.810	8	0.850
CESM1-BGC (6)	8	0.823	16	0.781	6	0.865
ACCESS1-0 (1)	9	0.822	13	0.790	7	0.853
CSIRO-MK3-6-0 (8)	10	0.819	2	0.887	21	0.751
BCC-CSM1-1 (2)	11	0.797	7	0.830	17	0.763
Eta-BESM (22)	12	0.791	19	0.771	12	0.810
Eta-HadGEM2-ES (24)	13	0.786	12	0.794	14	0.779
MPI-ESM-LR (18)	14	0.786	21	0.739	10	0.834
GFDL-ESM2G (10)	15	0.780	15	0.784	15	0.775
GFDL-CM3 (9)	16	0.779	11	0.796	18	0.761
INMCM4 (2)	17	0.778	21	0.739	11	0.818
BNU-ESM (3)	18	0.774	9	0.825	24	0.723
IPSL-CM5A-MR (14)	19	0.753	18	0.773	22	0.732
Eta-MIROC5 (25)	19	0.753	23	0.736	16	0.769
NorESM1-M (21)	21	0.750	17	0.776	23	0.724
IPSL-CM5A-LR (13)	22	0.744	14	0.789	26	0.699
MIROC-ESM-CHEM (17)	23	0.741	24	0.731	20	0.752
Eta-CanESM2 (23)	24	0.736	27	0.693	13	0.780
MIROC-ESM (16)	25	0.733	20	0.751	25	0.715
GFDL-ESM2M (11)	26	0.730	26	0.707	19	0.754
CanESM2 (4)	27	0.705	25	0.724	27	0.687

Downscaled models or MMEs in bold achieve a skill score ≥ 0.85 . The optimal value of M_R is 1.0

M_R are Eta-CanESM2 (23) and GFDL-ESM2M (11). When considering the precipitation indices, the top three models are CCSM4 (5) followed by MRI-CGCM3 (20), and CNRM-CM5 (7), whereas models with the worst M_R are MIROC-ESM (16), IPSL-CM5A-LR (13) and CanESM2 (4). Considering all climate indices over all basins, the best individual models are CNRM-CM5 and CCSM4, followed by MRI-CGCM3, and the worst on the overall ranking are MIROC-ESM (16), GFDL-ESM2M (11) and CanESM (4). Furthermore, analyzing the country as a whole (Table 3), the multi-model ensemble of NEX-GDDP models (MME-Sta, $M_{R\text{overall}}=0.927$) generally leads to better skill scores than individual models and ensemble of Eta-INPE models (MME-Dyn, $M_{R\text{overall}}=0.872$).

Furthermore, the ranking of the downscaled ESMs obtained at the seasonal scale is very similar to those presented at the annual scale. For instance, the best models are MRI-CGCM3, CNRM-CM5, and CCSM4, using the overall ranking of the selected temperature (TXx, TNn, TN10p, and

TX90p) and precipitation (PRCPTOT, RX1day, RX5day, CDD) indices. Also, the MMEs-Sta performed better than MMEs-Dyn and individual downscaled ESMs in both summer and winter. Being aware of these results, we decided to emphasize the ranking discussion on an annual scale. Readers interested in the ranking for summer and winter can refer to Table S1.

It is important to note that the top three models in Table 3 have a native horizontal resolution finer than $2^\circ \times 2^\circ$ —latitude/longitude (Table 1), which could indicate that a finer resolution allows the models to resolve better processes associated with climate extremes. Although models with coarser resolutions do tend to perform poorly, having a finer resolution is not necessarily a determining factor to choose the best performing model. For instance, the downscaled results of models with very fine native horizontal resolutions (e.g., CESM1-BGC: $0.924^\circ \times 1.250^\circ$) do not perform better in temperature indices than coarser resolution models such as BNU-ESM ($2.8^\circ \times 2.8^\circ$) or BCC-CSM1-1 ($2.8^\circ \times 2.8^\circ$).

