



# Association between El Niño and extreme temperatures in southern South America in CMIP5 models. Part 1: model evaluation in the present climate

Soledad Collazo<sup>1,\*</sup>, Mariana Barrucand<sup>1,2</sup>, Matilde Rusticucci<sup>1,2</sup>

<sup>1</sup>Departamento de Ciencias de la Atmósfera y los Océanos, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires (DCAO-FCEN-UBA), Ciudad Universitaria, 1428 Buenos Aires, Argentina

<sup>2</sup>Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), 1425 Buenos Aires, Argentina

**ABSTRACT:** Climate variability might temporarily improve or mitigate the effects of increasing global warming. Understanding and estimating internal variability is just as important as understanding the role of anthropogenic forcing, as the combination of both drives climate events in the real world. The objective of this work is to analyze the relationship between the SST of the equatorial Pacific and 4 extreme temperature indices in southern South America considering gridded observational data (HadEX3), reanalyses (ERA-Interim, NCEP1, NCEP2), and global climate models participating in the Coupled Model Intercomparison Project Phase 5 (CMIP5) in the historical period 1979–2005. For this, correlations and quantile regression for the 90th percentile were estimated between the variables. Moreover, to assess the performance of the reanalysis and CMIP5 models, multiple metrics were calculated. The observations showed that warm conditions in the equatorial Pacific are mainly associated with a higher occurrence of warm nights in the north and center of Argentina and Chile in winter and spring. However, the different reanalyses considered in this study showed discrepancies in representing these relationships. Several CMIP5 models were generally able to simulate correlation patterns for the warm extremes of the minimum and maximum temperature in comparison to HadEX3. These types of studies are critical to understanding whether climate models simulate temperature extremes in association with physical processes, providing greater confidence in their future projections.

**KEY WORDS:** Climate extremes · Global climate models · Attribution · Correlations · Quantile regression · Argentina

*Resale or republication not permitted without written consent of the publisher*

## 1. INTRODUCTION

Climate extreme events have devastating impacts on human society and ecosystems. A better understanding of observed changes in climate extremes is crucial for a reliable future prediction (Kim et al. 2016). To improve our knowledge of extreme events, it is necessary to study whether the processes that favor or inhibit their occurrence in the current climate are the same in the future, when the consequences of global warming are going to be even more drastic.

Despite the hiatus in the mean surface temperature recorded in recent years, the frequency of warm extremes has continued to increase and that of cold extremes has continued to decrease globally (Alexander 2016). This difference between mean and extreme temperatures could indicate that the processes that drive the mean temperature are not necessarily the same as those that generate the extreme temperatures. Possible causes include land-sea contrast in response to radiative forcing as well as feedbacks (for example, from decreased soil moisture, snow, or ice), which further amplifies the changes in

\*Corresponding author: scollazo@at.fcen.uba.ar

extreme temperatures in some key regions (Seneviratne et al. 2016).

SST in the equatorial Pacific is one of the primary drivers of the seasonal climate globally (Ropelewski & Halpert 1987). Kenyon & Hegerl (2008) found that different phases of ENSO influence temperature extremes worldwide. Cooler (warmer) extreme maximum temperatures over Australia, southern Asia, Canada, and South Africa (North America, Eastern Europe–Central Asia, northeastern Asia, and southern South America) were observed using gridded observations and coupled climate model simulations during strong La Niña events compared to El Niño events (Arblaster & Alexander 2012, Luo & Lau 2020). Various China regions showed significant correlations between summer high-temperature extremes and 2 ENSO types (Gao et al. 2020). Furthermore, El Niño significantly amplifies the heatwave activities in most areas of China (Luo & Lau 2019) and India (Murari et al. 2016).

To understand the variability of extreme temperatures in southern South America under the present climate, it is necessary to examine their relationship with the associated atmospheric circulation. Rusticucci et al. (2017) studied the co-variability between temperature extremes in Argentina and regional circulation. They found that the subtropical jet intensity presented significant positive (negative) correlations with the frequency of cold (warm) extremes in many months. Moreover, Loikith et al. (2017) analyzed the associations between extreme temperature in South America and the 4 main modes of climate variability. In particular, ENSO is the mode that presented the strongest relationships with all extreme indices. The most recent advances are described in Collazo et al. (2019a) and Collazo et al. (2019b), in which predictors of hot days in summer and cold nights in winter, respectively, were sought. In the first of those studies, the principal component regression technique was successfully applied to predict the warm extremes of the maximum temperature.

Several previous works have found that SST anomalies in the equatorial Pacific associated with ENSO explain much of the interannual variability of the Southern Hemisphere (Vera et al. 2004 and references within). Through teleconnections, the El Niño phase is associated with an increase in precipitation over southeastern South America (SESA) (Ropelewski & Halpert 1987, Vargas et al. 1999, Grimm et al. 2000, among others). In terms of mean temperature, Barros et al. (2002) found positive temperature anomalies in the SESA region during winter for the El Niño phase. Opposite temperature patterns tend

to occur in La Niña years over South America. The relationship between the number of days with frost and ENSO in the Humid Pampa region was studied by Müller et al. (2000), who concluded that part of the variability in the frequency of frost occurrence is explained by the ENSO cycle, with a lower average of frost than the climatology during El Niño events.

Among the studies that analyze the ENSO impacts on temperature extremes in Argentina, Rusticucci & Vargas (2002) observed more homogeneous effects over temperature extremes during La Niña episodes than during the El Niño phase, which makes their prediction easier, especially for cold events. Kenyon & Hegerl (2008) analyzed the influence of large-scale climate variability in extreme summer and winter temperatures at a global level. In particular, they found a higher frequency of warm extremes of the minimum temperature in El Niño years during the extended winter in Argentina. Furthermore, Garreaud et al. (2009) detected a positive relationship between temperature anomalies and ENSO phases over much of South America, i.e. El Niño is associated with temperatures above normal, while La Niña is associated with cold anomalies. More recently, Agosta & Barrucand (2012) described the modulation of the positive ENSO phase in the frequency of warm nights over subtropical Argentina. Finally, Rusticucci et al. (2017) demonstrated that the impact of the El Niño event on extreme temperatures in Argentina presents monthly differences, favoring warming during winter (more nights and warm days) and colder conditions in summer (fewer warm days and more cold days). The opposite pattern occurred during La Niña; in summer, almost every month exhibited more (less) frequency of warm (cold) days compared to their climatology.

To study future climate projections towards the end of the century, we must resort to global climate models (GCMs) that are capable of simulating the entire Earth System and its interactions. First, GCMs must be evaluated on their ability to represent the current climate, which then provides some confidence in their future projections. Sillmann et al. (2013a) provide a first overview of the performance of the CGMs of the penultimate generation, participating in Phase 5 of the Coupled Model Intercomparison Project (CMIP5), in the simulation of extreme climate indices. The authors showed that CMIP5 models are generally capable of simulating climatic extremes and their trend patterns.

To adequately represent temperature extremes, models must simulate synoptic-scale phenomena, low-frequency modes that provide a large-scale

meteorological context, small-scale atmospheric processes, and land surface processes that influence surface heat fluxes (Grotjahn et al. 2016). However, atmospheric models have experienced historical difficulties in representing some types of low-frequency intraseasonal variability (Black & Evans 1998). Tedeschi & Collins (2016) investigate how CMIP5 models display different types of ENSOs and how they represent teleconnections with South American rainfall. They observed that CMIP5 models were capable of simulating the basic structure of SST anomalies that occur during different types of ENSO events and reproducing the correct sign of precipitation anomalies in northeastern South America during the southern summer and autumn. However, in eastern and southeastern South America, many models do not reproduce the correct signal during ENSO events. Finally, they concluded that errors in teleconnections are primarily due to the atmospheric component of the models and not to the bias in the SST.

Internal variability in the oceans influences the global climate in diverse ways; nevertheless, to date, few studies have systematically investigated the role of SST variability on extreme temperatures and its representation in GCMs (Dittus et al. 2018). SST variability could explain approximately 50% of the interannual variability of global averaged extreme temperatures (Donat et al. 2016); thus, it is relevant that the physical component of the models might be able to simulate it correctly.

Large-scale atmospheric circulation patterns are relevant drivers for local and regional extremes, especially on the interannual time scale (Horton et al. 2015, Grotjahn et al. 2016, King et al. 2016, Rusticucci et al. 2017, Collazo et al. 2019a,b). Although there is some confidence about future changes in these patterns (Collins et al. 2010, Guilyardi et al. 2012, Kim & Yu 2012, Stevenson 2012, Christensen et al. 2013, Power et al. 2013, Cai et al.

2014, Perry et al. 2017), little is known about the projected responses of the extremes caused by these changes in atmospheric circulation. The objective of this initial part of the work is to assess the representation made by different reanalyses and GCMs of the relationships between the extreme temperature indices in southern South America and SST in the El Niño 3.4 region under the present climate (1979–2005). The association was evaluated by correlations and quantile regression. Ultimately, the strength of agreement between models and observations was measured by different metrics such as Taylor diagrams, Cohen's  $\kappa$  coefficients, and Mapcurves score.

## 2. DATA AND METHODOLOGY

### 2.1. Data

To represent extreme temperature events, we considered 4 climate indices used internationally (Table 1), defined by the Expert Team on Climate Change Detection and Indices, for the period 1979–2005 in southern South America ( $20^{\circ}$ – $60^{\circ}$ S,  $285^{\circ}$ – $310^{\circ}$ E). These indices are based on a percentile threshold, i.e. they describe the exceedance rates above or below a threshold defined as the 10th or 90th percentile derived from the 1961–1990 base period. The validation of climate extremes indices and the analysis of their projected future changes simulated by the CMIP5 models are presented in Sillmann et al. (2013a,b). This data is available at <http://climate-modelling.canada.ca/climatemodeleddata/climdex/climdex.shtml> (accessed 12 August 2020).

On the other hand, the observed and CMIP5 modeled SST data in the El Niño 3.4 region (SST3.4,  $5^{\circ}$ N– $5^{\circ}$ S,  $190^{\circ}$ – $240^{\circ}$ E) was obtained from KNMI Climate Explorer (<https://climexp.knmi.nl/start.cgi>, accessed 12 Aug 2020).

Table 1. Four extreme temperature indices recommended by the Expert Team on Climate Change Detection and Indices

Label	Index name	Description
TN10p	Cold nights	Percentage of days in a month when daily minimum temperature is below the 10th percentile centered on a 5 d window
TN90p	Warm nights	Percentage of days in a month when daily minimum temperature is above the 90th percentile centered on a 5 d window
TX10p	Cold days	Percentage of days in a month when daily maximum temperature is below the 10th percentile centered on a 5 d window
TX90p	Warm days	Percentage of days in a month when daily maximum temperature is above the 90th percentile centered on a 5 d window

### 2.1.1. Observations and reanalysis

We used the gridded HadEX3 data set of observation-based extreme temperature indices (Dunn et al. 2020). HadEX3 indices are calculated directly from station-based observations and then interpolated to a global grid only over land on a  $1.25^\circ \times 1.875^\circ$  grid from 1901 to 2018. This data set allows a comparison between model-simulated and observed indices.

Reanalyses data sets are also often used for model evaluation. In this study, we employed extreme temperature indices in southern South America for 3 widely used reanalyses: ERA-Interim (Dee et al. 2011), NCEP/NCAR Reanalysis 1 (NCEP1) (Kistler et al. 2001, Kalnay et al. 1996), and NCEP-DOE Reanalysis 2 (NCEP2) (Kanamitsu et al. 2002). The extreme temperature indices estimated from ERA-Interim are available on a regular  $1.5^\circ \times 1.5^\circ$  grid for the 1979–2019 period. Both NCEP1 and NCEP2 reanalysis extreme temperature indices are available for  $192 \times 94$  Gaussian grid points for the years 1948–2020 and 1979–2020, respectively. We considered both NCEP reanalyses and not just the latest version because the NCEP1 reanalysis is still widely used in the meteorological community for its prolonged temporal coverage. Moreover, we were also interested in comparing both performances and determining if NCEP2 achieves improvements concerning the first version.

The observed monthly average SST data set used in this work was HadISST version 1.1 (Rayner et al. 2003). HadISST is a combination of monthly globally complete fields of SST and sea ice concentration for 1871–2019 present on a regular  $1^\circ \times 1^\circ$  grid. This data was spatially averaged in the region  $5^\circ\text{N}$ – $5^\circ\text{S}$ ,  $190^\circ$ – $240^\circ\text{E}$ .

### 2.1.2. GCMs

The results of the climate model run depend on the starting point of the calculation, the initialization method, and the model physics. Ensemble calculations facilitate quantifying the variability of simulation data concerning a single model. The CMIP is a standard experimental framework for studying the output of coupled atmosphere–ocean general circulation models. In the CMIP5 project, the ensembles include multiple runs from some GCMs, and the simulations are distinguished by their rip nomenclature, r for realization, i for initialization method, and p for physics version followed by an integer, e.g. r1i1p1. We used extreme temperature indices from 23

CMIP5 GCMs for the experiment r1i1p1 and its ensemble (Table 2). The historical simulation period of the CMIP5 models is from about 1850 to 2005. These simulations employ historical changes in the atmospheric composition reflecting both anthropogenic and natural sources (Sillmann et al. 2013a).

