RMetS
Royal Meteorological Society

Evaluation of CMIP5 retrospective simulations of temperature and precipitation in northeastern Argentina

Miguel A. Lovino,[a]*[ORCID] Omar V. Müller,[a] Ernesto H. Berbery[b] and Gabriela V. Müller[a]

[a] *Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Centro de Estudios de Variabilidad y Cambio Climático (CEVARCAM), Universidad Nacional del Litoral, Facultad de Ingeniería y Ciencias Hídricas, Santa Fe, Argentina*
[b] *Earth System Science Interdisciplinary Center/Cooperative Institute for Climate and Satellites-Maryland, University of Maryland, College Park, MD, USA*

ABSTRACT: It is generally agreed that models that better simulate historical and current features of climate should also be the ones that more reliably simulate future climate. This article describes the ability of a selection of global climate models (GCMs) of the Coupled Model Intercomparison Project Phase 5 (CMIP5) to represent the historical and current mean climate and its variability over northeastern Argentina, a region that exhibits frequent extreme events. Two types of simulations are considered: Long-term simulations for 1901–2005 in which the models respond to climate forcing (e.g. changes in atmospheric composition and land use) and decadal simulations for 1961–2010 that are initialized from observed climate states. Monthly simulations of precipitation and temperature are statistically evaluated for individual models and their ensembles.

Subsets of models that best represent the region's climate are further examined. First, models that have a Nash–Sutcliffe efficiency of at least 0.8 are taken as a subset that best represents the observed temperature fields and the mean annual cycle. Their temperature time series are in phase with observations ($r > 0.92$), despite systematic errors that if desired can be corrected by statistical methods. Likewise, models that have a precipitation Pearson correlation coefficient of at least 0.6 are considered that best represent regional precipitation features. GCMs are able to reproduce the annual precipitation cycle, although they underestimate precipitation amounts during the austral warm season (September through April) and slightly overestimate the cold season rainfall amounts. The ensembles for the subsets of models achieve the best evaluation metrics, exceeding the performance of the overall ensembles as well as those of the individual models.

KEY WORDS   temperature; precipitation; global climate models; Coupled Model Intercomparison Project Phase 5 (CMIP5); long-term simulations; decadal simulations

Received 12 April 2017; Revised 13 December 2017; Accepted 2 January 2018

1. Introduction

Southeastern South America presents large departures from the mean climate at different time scales leading to frequent extreme events, including floods, droughts, and heat waves that affect natural and human systems (Magrin *et al.,* 2014). Several observational studies have found that extreme events have increased in their frequency and severity. This trend may become more pronounced in the coming decades as discussed by Seneviratne *et al.* (2012), Cavalcanti *et al.* (2015) and Carril *et al.* (2016). If so, these changes are likely to disrupt hydrological systems affecting food production in one of the most fertile plains in the world. Research on regional climate variability and change then becomes a necessity in order to transfer scientific knowledge to decision-making processes.

Projections of future climate are produced with numerical simulations based on global climate models (GCMs), whose reliability depends on their ability to reproduce historical and current features of climate. To this end, the

World Climate Research Programme's Working Group on Coupled Modelling promoted a set of experiments known as the fifth phase of the Coupled Model Intercomparison Project (CMIP5; Taylor *et al.,* 2012a), following a previous version known as CMIP3 (Meehl *et al.,* 2007). In general, models have less skill to simulate precipitation than they do for temperature because the temperature is obtained from a thermodynamical balance, while precipitation results from simplified parameterizations approximating actual processes (Flato *et al.,* 2013; see also references therein). These authors showed that the CMIP5 models reproduce global scale patterns of surface temperature and that at regional scales, the CMIP5 simulations of temperature and large-scale precipitation have improved over those of CMIP3.