However, this association is stronger when considering precipitation indices, as the top three models have native horizontal resolutions less than $1.5^\circ \times 1.5^\circ$ (Table 3), and the worst ones greater than $2^\circ \times 2^\circ$ such as CanESM2 ($2.8^\circ \times 2.8^\circ$) and MIROC-ESM ($2.791^\circ \times 2.813^\circ$). This is likely due to the higher spatial heterogeneity of the precipitation and has also been observed with raw ESM results over Australia (Alexander and Arblaster 2017) and East Asia (Kusunoki and Arakawa 2015). The number and extent of vertical layers in the model does not seem to be an important factor for either temperature or precipitation indices, as was previously observed over higher altitude regions such as the Equatorial Andes (Campozano et al. 2017).

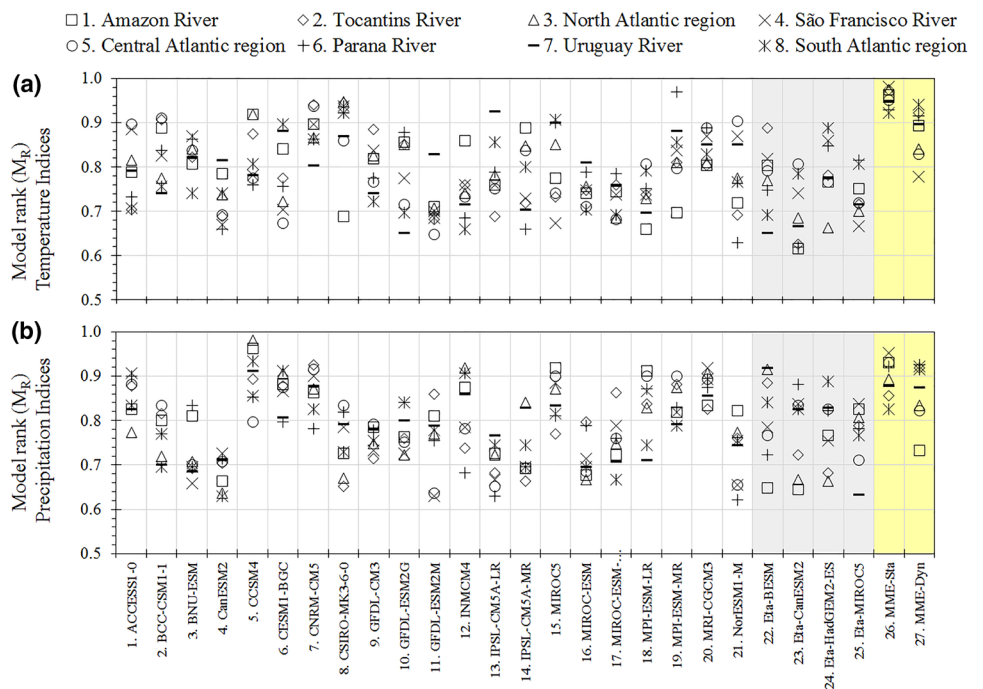
Although the ensembles generally perform better in a larger number of basins than individual models, some models are better than both ensembles for some particular basins, especially for precipitation (Fig. 12). For example, although MME-Sta ranks better than most models for precipitation for the South Atlantic basin, individual models like CCSM4 (5), CESM1-BGC (6), and INMCM4 (12) rank considerably better than the ensemble. Although an improvement over most individual dynamically downscaled models for precipitation in most basins, MME-Dyn does not rank better than the best model in half of the basins and ranks especially poorly in the Amazon River basin. For temperature indices, using MMEs more consistently leads to better results than individual models, though not always. For example, individual models such as MPI-ESM-MR (19) show the highest values of M_R among models and ensembles for the Parana River basin. It is

important to note that MME-Dyn is considerably worse than MME-Sta for most basins. However, MME-Dyn ranks better in the South Atlantic basin at both temperature and precipitation indices, and in the Tocantins, Parana, and South Atlantic basins for precipitation indices.

Noteworthy, the most successful downscaling simulations based on the Eta Regional Climate Model, are the ones forced by BESM (22) and HadGEM2-ES (24) (Table 3). The Eta-HadGEM2-ES appear to be better than Eta-BESM for temperature indices over hydrological basins located in the southern part of Brazil, and worse for precipitation indices over basins on the northeast of the country (Fig. 11).