In addition, we used spatially averaged SST in the El Niño 3.4 region for each GCM and ensemble.

## 2.2. Methodology

### 2.2.1. Processing

Before any calculation, we regredded the extreme temperature indices from all the different data sets to a common grid ( $2.5^\circ \times 2.5^\circ$ ) using a first-order conservative remapping procedure (Jones 1999) implemented in Climate Data Operators version 1.9.6rc3 (<https://code.zmau.de/projects/cdo>), following Sillmann et al. (2013a). The resulting grid has  $10 \times 16$  points in the southern South America region. This processing allowed us to compare the different models and contrast them with the observations. For the comparison between HadEX3 indices and GCM indices, only land grid points were considered.

Then, the extreme temperature indices were averaged quarterly to represent the 4 seasons of the year: summer (DJF), autumn (MAM), winter (JJA), and spring (SON). The extremes were seasonally analyzed, as they are often more relevant to impacts (Menzel et al. 2006, Orlowsky & Seneviratne 2012, Alexander 2016).

### 2.2.2. Association between extreme temperature indices and El Niño 3.4 SST

Initially, the linear trends of the extreme temperature indices and the spatial average of the SST in the central equatorial Pacific were filtered, since we seek to analyze the component associated with interannual variability instead of long-term forcing. To achieve this, we subtracted the linear regression adjusted by least squares.

Subsequently, the Spearman correlation coefficient (Spearman 1904) and quantile regression for the 90th percentile (Koenker 2005) were calculated between SST3.4 and extreme temperature indices for the 4 seasons. Specifically, HadEX3 and reanalyses indices were correlated with El Niño 3.4 SST derived from the HadISST1.1 data set and GCM indices with SST of the model itself. The statistical

Table 2. List of CMIP5 models used in this work

Model	Institution, country	Atmospheric resolution ( $^{\circ}$ long $\times$ $^{\circ}$ lat)	Reference
1 ACCESS 1.0	CSIRO-BOM, Australia	$1.875 \times 1.25$	Bi et al. (2013)
2 bcc-csm1-1-m	BCC-CMA, China	$2.81 \times 2.79$	Wu et al. (2014)
3 bcc-csm1-1	BCC-CMA, China	$2.81 \times 2.79$	Xin et al. (2013)
4 CanESM2	CCCMA, Canada	$2.81 \times 2.79$	Arora et al. (2011)
5 CCSM4	NCAR, USA	$1.25 \times 0.94$	Gent et al. (2011)
6 CMCC-CM	CMCC, Italy	$0.75 \times 0.75$	Scoccimarro et al. (2011)
7 CMCC-CMS	CMCC, Italy	$3.75 \times 3.71$	CMCC (2013)
8 CNRM-CM5	CNRM-CERFACS, France	$1.41 \times 1.40$	Voldoire et al. (2013)
9 CSIRO-Mk3-6-0	CSIRO-QCCCE, Australia	$1.88 \times 1.87$	Rotstayn et al. (2010)
10 GFDL-CM3	NOAA GFDL, USA	$2.50 \times 2.00$	Donner et al. (2011)
11 GFDL-ESM2G	NOAA GFDL, USA	$2.00 \times 2.02$	Dunne et al. (2012)
12 GFDL-ESM2M	NOAA GFDL, USA	$2.50 \times 2.02$	Dunne et al. (2012)
13 HadGEM2-CC	MOHC, UK	$1.88 \times 1.25$	Collins et al. (2011)
14 HadGEM2-ES	MOHC, UK	$1.88 \times 1.25$	Collins et al. (2011)
15 inmcm4	INM, Russia	$2.00 \times 1.50$	Volodin et al. (2010)
16 IPSL-CM5A-LR	IPSL, France	$3.75 \times 1.89$	Dufresne et al. (2013)
17 IPSL-CM5A-MR	IPSL, France	$2.50 \times 1.25$	Dufresne et al. (2013)
18 IPSL-CM5B-LR	IPSL, France	$3.75 \times 1.89$	Dufresne et al. (2013)
19 MIROC5	MIROC, Japan	$1.41 \times 1.40$	Watanabe et al. (2010)
20 MPI-ESM-LR	MPI-M, Germany	$1.88 \times 1.87$	Zanchettin et al. (2013)
21 MPI-ESM-MR	MPI-M, Germany	$1.88 \times 1.87$	Zanchettin et al. (2013)
22 MRI-CGCM3	MRI, Japan	$1.13 \times 1.12$	Yukimoto et al. (2012)
23 NorESM1-M	NCC, Norway	$2.5 \times 1.89$	Bentsen et al. (2013)

significance of the correlations and regressions was evaluated at a 95 % confidence level.

The Spearman correlation coefficient is a measure of non-parametric association between 2 continuous random variables that do not follow a normal distribution. The statistical significance of the correlation coefficient was tested using the AS 89 algorithm (Best & Roberts 1975). The null hypothesis states that no monotonic association exists between the 2 variables in the population. The alternative hypothesis affirms that a monotonic correlation is present. We used this Spearman correlation coefficient instead of the more traditional Pearson correlation coefficient because the extreme temperature indices have a non-Gaussian distribution (Klein Tank et al. 2009).

To assess the impact of one variable on the tails of the distribution of another, the slope of the 90th percentile quantile regression between the extreme temperature indices and SST3.4 was estimated at each grid point (Mueller & Seneviratne 2012). An advantage of the quantile regression methodology over the least squares regression is that quantile regression does not assume a particular parametric distribution for the response variable nor a constant variance (Rodriguez & Yao 2017). Moreover, quantile regression allows modeling multiple percentiles of the distribution, which provides a deep understanding of the relationship between the dependent variable and the

predictors. To test the statistical significance, we estimated the SE and the p-value using a bootstrap method since it does not require any *a priori* assumption of the data distribution (Hao & Naiman 2007).

### 2.2.3. Spatial performance metrics

Comparison of the observed and modeled correlation and slope fields makes it possible to identify models with the skill to represent the relationship between SST3.4 and temperature extremes. This assessment was performed quantitatively by Taylor diagrams and categorically by Cohen's kappa coefficient and Mapcurves score.

Taylor diagrams (Taylor 2001) provide a visual framework and a concise statistical summary of how well patterns match each other in terms of their correlation, their root mean square error, and the ratio of their variances. The Taylor diagram allows comparing a suite of variables from 1 or more test data sets to 1 or more reference data sets. Commonly, the test data sets are model experiments, while the reference data set is a control experiment or some reference observations. All variables must be on the same grid, so regridding may be necessary. The Taylor diagram is very useful for evaluating climate models and has been used extensively in model assessment and

intercomparison studies in recent years (e.g. Hellström & Chen 2003, Töyrä et al. 2005, Giorgi & Gutowski 2015, Jiang et al. 2015, Katragkou et al. 2015, Tang et al. 2016, Tencer et al. 2016, Lovino et al. 2018, Molina & Bernhofer 2019).

In the categorical approach, we evaluated whether the statistically significant regions coincide, as well as their sign. These fields can be analyzed as categorical maps (with positive and significant, negative and significant, and non-significant regions) with 2 types of errors (Pontius 2000): (1) the quantization error that occurs when the number of cells of a particular category in one map is different from the number of cells in that category on the other map, and (2) the location error that occurs when the location of a category on one map is different from the location of that category on the other map.

The most direct comparison involves determining whether each of the grid points is in the same category as the one observed. This analysis is useful for the construction of contingency tables (Congalton & Green 1993). From the contingency tables, Cohen's kappa coefficient was calculated. This coefficient adjusts the effect of chance in the proportion of correct answers for categorical variables (Cohen 1960).

Cohen's kappa ( $\kappa$ ) statistic is a chance-corrected method for assessing agreement (rather than association) among raters (DeVellis 2005).  $\kappa$  is defined as follows (Eq. 1):

$$\kappa = \frac{p_0 - p_e}{1 - p_e} \quad (1)$$

where  $p_0$  is the observed proportion of agreements between raters, and  $p_e$  is the probability of agreements expected by chance. In this case, one rater is the observations, and the other rater is each model. The probability  $p_0$  is the hit ratio (number of hits divided by the total number of observations).

For  $k$  categories,  $N$  observations to categorize and  $n_{ki}$  the number of times rater  $i$  predicted category  $k$  (Eq. 2):

$$p_e = \frac{1}{N^2} \sum_k n_{k1} n_{k2} \quad (2)$$

Cohen's  $\kappa$  is always less than or equal to 1. Values of zero or less indicate the classifier is useless, i.e. the observations and models agreed less than would be expected just by chance. A value of 1 implies perfect agreement, and values less than 1 imply less than perfect agreement. Landis & Koch (1977) provide a way to characterize values. According to their scheme, a value  $<0$  indicates no agreement, 0–0.20 slight agreement, 0.21–0.40 fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 substantial agreement, and 0.81–1 almost perfect agreement.

The  $\kappa$  coefficient was also tested by considering a 2-tailed Z-test with a 5 % significance level (Eq. 3). The null hypothesis states that the empirical  $\hat{\kappa}$  coefficient is equal to a population  $\kappa$ . In this work, we considered  $\kappa_0 = 0$ , because a negative value of  $\kappa$  does not normally have a meaningful interpretation (Sim & Wright 2005):

$$Z = \frac{\hat{\kappa} - \kappa_0}{SE(\hat{\kappa})} \quad (3)$$

where  $SE(\kappa) = \sqrt{\frac{p_0(1-p_0)}{N(1-p_e)^2}}$

The  $\kappa$  coefficient was previously used to evaluate the performance of regional and global climate models in simulating the climate classifications (Sparovek et al. 2007, Tapiador et al. 2019) and to measure the agreement between predicted and observed classes of wet days (Poggio & Gimona 2015).

Contingency table methods fail to distinguish between a near miss and a far miss and are not designed to account for partial success (Hargrove et al. 2006). To complement the information provided by the  $\kappa$  coefficient, a Mapcurves score was estimated. Mapcurves make it possible to unequivocally determine the degree of spatial agreement between 2 or more categorical maps (Hargrove et al. 2006), i.e. Mapcurves are a measure of goodness of fit (GOF) (Demirel et al. 2018). The Mapcurves score has an optimal value of 1, whereas the lowest value is zero since it represents the proportion of GOF. In this study, the function 'Mapcurves (x, y)' available in the 'sabre' package version 0.3.2 (Nowosad & Stepinski 2018), written in the R programming language, was used for estimating Mapcurves values. R version 3.6.0 was employed for the calculations. In that equation,  $x$  and  $y$  are vectors representing the categorical values of historical observed data and categorical values of simulated data by a GCM, respectively (Ahmed et al. 2019). The Mapcurves value ( $MC_X$ ) is calculated by Eq. (4):

$$MC_X = \sum_{Y=1}^n \left[ \left( \frac{C}{A} \cdot \frac{C}{B} \right) \right] \quad (4)$$

where  $A$  is the total area of a given class  $X$  on the map being compared,  $B$  is the total area of a given class  $Y$  on the observed map,  $C$  is the area of intersection between  $X$  and  $Y$  when the maps are overlaid, and  $n$  is the number of classes in the observed map.

There are different sources of error in the GCMs that can derive good and bad performance in simulating the correlations between SST3.4 and extreme

indices. One of the most prominent errors is a poor representation of the teleconnections by the models. To analyze this source of uncertainty, we must consider the physical processes which may be responsible for the occurrence of extreme temperatures.

Minimum and maximum temperatures are sensitive to cloudiness. Radiative cooling is reduced under cloudy nights, since some of the heat emitted from the earth's surface is trapped by the clouds and reemitted back towards the surface. During a cloudy day, less of the sun's energy is able to reach the earth's surface, which causes the surface to heat up more slowly. In this sense, El Niño events favor the higher occurrence of rainfall in SESA. Silvestri (2005) shows that winter precipitation during the El Niño phase is significantly higher than La Niña and in neutral cases in SESA. More recently, Tedeschi & Collins (2016) showed that canonical El Niño produces positive precipitation anomalies and precipitation events more extreme than the climatology over SESA in the JJA season. Likewise, several authors found an increase in precipitation in SESA during El Niño years in spring (Vargas et al. 1999, Grimm et al. 2000, Grimm & Tedeschi 2009). This increase in precipitation and cloud cover could be the reason for warmer conditions at night and cooler conditions during diurnal hours. Therefore, we decided to evaluate how GCMs represent the anomalies of sea level pressure in the Southern Hemisphere and precipitation in southern South America by comparing them against ERA-Interim fields. For this purpose, the detrended SST3.4 index was standardized, and a threshold of  $\pm 0.5$  was considered to select El Niño and La Niña events.