Studies focused on southeastern South America have reported that CMIP5 models properly simulate regional mean temperature (Flato *et al.,* 2013; Kumar *et al.,* 2014). de Barros Soares *et al.* (2017) have shown that most CMIP5 models capture the positive trends in temperature. CMIP5 models simulate precipitation with lower relative errors and smaller inter-model dispersion than CMIP3 in annual and summer precipitation patterns (Gulizia

* Correspondence to: M. A. Lovino, CONICET, CEVARCAM, FICH, UNL, Ruta Nacional N° 168, Km 472.4, CC 217, Ciudad Universitaria, 3000 Santa Fe, Argentina. E-mail: mlovino@unl.edu.ar
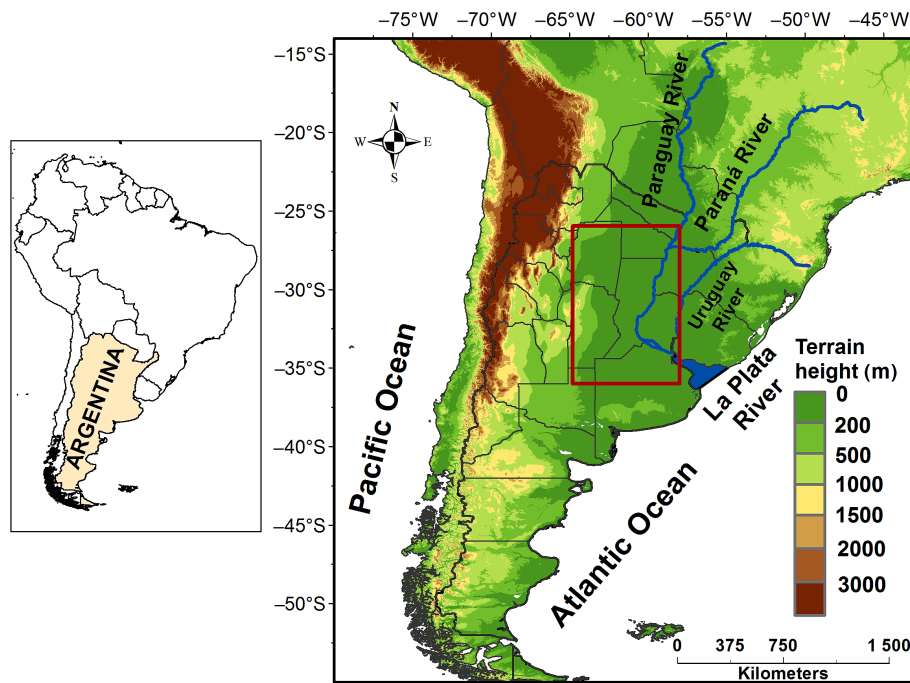
Figure 1. Topography map of the southern portion of South America together with the main rivers that drain into the La Plata River. The red box highlights the study region in northeastern Argentina.

and Camilloni, 2015; Díaz and Vera, 2017). Yet in the eastern portion of southeastern South America, a CMIP5 multi-model ensemble underestimates the annual cycle of precipitation as well as drought events (Penalba and Rivera, 2013, 2016). Consistent with these findings, Maenza *et al.* (2017) found that while 15 CMIP5 models and their ensemble capture the low-frequency variations of wet season precipitation, they underestimate precipitation values over the western Pampas of Argentina. Trends in precipitation for the period 1902–2005 have been discussed by Vera and Díaz (2015) who reported that an ensemble of CMIP5 models identified the positive change of precipitation, but with a weaker magnitude than that in observations. Furthermore, ensembles of regional climate models coordinated under the project CLARIS-LPB (see Solman *et al.,* 2013; Sánchez *et al.,* 2015; Boulanger *et al.,* 2016 and references therein) show that the spatial and temporal patterns provide more detail although they still reflect the biases of the driving GCMs.