The large difference in the number of models among different datasets used in the ensemble mean complicates a proper comparison between the dynamical and statistical downscaling techniques. Although none of the dynamically downscaled models are among the best in the overall ranking (Table 3), they perform very well in some aspects. Eta-BESM, for example, is ranked the best for precipitation indices in the Uruguay River basin, although it is ranked the worst for temperature indices in the same basin (Fig. 12). A more useful comparison can be made using the ESMs that were downscaled using both techniques, CanESM2 and MIROC5, but also show that one technique is not necessarily better than the other for evaluating climate extremes. Although the statistically downscaled CanESM2 is among the worst ranking models in all indices, the dynamically downscaled version performs reasonably well in all basins, except for the Amazon and Tocantins basins. On the other hand, the statistically downscaled MIROC5 is generally

Fig. 12 Model rank (MR) value for temperature (a) and precipitation indices (b) at the annual scale. Each symbol represents a given basin. White, gray and yellow areas refer to 21 NEX-GDDP climate models, four (4) Eta-INPE climate models and multi-model ensembles (MMEs), respectively



better than its dynamically downscaled counterpart, except for the Tocantins River basin.

4 Summary and conclusions

This paper provides an overview of the performance of 25 downscaled Earth System Models, generated by statistical (NEX-GDDP) and dynamical (Eta-INPE) downscaling techniques, to evaluate extreme climate indices during historical climate (1980–2005) over eight hydrological basins across Brazil. Performance was evaluated for annual and seasonal indices (summer and winter) by contrasting simulations with an observational gridded dataset at high horizontal resolution.

The GMFD dataset used as reference for the statistical downscaling is problematic for precipitation over the Amazon River basin in both annual and seasonal scale, with little capacity to simulate the climatology and temporal variability of most precipitation indices, except for PRCPTOT and CDD. GMFD also tends to reproduce much higher TXx and lower WSDI than the observed values, and shows trends with the wrong sign (positive or negative) for several indices and basins. These discrepancies point to the possibility of improvement of statistically downscaled products for Brazil by using denser observational networks as reference.

Although the CNRM-CM5, CCSM4, and MRI-CGCM3 (NEX-GDDP models) statistically downscaled products have the best results among individual models in an overall comparison for Brazil for annual, summer and winter indices, the results varied widely among basins. Finer horizontal resolutions of the original ESMs appear to be somewhat related, but not determinant, to the performance of the downscaled product in representing extreme climate events, especially precipitation. The use of multi-model ensembles, although improving the overall representation, does not always lead to the best results depending on the region considered. The multi-model ensembles also show considerable discrepancies, especially across northern Brazil, in several extremes climate indices, particularly ones related to the persistence of climate events such as cumulative wet and dry days and warm spell duration. These conclusions are generally valid at both annual and seasonal scales. However, some models and regions present conflicting behaviors at the annual scale and in different seasons, especially for consecutive dry days (CDD). Caution must be taken when selecting model products for applications that are particularly sensitive to extremes in specific seasons.

The downscaled ESMs appear to compare better with OBS-BR in terms of trend patterns than the GMFD dataset. Furthermore, downscaled product trends are much more spatially coherent in temperature than precipitation indices when compared with the observational dataset. In this sense,

the trend pattern in most climate indices is generally better captured by multi-model ensembles than individual downscaled ESMs (especially for precipitation indices).

In conclusion, despite some models being generally better than others, no single downscaled product or ensemble is the best choice for every region. The results presented in this paper can guide researchers in choosing the best data for each particular application, as well as inform climate modelers about the shortcomings of models and downscaling approaches over Brazil.

Acknowledgements The authors would like to thank the Universidade Federal de Viçosa. This work was supported by Minas Gerais Research Foundation (FAPEMIG) and to the Coordination for the Improvement of Higher Education Personnel (CAPES). The authors thank to the Byrd Polar and Climate Research Center. The authors are grateful to the climate simulations used from National Institute for Space Research (INPE), and we thank NEX-GDDP dataset, prepared by the Climate Analytics Group and NASA Ames Research Center using the NASA Earth Exchange, and distributed by the NASA Center for Climate Simulation (NCCS). The authors further thank Alexandre Xavier, Carey King, and Bridget Scanlon that provided a gridded observational dataset used in this work.