### 3. RESULTS AND DISCUSSION

#### 3.1. Observed association between SST3.4 and extreme temperature indices

The warm night index shows the largest region with significant correlations with SST3.4. During autumn, significant positive correlations are predominantly located in Chile and western Argentina (Fig. 1). In winter and spring, the direct correlations imply that warmer SSTs are associated with a higher frequency of warm nights in northern and central Argentina and Chile, Uruguay, Paraguay, and southern Brazil. In addition, during the spring, negative correlations are observed in southern Argentina. Warm days show a similar pattern in SON, but the negative correlations are more extended to the north, and the positive correlations are confined to the

northwest region. In contrast, cold nights presented negative correlations mainly in northwestern Argentina, and cold days are positively correlated with SST3.4 in Patagonia in SON. The opposite response to SST was observed between the cold and warm extremes in MAM and SON. A similar pattern was identified by Rusticucci et al. (2017) in central and northern Argentina during September and by Kenyon & Hegerl (2008) in winter. The ENSO impacts over southern South America by an extratropical teleconnection through a pattern of stationary Rossby wave trains which is most pronounced during SON. During an El Niño event, the atmospheric pattern culminates in negative and positive pressure anomalies over the eastern mid-latitudes and eastern subtropical South America, respectively (Cai et al. 2020). These atmospheric circulation anomalies favor northwesterly advection of moist warm air into SESA and an increase in precipitation (Silva & Ambrizzi 2006, Viale et al. 2018, Montini et al. 2019). In SESA, we only found a significant association between SST3.4 and warm nights, i.e. warm advection and an enhanced cloud cover during El Niño favor warmer minimum temperature.

The influence of SST on the highest values of the extreme indices is even more localized than what we observe for the correlations, even though the sign of the association remains (Fig. 2). The observed slopes tend to be smaller than  $10\% \text{ }^{\circ}\text{C}^{-1}$  in absolute value, i.e. for a positive trend, an increase of  $1^{\circ}$  in SST3.4 is associated with an increase less than 10% in the 90th percentile of the extreme temperature indices.

#### 3.2. Reanalyses performance assessment

To sum up the analysis, we focus the evaluation on the warm nights index in the winter (JJA) and spring (SON) seasons because they presented the most extended area with significant associations. Fig. 3 shows that ERA-Interim presents difficulties representing significant positive correlations in central Argentina for both seasons. NCEP1 and HadEX3 have a similar pattern of significant correlations in JJA, while NCEP1 performs poorly in western and southern Argentina in SON. Finally, the performance of NCEP2 is worse than NCEP1 in JJA and similar to its predecessor reanalysis in SON.

The slope of the quantile regressions between observed SST3.4 and the 90th percentile of observations and reanalyses can be found in Fig. 4. This figure shows a good agreement in JJA for NCEP2, but the reanalyses were unable to represent the signifi-

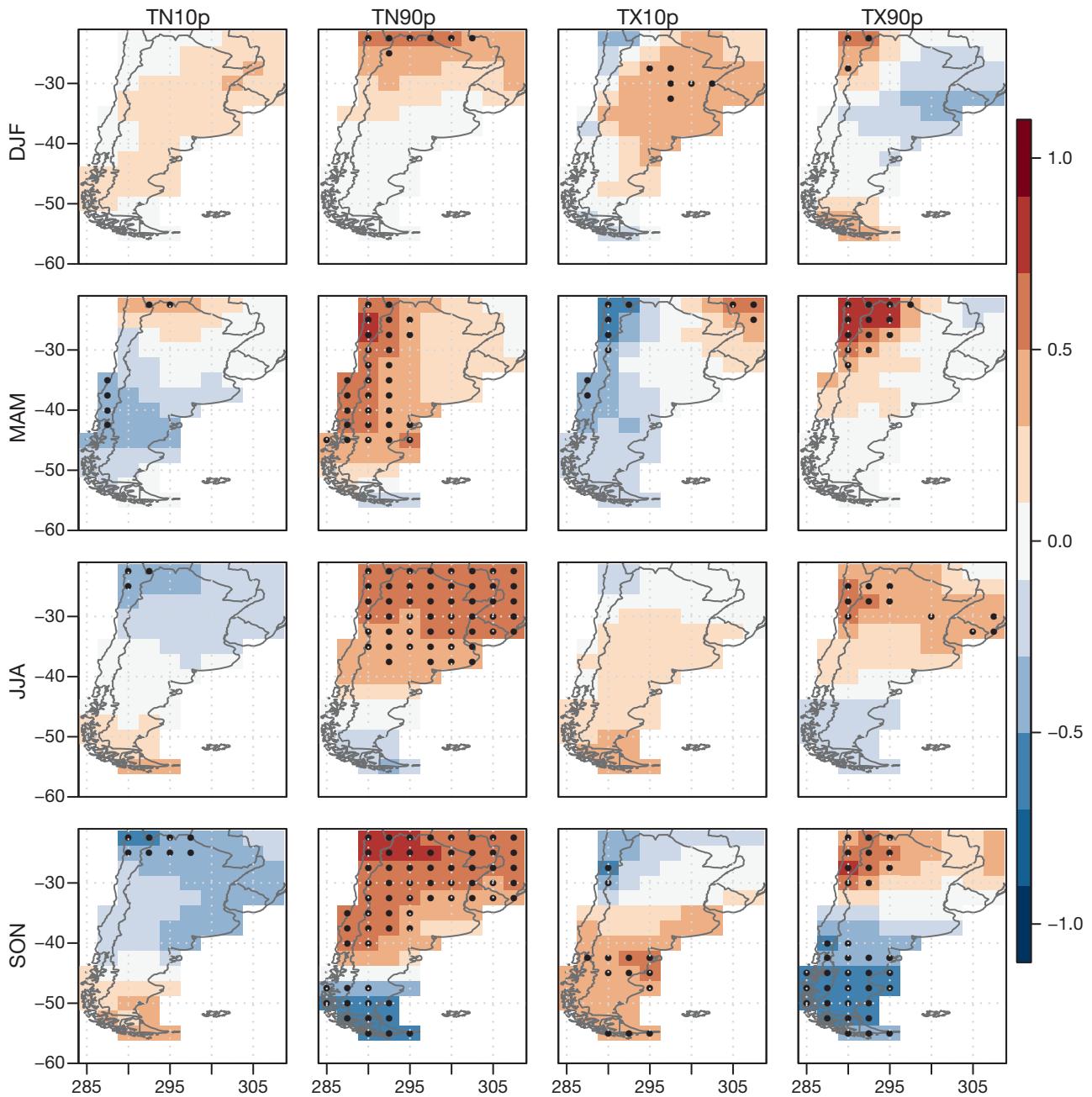


Fig. 1. Correlation between observed SST in the El Niño 3.4 region and extreme temperature indices of HadEX3 (dotted regions indicate significant correlation at the 0.05 significance level). In the x-axis, longitude is shown on a scale of 0°–360°E, the y-axis displays latitude in °N

cant slopes observed in northern and western Argentina in SON.

The performance of the reanalyses was objectively quantified through 2 metrics (Cohen's  $\kappa$  coefficient and Mapcurves values). Fig. 5 shows that the performance of the reanalysis varies considerably according to the season and the metric applied. Both metrics agree the NCEP1 is the best in representing

the observed correlations in JJA. Moreover, both NCEP reanalyses show a moderate performance for the slopes of the quantile regression in winter. During the spring, the metrics do not agree on what the best reanalysis is, even though all the reanalyses show a significant  $\kappa$  coefficient. The slight strength of agreement and the Mapcurves values below 0.5 are indicative of the poor performance of the reanalyses

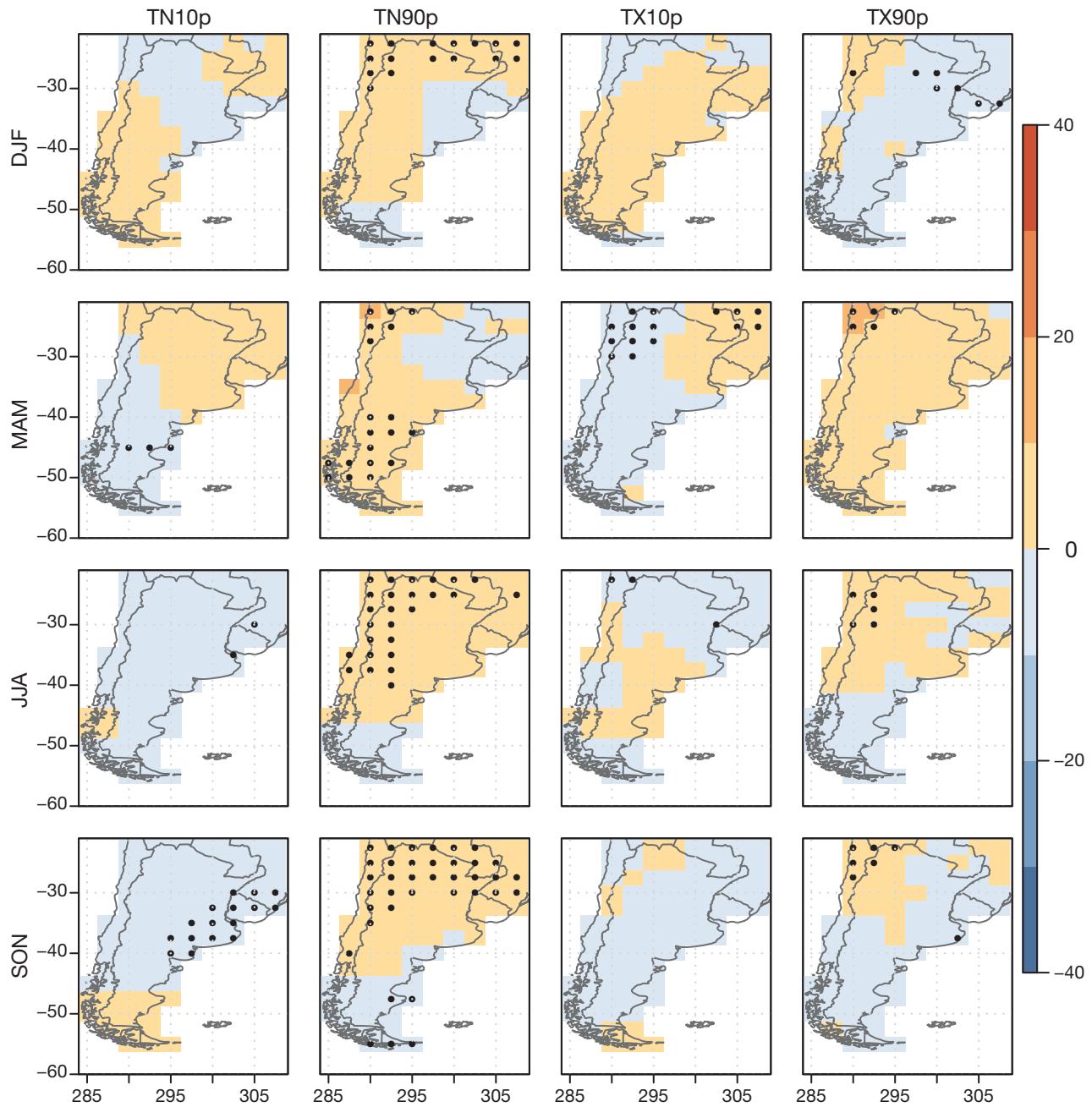


Fig. 2. Slope of the quantile regression for the 90th percentile between extreme temperature indices of HADEX3 and SST in the El Niño 3.4 region ( $\% \text{ } ^\circ\text{C}^{-1}$ ). Dots indicate significant slopes at the 95 % confidence level

in estimating the slopes of the 90th percentile in SON. In general, for all extreme indices and seasons, the reanalyses are more efficient in representing correlations than slopes of the 90th percentile, which might be caused by an incorrect depiction of the right tail of the distribution of the extreme indices.

*A priori*, we expected to obtain more similar results between the different reanalyses and observations; however, the discrepancies were substantial, indi-

cating considerable uncertainties regarding their simulation of extremes. Sillmann et al. (2013a) arrived at a similar conclusion and affirmed that reanalyses might be affected by, for instance, inaccurate representation of surface and boundary layer processes, convection, and its spinup. Moreover, we agree with Angélil et al. (2016) about the importance of using multiple reanalyses and observation products for attribution studies.