References

- Aerenson T, Tebaldi C, Sanderson B, Lamarque J-F (2018) Changes in a suite of indicators of extreme temperature and precipitation under 1.5 and 2 degrees warming. *Environ Res Lett* 13:035009. <https://doi.org/10.1088/1748-9326/aaafd6>
- Aguilar E, Peterson T, Obando P et al (2005) Changes in precipitation and temperature extremes in Central America and northern South America, 1961–2003. *J Geophys Res* 110:D23107. <https://doi.org/10.1029/2005JD006119>
- Alexander L, Arblaster J (2017) Historical and projected trends in temperature and precipitation extremes in Australia in observations and CMIP5. *Weather Climate Extrem* 15:34–56. <https://doi.org/10.1016/j.wace.2017.02.001>
- Alexander L, Zhang X, Peterson T et al (2006) Global observed changes in daily climate extremes of temperature and precipitation. *J Geophys Res* 111:1–22. <https://doi.org/10.1029/2005jd006290>
- Ambrizzi T, Reboita M, da Rocha R, Llopart M (2019) The state-of-the-art and fundamental aspects of regional climate modeling in South America. *Ann N Y Acad Sci* 1436:98–120. <https://doi.org/10.1111/nyas.13932>
- Ávila A, Justino F, Wilson A et al (2016) Recent precipitation trends, flash floods and landslides in southern Brazil. *Environ Res Lett* 11:114029. <https://doi.org/10.1088/1748-9326/11/11/114029>
- Avila A, Guerrero F, Escobar Y, Justino F (2019) Recent precipitation trends and floods in the Colombian Andes. *Water* 11:379. <https://doi.org/10.3390/w11020379>
- Avila-Diaz A, Benezoli V, Justino F et al (2020) Assessing current and future trends of climate extremes across Brazil based on Reanalyses and Earth System Model projections. *Climate Dyn* (**in press**)
- Bezerra B, Silva L, Santos e Silva C, de Carvalho G (2019) Changes of precipitation extremes indices in São Francisco River Basin, Brazil from 1947 to 2012. *Theor Appl Climatol* 135:565–576. <https://doi.org/10.1007/s00704-018-2396-6>
- Boulanger J, Martinez F, Segura E (2006) Projection of future climate change conditions using IPCC simulations, neural networks and

- Bayesian statistics. Part 1: temperature mean state and seasonal cycle in South America. *Climate Dyn* 27:233–259. <https://doi.org/10.1007/s00382-006-0182-0>
- Boulanger J, Martinez F, Segura E (2007) Projection of future climate change conditions using IPCC simulations, neural networks and Bayesian statistics. Part 2: precipitation mean state and seasonal cycle in South America. *Climate Dyn* 27:233–259. <https://doi.org/10.1007/s00382-006-0182-0>
- Bozkurt D, Rojas M, Boisier J et al (2019) Dynamical downscaling over the complex terrain of southwest South America: present climate conditions and added value analysis. *Climate Dyn* 53:6745–6767. <https://doi.org/10.1007/s00382-019-04959-y>
- Brito S, Cunha A, Cunningham C et al (2018) Frequency, duration and severity of drought in the Semiarid Northeast Brazil region. *Int J Climatol* 38:517–529. <https://doi.org/10.1002/joc.5225>
- Burger G, Murdock Q, Werner A et al (2011) Downscaling extremes—an intercomparison of multiple statistical methods for present climate. *J Climate* 25:4366–4388. <https://doi.org/10.1175/JCLI-D-11-00408.1>
- Campozano L, Vázquez-Patiño A, Tenelanda D et al (2017) Evaluating extreme climate indices from CMIP3&5 global climate models and reanalysis data sets: a case study for present climate in the Andes of Ecuador. *Int J Climatol* 37:363–379. <https://doi.org/10.1002/joc.5008>
- Carvalho J, Assad E, de Oliveira A, Pinto H (2014) Annual maximum daily rainfall trends in the midwest, southeast and southern Brazil in the last 71 years. *Weather Climate Extrem* 5:7–15. <https://doi.org/10.1016/j.wace.2014.10.001>
- Cheng C, Knutson T (2008) On the verification and comparison of extreme rainfall indices from climate models. *J Climate* 21:1605–1621. <https://doi.org/10.1175/2007JCLI1494.1>
- Chou S, Lyra A, Mourão C et al (2014a) Evaluation of the eta simulations nested in three global climate models. *Am J Climate Change* 03:438–454. <https://doi.org/10.4236/ajcc.2014.35039>
- Chou S, Lyra A, Mourão C et al (2014b) Assessment of climate change over South America under RCP 4.5 and 8.5 downscaling scenarios. *Am J Climate Change* 03:512–527. <https://doi.org/10.4236/ajcc.2014.35043>
- Coelho C, Cardoso D, Firpo M (2016a) Precipitation diagnostics of an exceptionally dry event in São Paulo, Brazil. *Theor Appl Climatol* 125:769–784. <https://doi.org/10.1007/s00704-015-1540-9>
- Coelho C, de Oliveira C, Ambrizzi T et al (2016b) The 2014 southeast Brazil austral summer drought: regional scale mechanisms and teleconnections. *Climate Dyn* 46:3737–3752. <https://doi.org/10.1007/s00382-015-2800-1>
- da Silva V, da Silva B, Albuquerque W et al (2013) Crop coefficient, water requirements, yield and water use efficiency of sugarcane growth in Brazil. *Agric Water Manag* 128:102–109. <https://doi.org/10.1016/j.agwat.2013.06.007>
- Darela-Filho J, Lapola D, Torres R, Lemos M (2016) Socio-climatic hotspots in Brazil: how do changes driven by the new set of IPCC climatic projections affect their relevance for policy? *Clim Change* 136:413–425. <https://doi.org/10.1007/s10584-016-1635-z>
- Debortoli N, Camarinha P, Marengo J, Rodrigues R (2017) An index of Brazil's vulnerability to expected increases in natural flash flooding and landslide disasters in the context of climate change. *Nat Hazards* 86:557–582. <https://doi.org/10.1007/s11069-016-2705-2>
- Dereczynski C, Silva W, Marengo J (2013) Detection and projections of climate change in Rio de Janeiro, Brazil. *Am J Climate Change* 02:25–33. <https://doi.org/10.4236/ajcc.2013.21003>
- Donat M, Alexander L, Yang H et al (2013) Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: the HadEX2 dataset. *J Geophys Res Atmos* 118:2098–2118. <https://doi.org/10.1002/jgrd.50150>