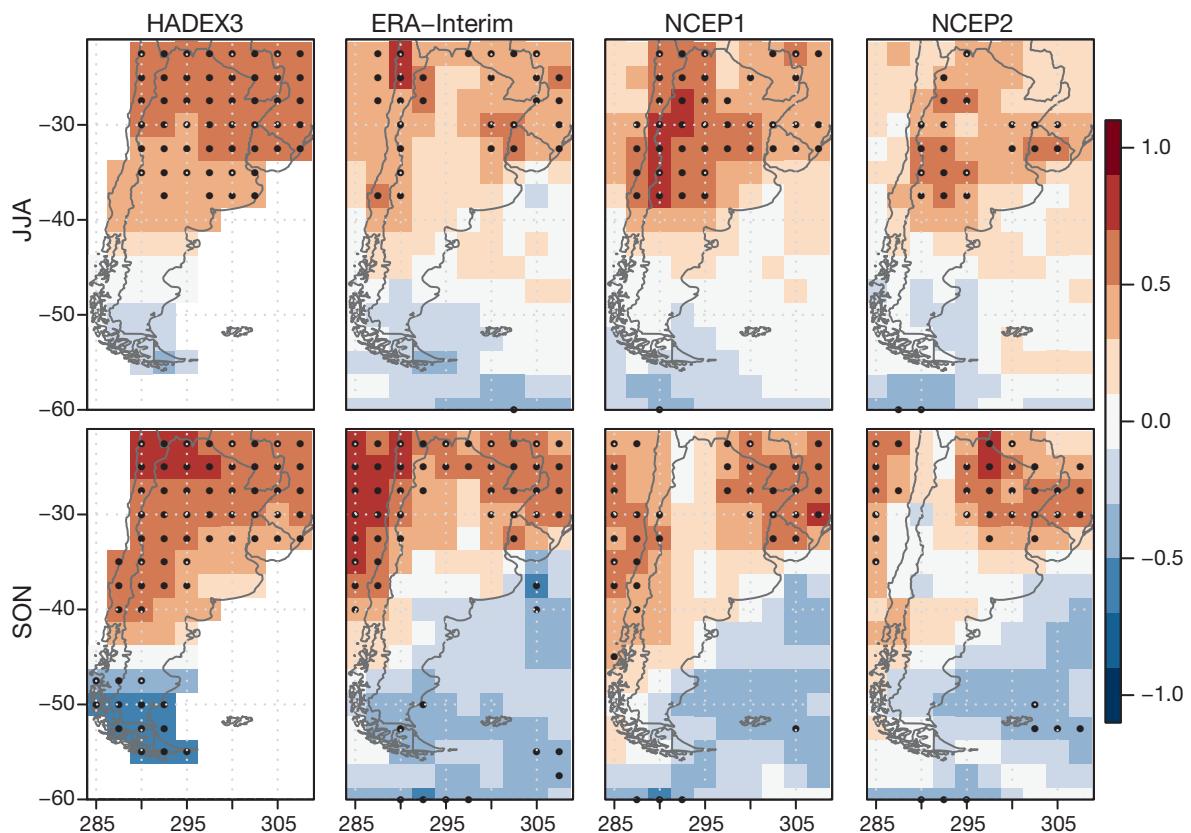


Fig. 3. Correlation between observed SST in the El Niño 3.4 region and warm nights index of HadEX3 and reanalyses (dotted regions indicate significant correlation at the 0.05 significance level)

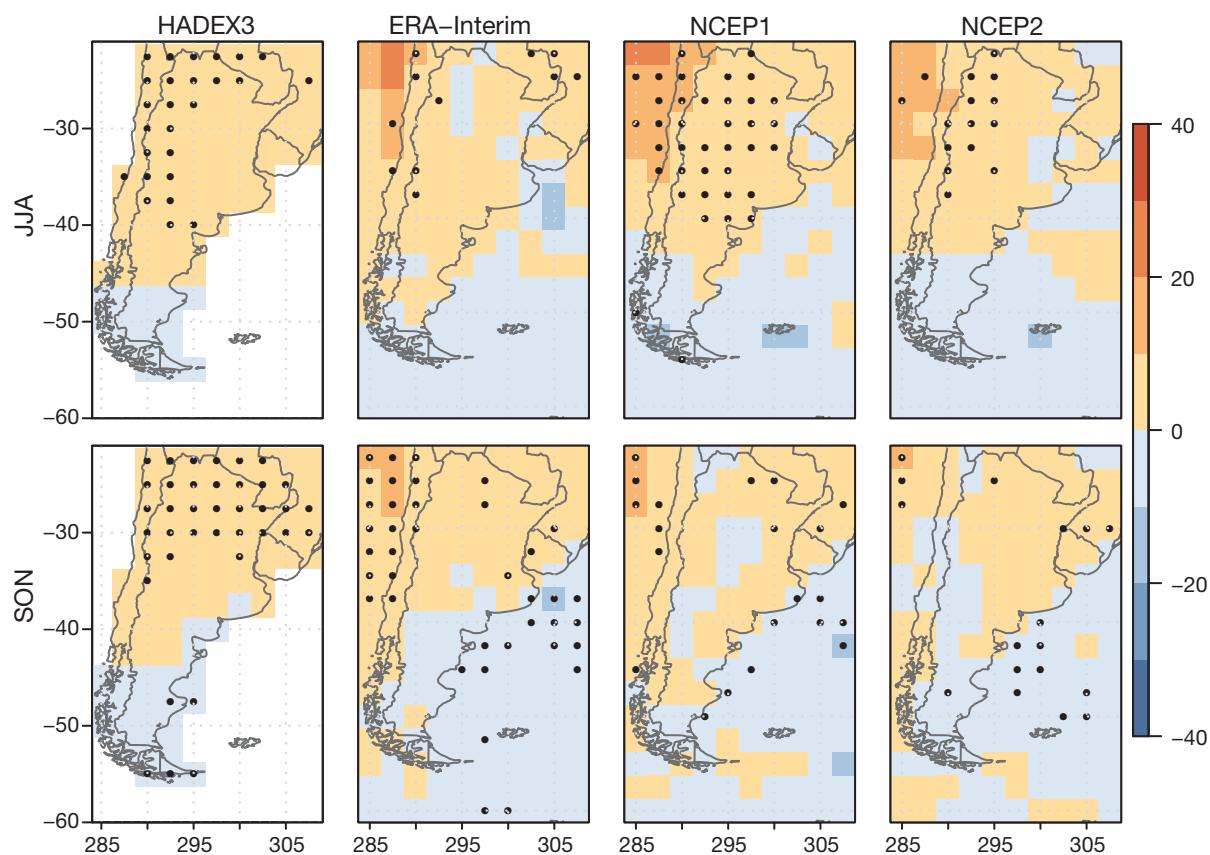


Fig. 4. Slope of the quantile regression for the 90th percentile between observed SST in the El Niño 3.4 region and warm nights index of HadEX3 and reanalyses (% °C<sup>-1</sup>) (dotted regions indicate significant correlation at the 0.05 significance level)

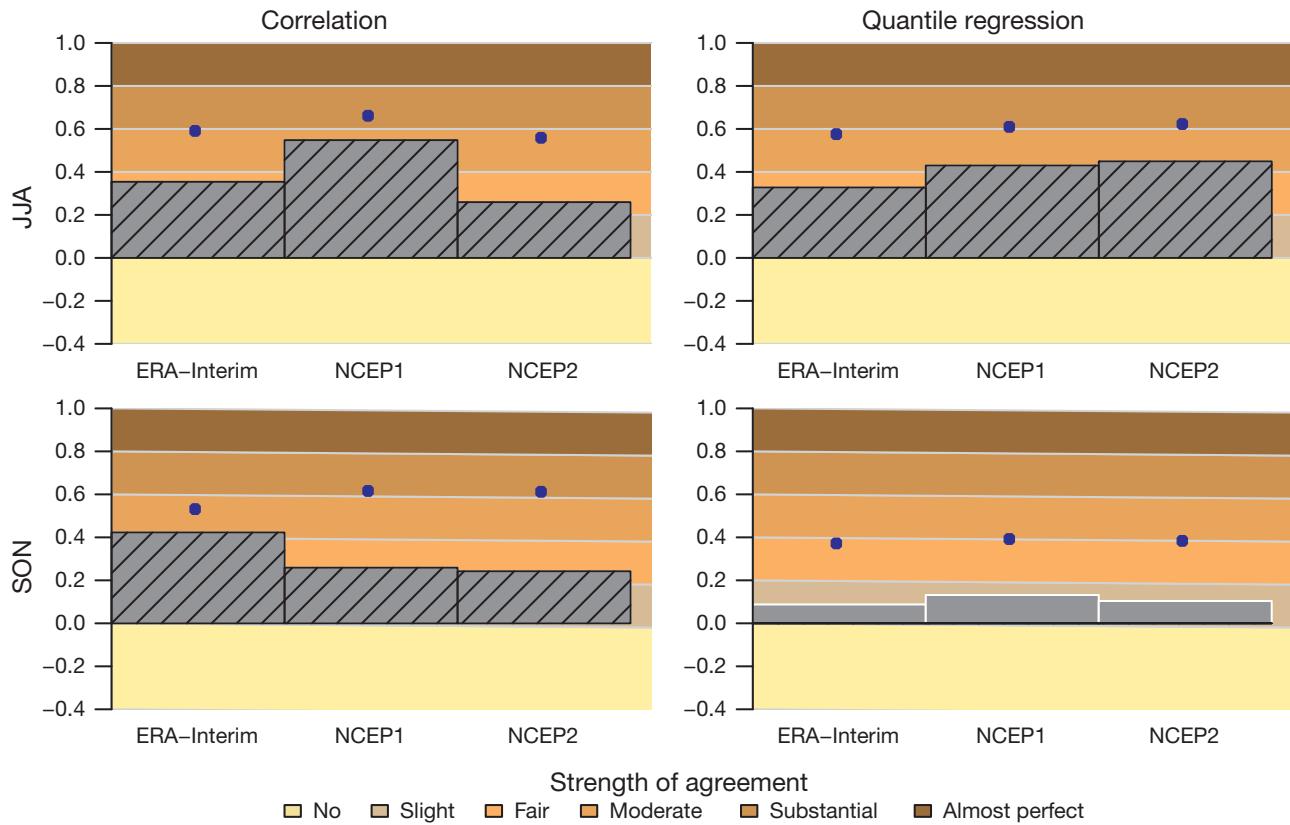


Fig. 5. Cohen's  $\kappa$  (bars) and Mapcurves (points) scores comparing the HadEX3 data set against reanalyses for correlations (left) and the slope of the quantile regression (right) between the El Niño 3.4 SST and TN90p. The strength of agreement of Cohen's  $\kappa$  coefficient is shown in colors, and diagonal lines indicate significant  $\kappa$  at 5%

### 3.3. GCM performance

Figs. 6–9 show simulations by the GCMs of the correlations and slopes of the 90th percentile between SST3.4 and TN90p in the JJA and SON seasons. Highly varied results were found between the models. The GCMs that could simulate with some precision the significant positive correlations observed in JJA in central and northern Argentina and Chile, Uruguay, Paraguay, and southern Brazil are CanESM2, CSIRO-Mk3-6-0, GFDL-ESM2G, GFDL-ESM2M, IPSL-CM5A-LR, IPSL-CM5B-LR, MPI-ESM-LR, MPI-ESM-MR, and NorESM1-M (Fig. 6). These models were capable of detecting that warmer SST conditions in the equatorial Pacific are associated with a higher frequency of warm nights in the mentioned region. The only model that represented opposite conditions to those observed is CMCC-CMS, with significant negative correlations in eastern Argentina. On the other hand, the ensemble mean only shows significant positive correlations in the north of the study region.

In SON, only the CNRM-CM5 model was able to simulate the observed dipole of correlation with neg-

ative correlations in southern Argentina and positive correlations in central and northern Argentina, even though this model did not simulate the relationship between SST3.4 and TN90p in central Argentina and Chile (Fig. 7). Associated with the poor representation of the models, the ensemble mean of the models cannot achieve an adequate representation of the correlations, even though it can successfully simulate the negative correlations in southern Argentina. Finally, among the worst models are bcc-csm1-1-m, IPSL-CM5A-LR, and MIROC5, since they represent an opposite pattern to what was observed.

The comparison for the quantile regression shows that the ensemble mean and several models simulated significant positive slopes in northern Chile and western Argentina in JJA as the observations (Fig. 8). In the spring, the results are more variable between models, being difficult to visually identify which was the best (Fig. 9).

The objective comparison between observations and simulations through Taylor diagrams is shown in Fig. 10 for TN90p in JJA and SON. In addition, in the Supplement (available at [www.int-res.com/articles/suppl/c083p111\\_supp.pdf](http://www.int-res.com/articles/suppl/c083p111_supp.pdf)), the Taylor diagrams and

Author copy

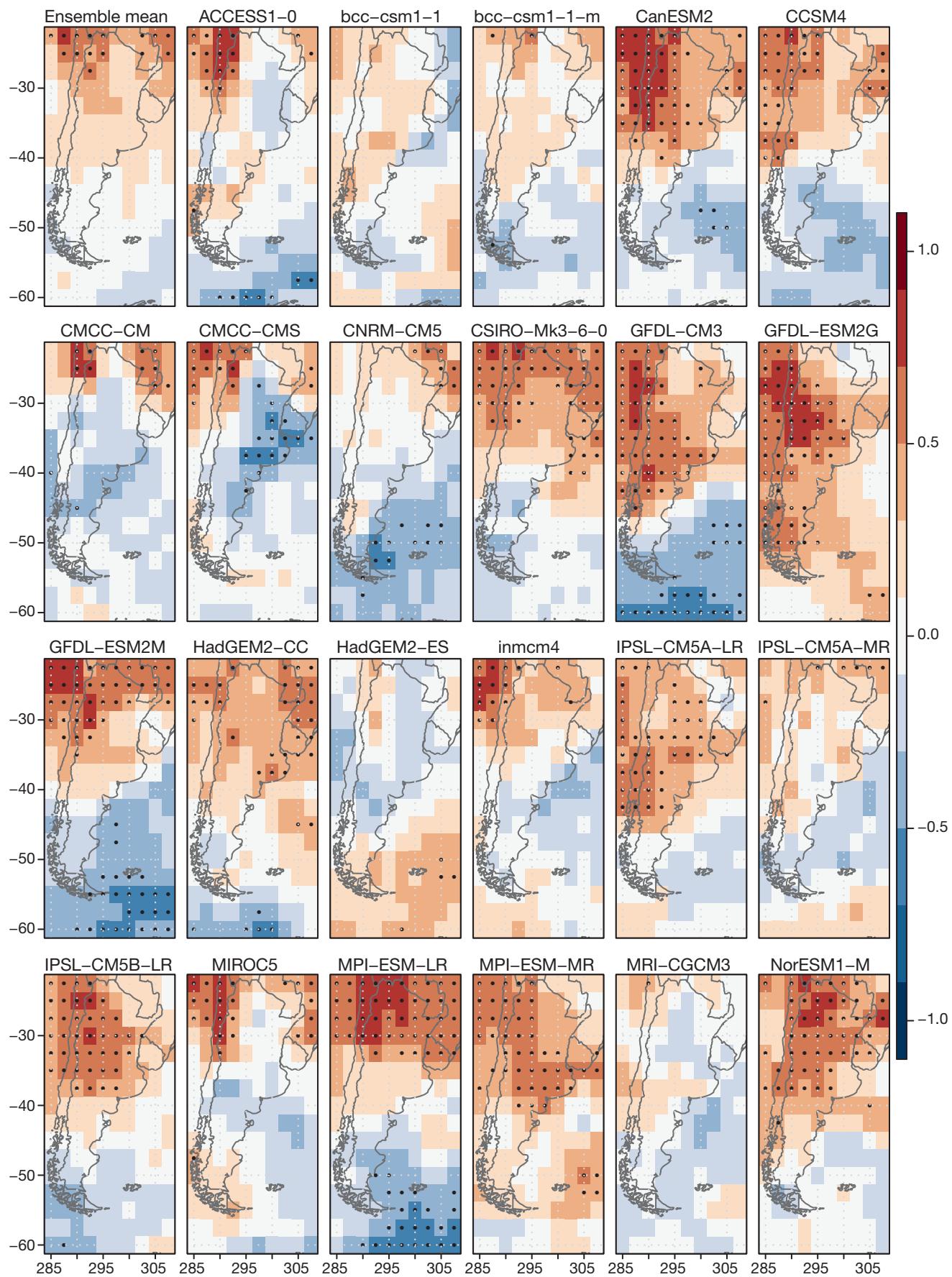


Fig. 6. Correlation between modeled SST in the El Niño 3.4 region and warm nights index of global climate models in JJA  
(dotted regions indicate significant correlation at the 0.05 significance level)

Author copy

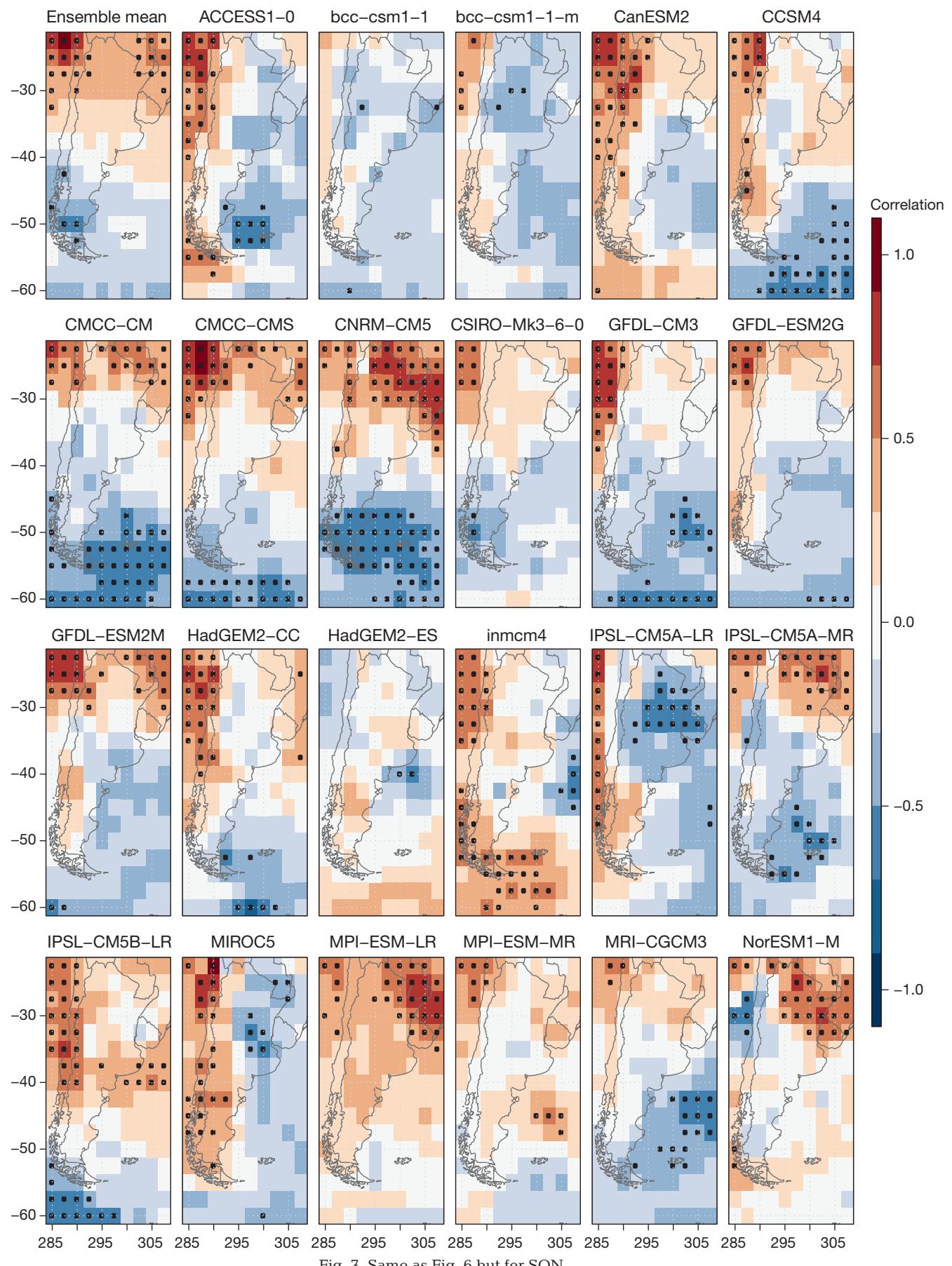


Fig. 7. Same as Fig. 6 but for SON

Author copy

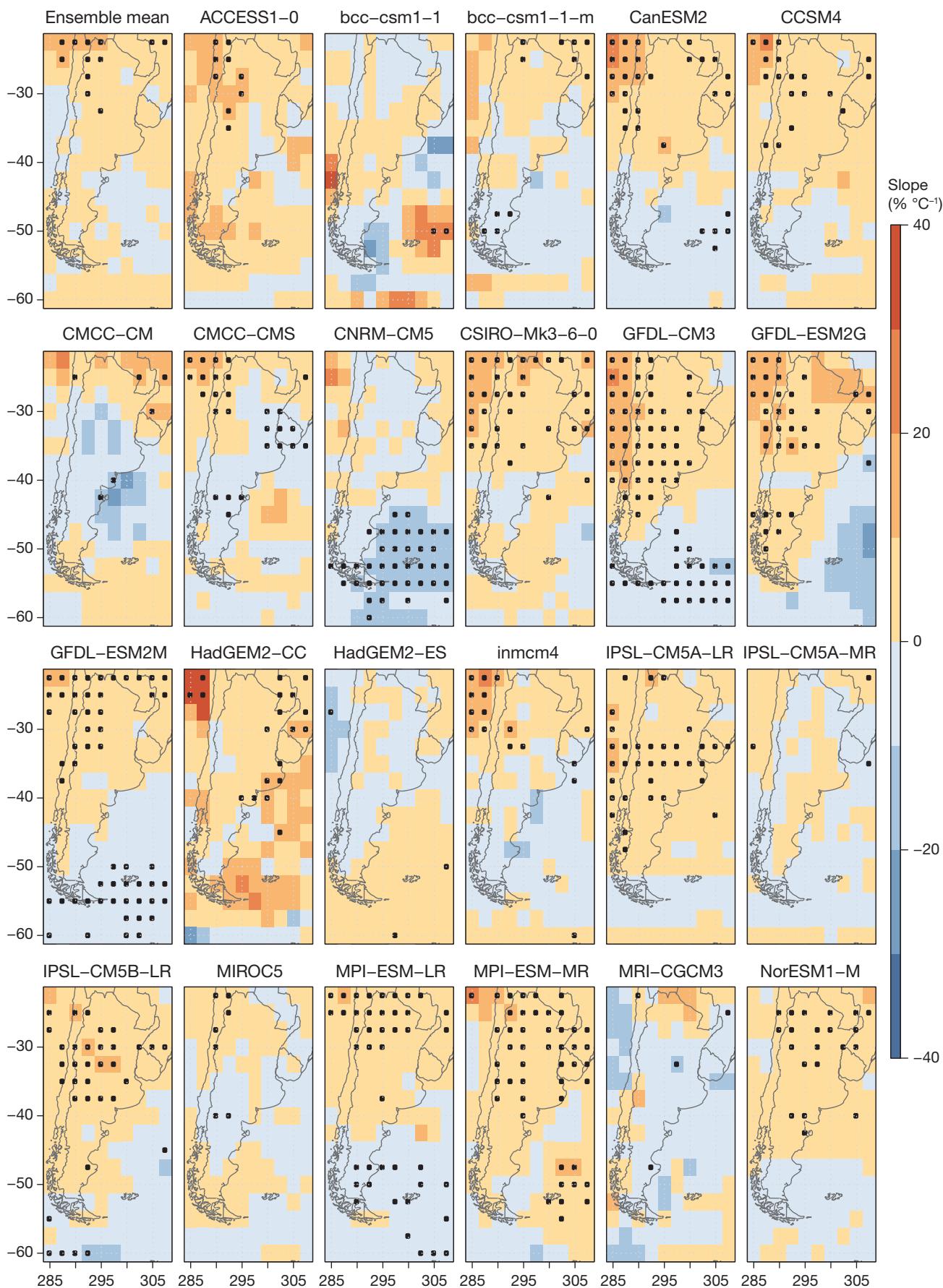


Fig. 8. Slope of the quantile regression for the 90th percentile between modeled SST in the El Niño 3.4 region and warm nights index of global climate models in JJA ( $\% \text{ } ^\circ\text{C}^{-1}$ ) (dotted regions indicate significant correlation at the 0.05 significance level)