- Dosio A, Jones R, Jack C et al (2019) What can we know about future precipitation in Africa? Robustness, significance and added value of projections from a large ensemble of regional climate models. *Climate Dyn*. <https://doi.org/10.1007/s00382-019-04900-3>
- Espinoza J, Marengo J, Ronchail J et al (2014) The extreme 2014 flood in south-western Amazon basin: the role of tropical-subtropical South Atlantic SST gradient. *Environ Res Lett* 9:1–9. <https://doi.org/10.1088/1748-9326/9/12/124007>
- Giorgi F (2006) Climate change hot-spots. *Geophys Res Lett* 33:1–4. <https://doi.org/10.1029/2006GL025734>
- Gupta H, Sorooshian S, Yapo P (1999) Status of automatic calibration for hydrologic models: comparison with multilevel expert calibration. *J Hydrol Eng* 4:135–143. [https://doi.org/10.1061/\(ASCE\)1084-0699\(1999\)4:2\(135\)](https://doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135))
- Jiang Z, Li W, Xu J, Li L (2015) Extreme precipitation indices over China in CMIP5 models. Part I: model evaluation. *J Climate* 28:8603–8619. <https://doi.org/10.1175/JCLI-D-15-0099.1>
- Jones PW (1999) First- and second-order conservative remapping schemes for grids in spherical coordinates. *Mon Weather Rev* 127:2204–2210. [https://doi.org/10.1175/1520-0493\(1999\)127%3c2204:FASOCR%3e2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127%3c2204:FASOCR%3e2.0.CO;2)
- Kendall MG (1975) Rank correlation methods, 4th edn. Charles Griffin, London
- Knutti R, Furrer R, Tebaldi C et al (2010) Challenges in combining projections from multiple climate models. *J Climate* 23:2739–2758. <https://doi.org/10.1175/2009JCLI3361.1>
- Kusunoki S, Arakawa O (2015) Are CMIP5 models better than CMIP3 models in simulating precipitation over East Asia? *J Climate* 28:5601–5621. <https://doi.org/10.1175/JCLI-D-14-00585.1>
- Liao X, Xu W, Zhang J et al (2019) Global exposure to rainstorms and the contribution rates of climate change and population change. *Sci Total Environ* 663:644–653. <https://doi.org/10.1016/j.scitotenv.2019.01.290>
- Lin L, Wang Z, Xu Y et al (2018) Larger sensitivity of precipitation extremes to aerosol than greenhouse gas forcing in CMIP5 models. *J Geophys Res Atmos* 123:8062–8073. <https://doi.org/10.1029/2018JD028821>
- Loaiza W, Kayano M, Andreoli R et al (2020) Streamflow intensification driven by the Atlantic multidecadal oscillation (AMO) in the Atrato River Basin, Northwestern Colombia. *Water* 12:216. <https://doi.org/10.3390/w12010216>
- Lyra A, Tavares P, Chou S et al (2018) 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>
- Mann HB (1945) Nonparametric tests against trend. *Econometrica* 13:245–259. <https://doi.org/10.2307/1907187>
- Marengo J, Camargo C (2008) Surface air temperature trends in Southern Brazil for 1960–2002. *Int J Climatol* 28:893–904. <https://doi.org/10.1002/joc.1584>
- Marengo J, Jones B, Alves L et al (2009) Future change of temperature and precipitation extremes in south america as derived from the PRECIS regional climate modeling system. *Int J Climatol* 29:2241–2255. <https://doi.org/10.1002/joc.1863>
- Marengo J, Ambrizzi T, da Rocha R et al (2010a) Future change of climate in South America in the late twenty-first century: intercomparison of scenarios from three regional climate models. *Climate Dyn* 35:1089–1113. <https://doi.org/10.1007/s00382-009-0721-6>
- Marengo J, Rusticucci M, Penalba O, Renom M (2010b) An intercomparison of observed and simulated extreme rainfall and temperature events during the last half of the twentieth century: part 2: historical trends. *Clim Change* 98:509–529. <https://doi.org/10.1007/s10584-009-9743-7>
- Marengo J, Chou S, Kay G et al (2012) Development of regional future climate change scenarios in South America using the Eta CPTEC/HadCM3 climate change projections: climatology and regional analyses for the Amazon, São Francisco and the Paraná River