Author copy

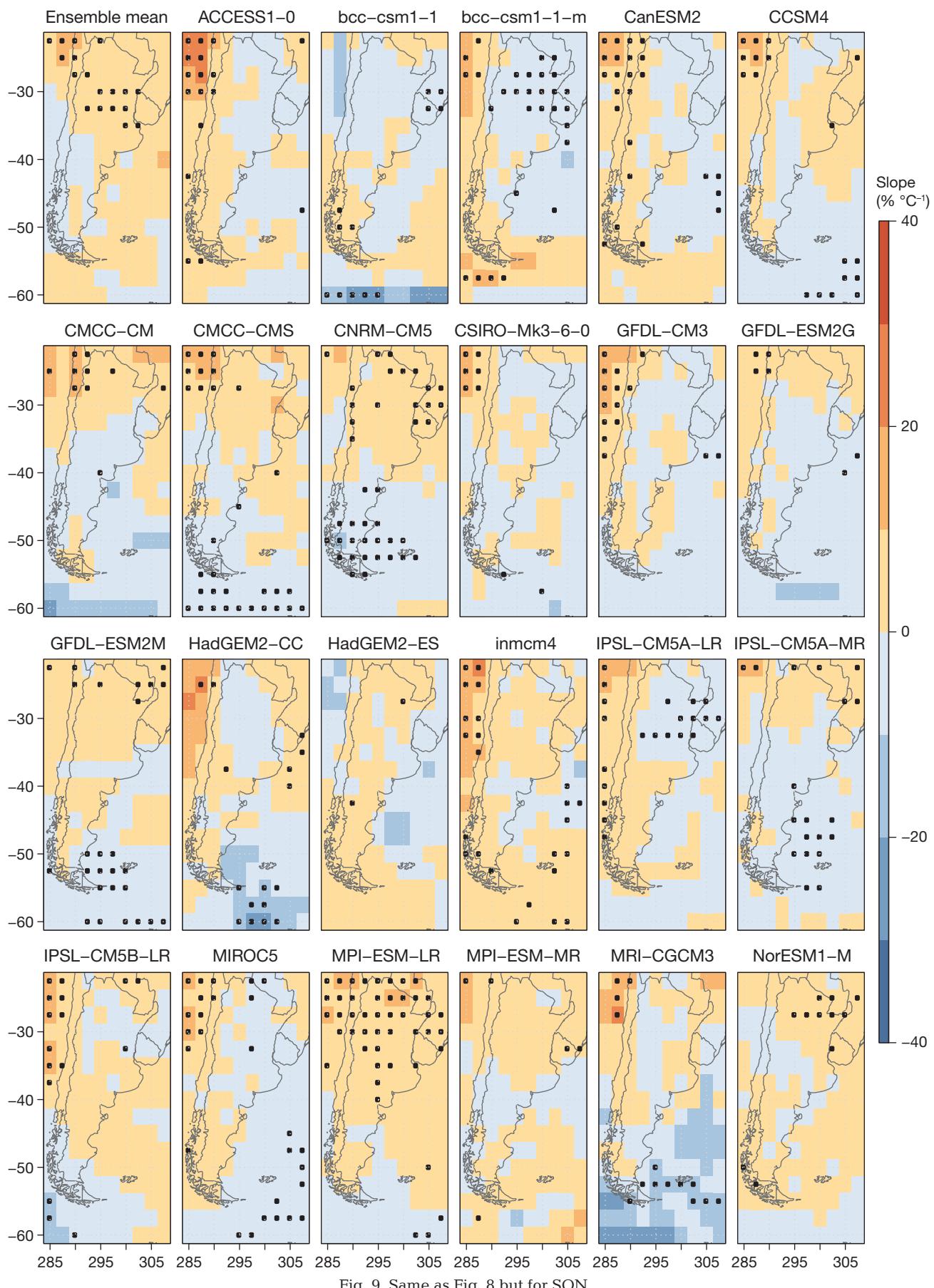


Fig. 9. Same as Fig. 8 but for SON

other metrics are presented for all seasons and extreme indices (Figs. S1–S4). Moreover, in the Supplement, we evaluated the skill of the GCMs in representing the individual variability of the SST3.4 and extreme indices (Text S1, Table S1, Figs. S5 & S6). We found the correlations between SST3.4 and warm days were correctly simulated for many models and the ensembles in DJF and MAM, with similar performances according to the metrics (Figs. S1 & S2). This result will allow more robust analysis of the future projections of these relationships.

Among the models identified with correlation patterns similar to the observed, CSIRO-Mk3-6-0 is highlighted for presenting the best performance, taking into account the Taylor diagram and the other metrics. Furthermore, the model has a similar skill to the ERA-Interim and NCEP1 reanalyses. Other GCMs

which maintain a substantial agreement between observed and modeled correlations for TN90p in JJA are NorESM1-M, HadGEM2-CC, IPSL-CM5B-LR, and CanESM2. On the other hand, MPI-ESM-LR showed higher values of the  $\kappa$  coefficient, map-curves, and spatial correlation but presented an SD higher than the observed, as shown in the Taylor diagram. In spring, the visual analysis matched with the quantitative analysis well, since both analyses indicated that the CNRM-CM5 model has the most outstanding performance.

The GCMs presented even more difficulties in representing the link between the higher frequencies of occurrences of extreme indices and SST3.4 (Fig. 10, Figs. S3 & S4). The regressions for the warm days and warm nights indices in autumn and winter, respectively, were best simulated by the models. As we

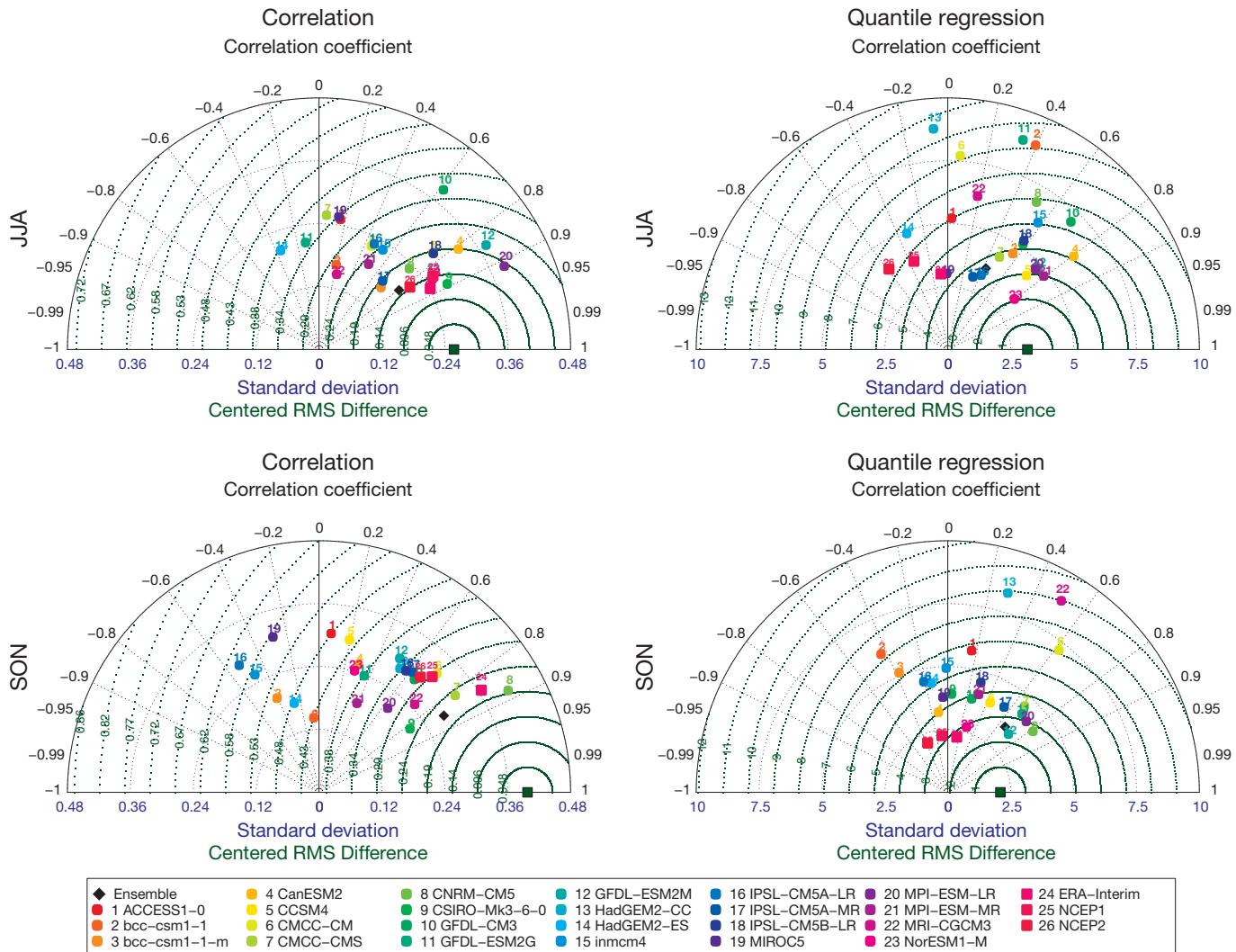


Fig. 10. Taylor diagrams comparing the observed correlation (quantile regression) between the SST3.4 and TN90p of HADEX3 against reanalyses and modeled correlations in the left (right) column. RMS: root mean square

mentioned before, several models could represent the observed pattern of the slopes for TN90p in JJA. However, we found some differences according to the metrics, e.g. the Taylor diagram showed a good performance of NorESM1-M, while the  $\kappa$  coefficient and Mapcurves indicated a fair strength of agreement. The reanalyses also present discrepancies among metrics since the Taylor diagram represents negative spatial correlations. In SON, on the contrary, fewer models presented significant  $\kappa$  coefficients, and only MPI-ESM-LR had a moderate strength of agreement; nevertheless, the Taylor diagram indicates several GCMs and the ensemble with a good performance, e.g. GFDL-ESM2M, GFDL-ESM2G, and CNRM-CM5, among others.

As we mentioned before, 1 source of error in the GCMs that can lead to a bad performance in simulating the correlations between SST3.4 and extreme indices is a poor representation of the teleconnections, as Tedeschi & Collins (2016) concluded. In this case, it is to be expected that the GCM presented a poor representation of the correlations for all extreme indices.

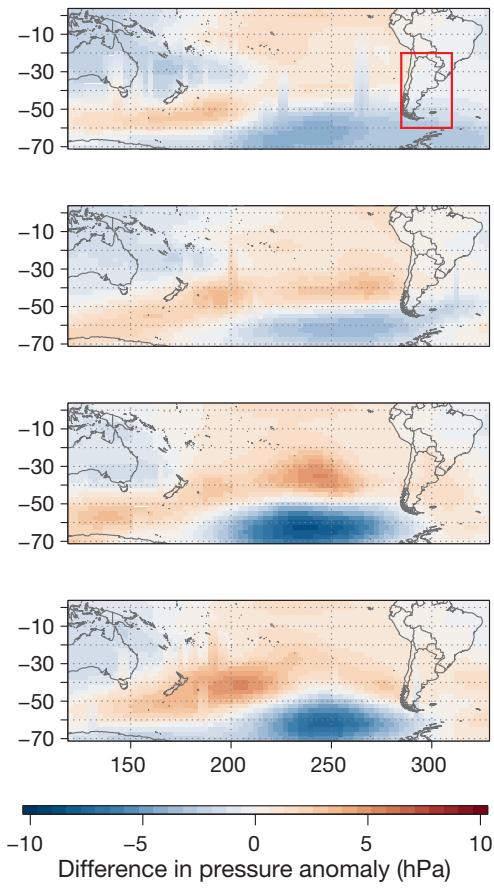


Fig. 11. Difference in the anomalies of sea level pressure (hPa) between La Niña and El Niño events for the ERA-Interim reanalysis (left panels). Spatial correlation between ERA-Interim and global climate models for the difference in sea level pressure anomalies in the Southern Hemisphere (middle panels) and only over southern South America is indicated with a red rectangle (right panels)

Fig. 11 shows the difference in sea level pressure anomalies (SLPAs) between La Niña and El Niño events, estimated for the ERA-Interim reanalysis. Cyclonic anomalies predominate in the South Pacific (centered at 60°S–250°E) during La Niña events, with an opposite pattern during El Niño. SLPAs are particularly intense during the winter and spring, which coincides with the previously detected seasons with a higher ENSO signal. For these seasons, the GCMs were able to simulate the atmospheric circulation pattern in the Southern Hemisphere, except HadGEM2-ES. Among the GCMs which were able to correctly simulate the correlation between SST3.4 and JJA TN90p, HadGEM2-CC, IPSL-CM5B-LR, and NorESM1-M stand out for their highest agreement of the spatial atmospheric circulation pattern in the Southern Hemisphere. During SON, CNRM-CM5 adequately simulates the Southern Hemisphere circulation. When we focus on how the GCMs represent the atmospheric circulation in southern South America, the performance of the models declines, and some cannot even represent it, e.g. IPSL-CM5A-LR in JJA. How-

ever, all the GCMs which presented a good performance, according to the Taylor diagrams, have a significant spatial correlation of the SLPA pattern in southern South America.

According to ERA-Interim, less precipitation is observed during La Niña events in southern Brazil, northeastern Argentina, and Uruguay compared with El Niño events (Fig. 12). The performance of the GCMs in representing the precipitation pattern varies according to the season. In particular, the model IPSL-CM5B-LR, which simulates well both the correlation between SST3.4 and JJA TN90p and the JJA SLPA patterns, fails to represent the precipitation field. A possible hypothesis about its good performance in simulating the relationship between SST3.4 and JJA TN90p is that this model simulates well the presence of some cloudiness, since it manages to represent the SLPAs but fails with the associated precipitation. Since minimum temperatures are influenced by cloud cover, IPSL-CM5B-LR manages to capture the association between warm nights and ENSO events.

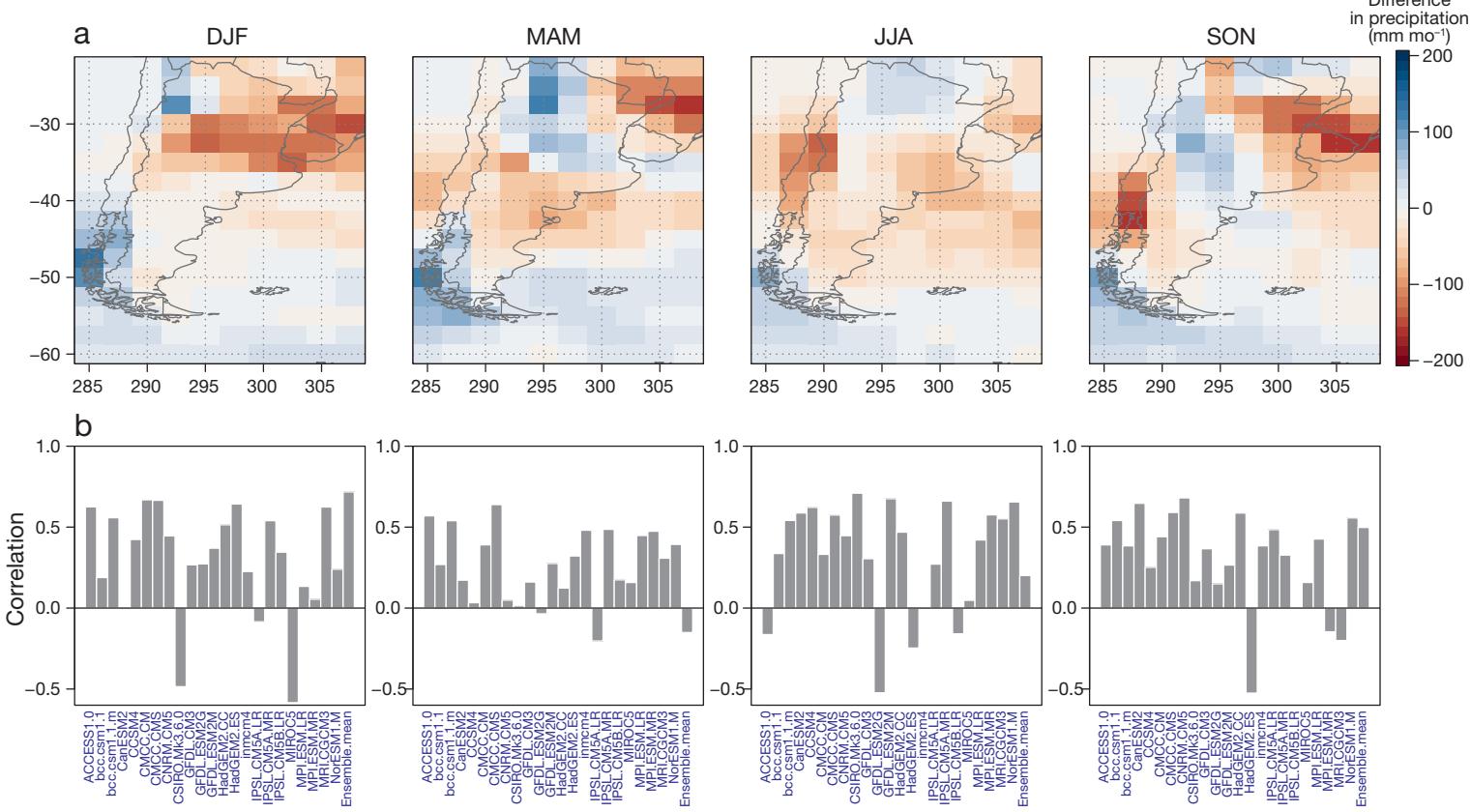


Fig. 12. (a) Difference in precipitation (mm/month) between La Niña and El Niño events for the ERA-Interim reanalysis. (b) Spatial correlation between ERA-Interim and global climate models for the difference in precipitation over southern South America

#### 4. CONCLUSIONS

In this work, we evaluated GCM performance in representing the joint variability between El Niño 3.4 SST and different extreme temperature indices. This type of analysis is of vital importance to determine whether the models simulate the occurrence of extreme events for the correct reasons, i.e. whether they occur in response to different physical processes, such as teleconnections.

ENSO is one of the main drivers of global circulation. We found that SST3.4 is mainly associated with the observed warm nights in winter and spring, and this result was consistent for the 2 methodologies: correlations and quantile regressions. Warmer conditions in the equatorial Pacific are associated with warmer nights in central and northern Argentina and Chile through teleconnections. We also observe discrepancies in the representation of these relationships between the different reanalyses. In this sense, we cannot determine which of the reanalyses is more similar to the

observed patterns because it depends on the season and methodology applied.

The performance of the GCMs in representing the associations between SST3.4 and temperature extremes was assessed through Taylor diagrams and 2 metrics: Cohen's  $\kappa$  coefficients and Mapcurves values. Most of the models were able to adequately simulate the correlations between SST3.4 and TX90p. For TN90p in JJA and SON, several models had acceptable performances, especially CNRM-CM5, CSIRO-MK3-6-0, HadGEM2-CC, and NorESM1-M. Moreover, we observed that these models were able to simulate the atmospheric circulation and precipitation patterns during ENSO events.

In general, the ensemble mean does not have an outstanding performance or clearly superior performance compared to an individual model, probably because the ensemble tends to underestimate the variability of both SST3.4 and the extreme indices (Table S1, Figs. S5 & S6).

The results achieved in this work allowed us to identify the models that can accurately simulate the link between temperature extremes and equatorial Pacific SST. However, we could not determine a single set of models with the best skills because these vary according to the extreme index and the season that is being considered. An adequate representation of the present climate provides higher reliability in future projections. In the second part of this study, these relationships will be analyzed in the future climate under different scenarios.

Moreover, additional research is pending to evaluate and compare these relationships with the new generation of CMIP6 models and considering different types of ENSO events. Several of the new CMIP6 models are more sensitive to greenhouse gas emissions than the previous generation of models. Tokarska et al. (2020) found that the models tend to overestimate recent global warming, which means that their sensitivity may be too high. These discrepancies between the CMIP5 and CMIP6 models make it necessary to evaluate and contrast both projections. In addition, other forcings can be assessed, for example, soil moisture conditions, which are strongly associated with the warm extremes of maximum temperature (Hirschi et al. 2011, Mueller & Seneviratne 2012, Whan et al. 2015, Collazo et al. 2019a). These types of studies are crucial to understanding whether climate models simulate temperature extremes, driven by physical processes, which provide higher confidence in their future projections.

**Acknowledgements.** This research was supported by CONICET PIP 0137-Res 4248/16 from the National Council of Scientific and Technical Research, Argentina, and UBACyT 2018 20020170100357BA from the University of Buenos Aires, Argentina. We acknowledge the Canadian Centre for Climate Modelling and Analysis for providing climate extreme indices. ORCID S.C.: 0000-0001-7785-2998; ORCID M.B.: 0000-0001-5385-2491; ORCID M.R.: 0000-0003-2588-6234.

## LITERATURE CITED

- Agosta EA, Barrucand MG (2012) Mean winter conditions and quasi-stationary Rossby waves associated with the winter frequency of warm and cold nights in subtropical Argentina. *Geoacta* 37:147–166
- Ahmed K, Sachindra DA, Shahid S, Demirel MC, Chung ES (2019) Selection of multi-model ensemble of general circulation models for the simulation of precipitation and maximum and minimum temperature based on spatial assessment metrics. *Hydrol Earth Syst Sci* 23:4803–4824
- Alexander LV (2016) Global observed long-term changes in temperature and precipitation extremes: a review of progress and limitations in IPCC assessments and beyond. *Weather Clim Extrem* 11:4–16
- Angélil O, Perkins-Kirkpatrick S, Alexander LV, Stone D and others (2016) Comparing regional precipitation and temperature extremes in climate model and reanalysis products. *Weather Clim Extrem* 13:35–43
- Arblaster JM, Alexander LV (2012) The impact of the El Niño–Southern Oscillation on maximum temperature extremes. *Geophys Res Lett* 39:L20702
- Arora VK, Scinocca JF, Boer GJ, Christian JR and others (2011) Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. *Geophys Res Lett* 38:L05805
- Barros V, Grimm A, Doyle M (2002) Relationship between temperature and circulation in southeastern South America and its influence from El Niño and La Niña events. *J Meteorol Soc Jpn* 80:21–32
- Bentsen M, Bethke I, Debernard J, Iversen T and others (2013) The Norwegian Earth System Model, NorESM1–M—Part 1: description and basic evaluation of the physical climate. *Geosci Model Dev* 6:687–720
- Best DJ, Roberts DE (1975) Algorithm AS 89: the upper tail probabilities of Spearman's Rho. *J R Stat Soc Ser C Appl Stat* 24:377–379
- Bi D, Dix M, Marsland SJ, O'Farrell S and others (2013) The ACCESS coupled model: description, control climate and evaluation. *Aust Meteorol Oceanogr J* 63:41–64
- Black RX, Evans KJ (1998) The statistics and horizontal structure of anomalous weather regimes in the community climate model. *Mon Weather Rev* 126:841–859
- Cai W, Borlace S, Lengaigne M, van Renssch P and others (2014) Increasing frequency of extreme El Niño events due to greenhouse warming. *Nat Clim Chang* 4:111–116
- Cai W, McPhaden MJ, Grimm AM, Rodrigues RR and others (2020) Climate impacts of the El Niño–Southern Oscillation on South America. *Nat Rev Earth Environ* 1:215–231
- Christensen JH, Krishna Kumar K, Aldrian E, An SI and others (2013) Climate phenomena and their relevance for future regional climate change. In: Stocker TF, Qin D, Plattner GK, Tignor M and others (eds) *Climate change 2013: the physical science basis. Contribution of working*

- group I to the fifth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, and New York, NY
- CMCC (Centro euro-Mediterraneo sui Cambiamenti Climatici) (2013) CMCC-CMS model output prepared for CMIP5 pre-industrial control, served by ESGF. World Data Center for Climate (WDCC) at DKRZ. <https://doi.org/10.1594/WDCC/CMIP5.CMC2pc>
- Cohen J (1960) A coefficient of agreement for nominal scales. *Educ Psychol Meas* 20:37–46
- Collazo S, Barrucand M, Rusticucci M (2019a) Summer seasonal predictability of warm days in Argentina: statistical model approach. *Theor Appl Climatol* 138:1853–1876
- Collazo S, Barrucand M, Rusticucci M (2019b) Variability and predictability of winter cold nights in Argentina. *Weather Clim Extrem* 26:100236
- Collins M, An SI, Cai W, Ganachaud A and others (2010) The impact of global warming on the tropical Pacific Ocean and El Niño. *Nat Geosci* 3:391–397
- Collins WJ, Bellouin N, Doutriaux-Boucher M, Gedney N and others (2011) Development and evaluation of an Earth-System model—HadGEM2. *Geosci Model Dev* 4:1051–1075
- Congalton R, Green K (1993) A practical look at the sources of confusion in error matrix generation. *Photogramm Eng Remote Sensing* 59:641–644
- Dee DP, Uppala SM, Simmons AJ, Berrisford P and others (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol Soc* 137:553–597
- Demirel MC, Koch J, Mendiguren G, Stisen S (2018) Spatial pattern oriented multicriteria sensitivity analysis of a distributed hydrologic model. *Water* 10:1188
- DeVellis RF (2005) Inter-rater reliability. In: Kempf-Leonard K (ed) Encyclopedia of social measurement, Vol 2. Elsevier, San Diego, CA, p 317–322
- Dittus AJ, Karoly DJ, Donat MG, Lewis SC, Alexander LV (2018) Understanding the role of sea surface temperature-forcing for variability in global temperature and precipitation extremes. *Weather Clim Extrem* 21:1–9
- Donat MG, Alexander LV, Herold N, Dittus AJ (2016) Temperature and precipitation extremes in century-long gridded observations, reanalyses, and atmospheric model simulations. *J Geophys Res Atmos* 121:11174–11189
- Donner L, Wyman B, Hemler R, Horowitz L and others (2011) The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. *J Clim* 24:3484–3519
- Dufresne J, Foujols M, Denvil S, Caubel A and others (2013) Climate change projections using the IPSL-CM5 earth system model: from CMIP3 to CMIP5. *Clim Dyn* 40: 2123–2165
- Dunn RJH, Alexander LV, Donat MG, Zhang X and others (2020) Development of an updated global land in situ-based data set of temperature and precipitation extremes: HadEX3. *J Geophys Res Atmos* 125:e2019JD032263
- Dunne J, John J, Adcroft A, Griffies S and others (2012) GFDL's ESM2 global coupled climate–carbon earth system models. Part I: physical formulation and baseline simulation characteristics. *J Clim* 25:6646–6665
- Gao T, Luo M, Lau NC, Chan TO (2020) Spatially distinct effects of two El Niño types on summer heat extremes in China. *Geophys Res Lett* 47:e2020GL086982
- Garreaud RD, Vuille M, Compagnucci R, Marengo J (2009) Present-day South American climate. *Palaeogeogr Palaeoclimatol Palaeoecol* 281:180–195
- Gent PR, Danabasoglu G, Donner LJ, Holland MM and others (2011) The community climate system model version 4. *J Clim* 24:4973–4991
- Giorgi F, Gutowski WJ (2015) Regional dynamical downscaling and the CORDEX initiative. *Annu Rev Environ Resour* 40:467–490
- Grimm AM, Tedeschi RG (2009) ENSO and extreme rainfall events in South America. *J Clim* 22:1589–1609
- Grimm AM, Barros VR, Doyle ME (2000) Climate variability in southern South America associated with El Niño and La Niña events. *J Clim* 13:35–58
- Grotjahn R, Black R, Leung R, Wehner MF and others (2016) North American extreme temperature events and related large scale meteorological patterns: a review of statistical methods, dynamics, modeling, and trends. *Clim Dyn* 46:1151–1184
- Guilyardi E, Cai W, Collins M, Fedorov A and others (2012) New strategies for evaluating ENSO processes in climate models. *Bull Am Meteorol Soc* 93:235–238
- Hao L, Naiman D (2007) Quantile regression. Sage, Thousand Oaks, CA. <https://methods.sagepub.com/book/quantile-regression>
- Hargrove WW, Hoffman FM, Hessburg PF (2006) Map-curves: a quantitative method for comparing categorical maps. *J Geogr Syst* 8:187
- Hellström C, Chen D (2003) Statistical downscaling based on dynamically downscaled predictors: application to monthly precipitation in Sweden. *Adv Atmos Sci* 20:951–958
- Hirschi M, Seneviratne SI, Alexandrov V, Boberg F and others (2011) Observational evidence for soil moisture impact on hot extremes in southeastern Europe. *Nat Geosci* 4:17–21
- Horton DE, Johnson NC, Singh D, Swain DL, Rajaratnam B, Diffenbaugh N (2015) Contribution of changes in atmospheric circulation patterns to extreme temperature trends. *Nature* 522:465–469
- Jiang Z, Li W, Xu J, Li L (2015) Extreme precipitation indices over China in CMIP5 models. Part I: model evaluation. *J Clim* 28:8603–8619
- Jones PW (1999) First- and second-order conservative remapping schemes for grids in spherical coordinates. *Mon Weather Rev* 127:2204–2210
- Kalnay E, Kanamitsu M, Kistler R, Collins W and others (1996) The NCEP/NCAR 40-year reanalysis project. *Bull Am Meteorol Soc* 77:437–472
- Kanamitsu M, Ebisuzaki W, Woollen J, Yang SK, Hnilo J, Fiorino M, Potter GL (2002) NCEP-DOE AMIP-II Reanalysis (R-2). *Bull Am Meteorol Soc* 83:1631–1643
- Katragkou E, García-Díez M, Vautard R, Sobolowski S and others (2015) Regional climate hindcast simulations within EURO-CORDEX: evaluation of a WRF multi-physics ensemble. *Geosci Model Dev* 8:603–618
- Kenyon J, Hegerl GC (2008) Influence of modes of climate variability on global temperature extremes. *J Clim* 21: 3872–3889
- Kim ST, Yu JY (2012) The two types of ENSO in CMIP5 models. *Geophys Res Lett* 39:L11704
- Kim Y, Min S, Zhang X, Zwiers F, Alexander LV, Donat MG, Tung Y (2016) Attribution of extreme temperature changes during 1951–2010. *Clim Dyn* 46:1769–1782
- King AD, van Oldenborgh GJ, Karoly DJ (2016) Climate change and El Niño increase likelihood of Indonesian heat and drought. *Bull Am Meteorol Soc* 97:S113–S117