- basins. *Climate Dyn* 38:1829–1848. <https://doi.org/10.1007/s00382-011-1155-5>
- Marengo J, Torres R, Alves L (2017) Drought in Northeast Brazil—past, present, and future. *Theor Appl Climatol* 129:1189–1200. <https://doi.org/10.1007/s00704-016-1840-8>
- Marengo J, Alves L, Alvares R et al (2018) Climatic characteristics of the 2010–2016 drought in the semiarid Northeast Brazil region. *An Acad Bras Cienc* 90:1973–1985. <https://doi.org/10.1590/0001-3765201720170206>
- MCTI (2016) Third National Communication of Brazil to the United Nations framework convention on climate change. Ministry of Science Technology and Innovation, Brasilia
- Missirian A, Schlenker W (2017) Asylum applications respond to temperature fluctuations. *Science* 358:1610–1614. <https://doi.org/10.1126/science.aao0432>
- Moriassi D, Arnold J, Van Liew M et al (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans ASABE* 50:885–900. <https://doi.org/10.13031/2013.23153>
- Murara P, Acquavotta F, Garzena D, Fratianni S (2018) Daily precipitation extremes and their variations in the Itajaí River Basin, Brazil. *Meteorol Atmos Phys* 131:1145–1156. <https://doi.org/10.1007/s00703-018-0627-0>
- Mysiak J, Torresan S, Bosello F et al (2018) Climate risk index for Italy. *Phil Trans R Soc A*. <https://doi.org/10.1098/rsta.2017.0305>
- Nguyen P, Thorstensen A, Sorooshian S et al (2017) Evaluation of CMIP5 model precipitation using PERSIANN-CDR. *J Hydrometeorol* 18:2313–2330. <https://doi.org/10.1175/JHM-D-16-0201.1>
- Oliveira P, Silva C, Lima K (2014) Linear trend of occurrence and intensity of heavy rainfall events on Northeast Brazil. *Atmos Sci Lett* 15:172–177. <https://doi.org/10.1002/asl2.484>
- Oliveira P, Silva C, Lima K (2017) Climatology and trend analysis of extreme precipitation in subregions of Northeast Brazil. *Theor Appl Climatol* 130:77–90. <https://doi.org/10.1007/s00704-016-1865-z>
- Raghavan S, Hur J, Liang S (2018) Evaluations of NASA NEX-GDDP data over Southeast Asia: present and future climates. *Clim Change* 148:503–518. <https://doi.org/10.1007/s10584-018-2213-3>
- Rao V, Franchito S, Santo C, Gan M (2016) An update on the rainfall characteristics of Brazil: seasonal variations and trends in 1979–2011. *Int J Climatol* 36:291–302. <https://doi.org/10.1002/joc.4345>
- Ray D, Gerber J, Macdonald G, West P (2015) Climate variation explains a third of global crop yield variability. *Nat Commun* 6:1–9. <https://doi.org/10.1038/ncomms6989>
- Rosso F, Boiaski N, Ferraz S et al (2015) Trends and decadal variability in air temperature over Southern Brazil. *Am J Environ Eng* 5:85–95. <https://doi.org/10.5923/s.ajee.201501.12>
- Rusticucci M, Marengo J, Penalba O, Renom M (2010) An intercomparison of model-simulated extreme rainfall and temperature events during the last half of the twentieth century. Part 1: mean values and variability. *Clim Change* 98:493–508. <https://doi.org/10.1007/s10584-009-9742-8>
- Santos M, Fragoso M, Santos J (2017) Regionalization and susceptibility assessment to daily precipitation extremes in mainland Portugal. *Appl Geogr* 86:128–138. <https://doi.org/10.1016/j.apgeog.2017.06.020>
- Seinfeld J, Bretherton C, Carslaw K et al (2016) Improving our fundamental understanding of the role of aerosol–cloud interactions in the climate system. *Proc Natl Acad Sci* 113:5781–5790. <https://doi.org/10.1073/pnas.1514043113>
- Sen PK (1968) Estimates of the regression coefficient based on Kendall's tau. *J Am Stat Assoc* 63:1379–1389. <https://doi.org/10.2307/2285891>
- Sheffield J, Goteti G, Wood E et al (2006) Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *J Climate* 19:3088–3111. <https://doi.org/10.1175/JCLI3790.1>
- Sillmann J, Kharin V, Zhang X et al (2013a) Climate extremes indices in the CMIP5 multimodel ensemble: part 1. Model evaluation in the present climate. *J Geophys Res Atmos* 118:1716–1733. <https://doi.org/10.1002/jgrd.50203>
- Sillmann J, Kharin V, Zwiers W et al (2013b) Climate extremes indices in the CMIP5 multimodel ensemble: part 2. Future climate projections. *J Geophys Res Atmos* 118:2473–2493