- ↗ Kistler R, Kalnay E, Collins W, Saha S and others (2001) The NCEP-NCAR 50-year reanalysis: monthly means CD-ROM and documentation. *Bull Am Meteorol Soc* 82: 247–267
- Klein Tank AMG, Zwiers FW, Zhang X (2009) Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation. WMO/TD-No. 1500, WCDMP-No. 72. World Meteorological Organization, Geneva
- Koenker R (2005) Quantile regression. Econometric Society monographs, Vol 38. Cambridge University Press, Cambridge
- ↗ Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174
- ↗ Loikith PC, Detzer J, Mechoso CR, Lee H, Barkhordarian A (2017) The influence of recurrent modes of climate variability on the occurrence of monthly temperature extremes over South America. *J Geophys Res Atmos* 122: 10297–10311
- ↗ Lovino MA, Müller OV, Berbery EH, Müller GV (2018) Evaluation of CMIP5 retrospective simulations of temperature and precipitation in northeastern Argentina. *Int J Climatol* 38:e1158–e1175
- ↗ Luo M, Lau NC (2019) Amplifying effect of ENSO on heat waves in China. *Clim Dyn* 52:3277–3289
- ↗ Luo M, Lau NC (2020) Summer heat extremes in northern continents linked to developing ENSO events. *Environ Res Lett* 15:074042
- ↗ Menzel A, Sparks TH, Estrella N, Koch E and others (2006) European phenological response to climate change matches the warming pattern. *Glob Change Biol* 12: 1969–1976
- ↗ Molina OD, Bernhofer C (2019) Projected climate changes in four different regions in Colombia. *Environ Syst Res* 8:33
- ↗ Montini TL, Jones C, Carvalho LM (2019) The South American low-level jet: a new climatology, variability, and changes. *J Geophys Res Atmos* 124:1200–1218
- ↗ Mueller B, Seneviratne SI (2012) Hot days induced by precipitation deficits at the global scale. *Proc Natl Acad Sci USA* 109:12398–12403
- ↗ Müller GV, Nuñez MN, Seluchi ME (2000) Relationship between ENSO cycles and frost events within the Pampa Húmeda region. *Int J Climatol* 20:1619–1637
- ↗ Murari KK, Sahana AS, Daly E, Ghosh S (2016) The influence of the El Niño Southern Oscillation on heat waves in India. *Meteorol Appl* 23:705–713
- ↗ Nowosad J, Stepinski TF (2018) Spatial association between regionalizations using the information-theoretical V-measure. *Int J Geogr Inf Sci* 32:2386–2401
- ↗ Orlowsky B, Seneviratne SI (2012) Global changes in extreme events: regional and seasonal dimension. *Clim Change* 110:669–696
- ↗ Perry SJ, McGregor S, Gupta AS, England MH (2017) Future changes to El Niño–Southern Oscillation temperature and precipitation teleconnections. *Geophys Res Lett* 44:10608–10616
- ↗ Poggio L, Gimona A (2015) Downscaling and correction of regional climate models outputs with a hybrid geostatistical approach. *Spat Stat* 14:4–21
- Pontius RG Jr (2000) Quantification error versus location error in comparison of categorical maps. *Photogramm Eng Remote Sensing* 66:1011–1016
- ↗ Power S, Delage F, Chung C, Kociuba G, Keay K (2013) Robust twenty-first-century projections of El Niño and related precipitation variability. *Nature* 502:541–545
- ↗ Rayner NA, Parker DE, Horton EB, Folland CK and others (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J Geophys Res* 108:4407
- Rodriguez RN, Yao Y (2017) Five things you should know about quantile regression. [https://uisug.org.uiowa.edu/sites/uisug.org.uiowa.edu/files/wysiwyg\\_uploads/quantilesas0525-2017.pdf](https://uisug.org.uiowa.edu/sites/uisug.org.uiowa.edu/files/wysiwyg_uploads/quantilesas0525-2017.pdf)
- ↗ Ropelewski CF, Halpert MS (1987) Global and regional scale precipitation patterns associated with the El Niño–Southern Oscillation. *Mon Weather Rev* 115:1606–1626
- ↗ Rotstayn LD, Collier MA, Dix MR, Feng Y and others (2010) Improved simulation of Australian climate and ENSO-related rainfall variability in a global climate model with an interactive aerosol treatment. *Int J Climatol* 30: 1067–1088
- ↗ Rusticucci M, Vargas W (2002) Cold and warm events over Argentina and their relationship with the ENSO phases: risk evaluation analysis. *Int J Climatol* 22:467–483
- ↗ Rusticucci M, Barrucand M, Collazo S (2017) Temperature extremes in the Argentina central region and their monthly relationship with the mean circulation and ENSO phases. *Int J Climatol* 37:3003–3017
- ↗ Scoccimarro E, Gualdi S, Bellucci A, Sanna A and others (2011) Effects of tropical cyclones on ocean heat transport in a high-resolution coupled general circulation model. *J Clim* 24:4368–4384
- ↗ Seneviratne SI, Donat MG, Pitman AJ, Knutti R, Wilby RL (2016) Allowable CO<sub>2</sub> emissions based on regional and impact-related climate targets. *Nature* 529:477–483
- ↗ Sillmann J, Kharin VV, Zhang X, Zwiers FW, Branaugh D (2013a) Climate extremes indices in the CMIP5 multi-model ensemble: Part 1. Model evaluation in the present climate. *J Geophys Res Atmos* 118:1716–1733
- ↗ Sillmann J, Kharin VV, Zwiers FW, Zhang X, Branaugh D (2013b) Climate extremes indices in the CMIP5 multi-model ensemble: Part 2. Future climate projections. *J Geophys Res Atmos* 118:2473–2493
- ↗ Silva GAM, Ambrizzi T (2006) Inter-El Niño variability and its impact on the South American low-level jet east of the Andes during austral summer—two case studies. *Adv Geosci* 6:283–287
- ↗ Silvestri GE (2005) Comparison between winter precipitation in southeastern South America during each ENSO phase. *Geophys Res Lett* 32:L05709
- ↗ Sim J, Wright CC (2005) The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 85:257–268
- ↗ Sparovek G, De Jong Van Lier Q, Dourado Neto D (2007) Computer assisted Koeppen climate classification: a case study for Brazil. *Int J Climatol* 27:257–266
- ↗ Spearman C (1904) The proof and measurement of association between two things. *Am J Psychol* 15:72–101. [www.jstor.org/stable/1412159](http://www.jstor.org/stable/1412159)
- ↗ Stevenson SL (2012) Significant changes to ENSO strength and impacts in the twenty-first century: results from CMIP5. *Geophys Res Lett* 39:L17703
- ↗ Tang J, Li Q, Wang S, Lee DK and others (2016) Building Asian climate change scenario by multi-regional climate models ensemble. Part I: surface air temperature. *Int J Climatol* 36:4241–4252
- ↗ Tapiador FJ, Moreno R, Navarro A, Sánchez JL, García-Ortega E (2019) Climate classifications from regional and global models: performances for present climate estimates and expected changes in the future at high

- spatial resolution. *Atmos Res* 228:107–121
- Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. *J Geophys Res* 106: 7183–7192
- Tedeschi RG, Collins M (2016) The influence of ENSO on South American precipitation during austral summer and autumn in observations and models. *Int J Climatol* 36:618–635
- Tencer B, Bettolli ML, Rusticucci M (2016) Compound temperature and precipitation extreme events in southern South America: associated atmospheric circulation, and simulations by a multi-RCM ensemble. *Clim Res* 68: 183–199
- Tokarska KB, Stolpe MB, Sippel S, Smith CJ, Lehner F, Knutti R (2020) Past warming trend constrains future warming in CMIP6 models. *Sci Adv* 6:eaaz9549
- Töyrä J, Pietroniro A, Bonsal B (2005) Evaluation of GCM simulated climate over the Canadian Prairie provinces. *Can Water Resour J* 30:245–262
- Vargas W, Penalba O, Minetti J (1999) Monthly precipitation in areas of Argentina and the ENSO: a focus on decision problems. *Meteorologica* 24:3–22
- Vera C, Silvestri G, Barros V, Carril A (2004) Differences in El Niño response over the Southern Hemisphere. *J Clim* 17:1741–1753
- Viale M, Valenzuela R, Garreaud RD, Ralph RM (2018) Impacts of atmospheric rivers on precipitation in southern South America. *J Hydrometeorol* 19:1671–1687
- Voldoire A, Sanchez-Gomez E, Salas y Mélia D, Decharme B and others (2013) The CNRM-CM5.1 global climate model: description and basic evaluation. *Clim Dyn* 40: 2091–2121
- Volodin EM, Dianskii NA, Gusev AV (2010) Simulating present-day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations. *Izv, Atmos Ocean Phys* 46:414–431
- Watanabe M, Suzuki T, Oishi R, Komuro Y and others (2010) Improved climate simulation by MIROC5: mean states, variability, and climate sensitivity. *J Clim* 23: 6312–6335
- Whan K, Zscheischler J, Orth R, Shongwe M, Rahimi M, Asare EO, Seneviratne SI (2015) Impact of soil moisture on extreme maximum temperatures in Europe. *Weather Clim Extrem* 9:57–67
- Wu T, Song L, Li W, Wang Z and others (2014) An overview of BCC climate system model development and application for climate change studies. *J Meteor Res* 28:34–56
- Xin X, Wu T, Zhang J (2013) Introduction of CMIP5 experiments carried out with the climate system models of Beijing climate center. *Adv Clim Chang Res* 4:41–49
- Yukimoto S, Adachi Y, Hosaka M, Sakami T and others (2012) A new global climate model of the meteorological research institute: MRI-CGCM3—model description and basic performance. *J Meteorol Soc Jpn* 90A:23–64
- Zanchettin D, Rubino A, Matei D, Bothe O, Jungclaus J (2013) Multidecadal-to-centennial SST variability in the MPI-ESM simulation ensemble for the last millennium. *Clim Dyn* 40:1301–1318

*Editorial responsibility:* Eduardo Zorita,  
Geesthacht, Germany  
*Reviewed by:* 2 anonymous referees

*Submitted:* September 15, 2020

*Accepted:* January 20, 2021

*Proofs received from author(s):* April 15, 2021