- Skansi M, Brunet M, Sigró J et al (2013) Warming and wetting signals emerging from analysis of changes in climate extreme indices over South America. *Glob Planet Change* 100:295–307. <https://doi.org/10.1016/j.gloplacha.2012.11.004>
- Tang J, Niu X, Wang S et al (2016) Statistical downscaling and dynamical downscaling of regional climate in China: present climate evaluations and future climate projections. *J Geophys Res Atmos* 121:2110–2129. <https://doi.org/10.1002/2015JD023977>
- Thrasher B, Maurer E, McKellar C, Duffy P (2012) Technical note: bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrol Earth Syst Sci* 16:3309–3314. <https://doi.org/10.5194/hess-16-3309-2012>
- Tomasella J, Pinho P, Borma L et al (2013) The droughts of 1997 and 2005 in Amazonia: floodplain hydrology and its potential ecological and human impacts. *Clim Change* 116:723–746. <https://doi.org/10.1007/s10584-012-0508-3>
- Torres R, Marengo J (2013) Uncertainty assessments of climate change projections over South America. *Theor Appl Climatol* 112:253–272. <https://doi.org/10.1007/s00704-012-0718-7>
- Torres R, Marengo J (2014) Climate change hotspots over South America: from CMIP3 to CMIP5 multi-model datasets. *Theor Appl Climatol* 117:579–587. <https://doi.org/10.1007/s00704-013-1030-x>
- Valverde M, Marengo J (2014) Extreme rainfall indices in the hydrographic basins of Brazil. *Open J Mod Hydrol* 4:10–26. <https://doi.org/10.4236/ojmh.2014.41002>
- Vincent L, Peterson T, Barros V et al (2005) Observed trends in indices of daily temperature extremes in South America 1960–2000. *J Climate* 18:5011–5023. <https://doi.org/10.1175/JCLI3589.1>
- Wilby R, Dawson C (2013) The statistical downscaling model: insights from one decade of application. *Int J Climatol* 33:1707–1719. <https://doi.org/10.1002/joc.3544>
- Willmott C, Robeson S, Matsuura K (2012) A refined index of model performance. *Int J Climatol* 32:2088–2094. <https://doi.org/10.1002/joc.2419>
- Willmott C, Robeson S, Matsuura K, Ficklin D (2015) Assessment of three dimensionless measures of model performance. *Environ Model Softw* 73:167–174. <https://doi.org/10.1016/j.envsoft.2015.08.012>
- Xavier A, King W, Scanlon B (2015) Daily gridded meteorological variables in Brazil (1980–2013). *Int J Climatol* 36:2644–2659. <https://doi.org/10.1002/joc.4518>
- Xavier A, King C, Bridget S (2017) An update of Xavier, King and Scanlon (2016) daily precipitation gridded data set for the Brazil. Conference proceedings, pp 562–569. <https://proceedings.science/sbsr/papers/an-update-of-xavier-king-and-scanlon--2016--daily-precipitation-gridded-data-set-for-the-brazil>
- Xu K, Wu C, Hu B (2019) Projected changes of temperature extremes over nine major basins in China based on the CMIP5 multimodel ensembles. *Stoch Environ Res Risk Assess* 33:321–339. <https://doi.org/10.1007/s00477-018-1569-2>
- Yin L, Fu R, Shevliakova E, Dickinson R (2013) How well can CMIP5 simulate precipitation and its controlling processes over tropical South America? *Climate Dyn* 41:3127–3143. <https://doi.org/10.1007/s00382-012-1582-y>
- You Q, Jiang Z, Wang D et al (2017) Simulation of temperature extremes in the Tibetan Plateau from CMIP5 models and

- comparison with gridded observations. *Climate Dyn* 51:355–369. <https://doi.org/10.1007/s00382-017-3928-y>
- Yue S, Pilon P, Cavadias G (2002) Power of the Mann–Kendall and Spearman’s rho tests for detecting monotonic trends in hydrological series. *J Hydrol* 259:254–271. [https://doi.org/10.1016/S0022-1694\(01\)00594-7](https://doi.org/10.1016/S0022-1694(01)00594-7)
- Zhang X, Yang F, Canada E (2004) RCLimDex (1.0) User Guide. Climate Research Branch Environment Canada, Downsview, Ontario, Canada, pp 1–22
- Zhang X, Alexander L, Hegerl G et al (2011) Indices for monitoring changes in extremes based on daily temperature and precipitation data. *Wiley Interdiscip Rev Climate Change* 2:851–870. <https://doi.org/10.1002/wcc.147>
- Zhang Y, You Q, Chen C et al (2018) Evaluation of downscaled CMIP5 coupled with VIC model for flash drought simulation in a humid Subtropical Basin, China. *J Climate* 31:1075–1090. <https://doi.org/10.1175/JCLI-D-17-0378.1>

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