Many studies that evaluate the performance of CMIP5 models have focused on long-term simulations. However, previous observational studies in southeastern South America have shown that decadal variability plays a key role in the region's climate (Baethgen and Goddard, 2013; Grimm and Saboia, 2015). The CMIP5 decadal simulations are expected to reproduce the decadal variability since the simulations do not only respond to external forcing but also consider internal interactions within the climate system (Taylor *et al.,* 2012a). Thus, decadal simulations potentially track the actual trajectory of climate change and explore climate predictability on decadal to multi-decadal time scales.

Argentina is a leading cereal and meat exporter, considered one of the breadbaskets of the world that plays a strategic role in the global food security (Fischer *et al.,* 2014). This study will focus on an extensive portion of the Pampas in northeastern Argentina, that presents three remarkable features: (1) it produces more than 80% of the agricultural yield (mostly soybeans, maize, sunflower, and wheat) and cattle raising of the country (Agrofy News, 2017; Ministry of Agroindustry, 2017); (2) it experiences hydro-climate variability on different time scales (i.e. interannual and decadal) and experiences extreme events that may increase in frequency and severity in the coming years and (3) it concentrates more than 90% of Argentina's population, which together with the consequent concentration of economic activities make it particularly vulnerable to hydro-climate variability and extremes. The study region is mostly flat with almost no slopes and is delimited by 65–58°W and 36–26°S (see Figure 1). The region is relatively homogeneous from a climatic perspective, identified as a subtropical humid climate according to the Köppen−Trewartha climate classification (Jacob *et al.,* 2012; Gallardo *et al.,* 2016). The regional precipitation is mostly uniformly distributed throughout the year and the thermal gradient is latitudinal (Berbery and Barros, 2002; Caffera and Berbery, 2006).

The main objective of this article is to evaluate the ability of CMIP5 models to represent the historical and present climate over northeastern Argentina. In doing so, measures of statistical and spatio-temporal properties of historical long-term and decadal simulations are assessed and subsets of the best performing models are identified. Section 2 describes the CMIP5 and its models, the used data sets and the methodology of evaluation. Sections 3 and 4 assess

to what extent the GCMs reproduce the observed precipitation and temperature in long-term and decadal simulations. Finally, Section 5 presents the conclusions.

2. Data and methodology

2.1. CMIP5 models

CMIP5 provides an unprecedented collection of data sets produced with climate models. The CMIP5 experiments were performed in two groups: (1) long-term simulations with century time scales and (2) near-term simulations (also known as decadal simulations) with decadal time scales of about 10–30 years (Taylor *et al.,* 2012a). Long-term experiments simulate the response to climate forcing like the atmospheric composition (including CO_2) due to volcanic and anthropogenic influence, solar radiation, emissions or concentrations of short life species, natural and anthropogenic aerosols, and land use. The decadal experiments not only respond to climate forcing as the long-term experiments do but also they are initialized from observed states of the climate system. Thus, the decadal variability can enhance or moderate anthropogenic trends of climate change at regional scales mainly in short-term periods (Baethgen and Goddard, 2013). In this way, decadal simulations potentially track the actual trajectory of climate change (Taylor *et al.,* 2012a).

This paper evaluates a set of 27 GCMs of the CMIP5 from different modelling centres. Model types include ocean–atmosphere coupled models (AOGCMs), earth system models (ESM), and models with chemical atmospheric processes (ChemAO and ChemESM). Table 1 presents the 27 selected GCMs and their features: responsible institution, model type, horizontal and vertical resolution together with a main citation. Of the 27 GCMs, we evaluate 25 with historical long-term simulations and 7 with decadal simulations. Two GCMs (CanCM4 and CFSv2-2011) only perform decadal simulations and were included to strengthen the multi-model ensembles of decadal simulations. Following the coding established by the WCRP working group on coupled modelling, we use the simulations identified as r1i1p1, where r is the first realization number, i is initialization method indicator, and p the perturbed physics number. More details are found in Taylor *et al.* (2012b).

2.2. CMIP5 data and observations

The models evaluation is performed on monthly precipitation (*pr*) and monthly air surface temperature (*tas*) derived from historical long-term and decadal simulations. The long-term simulations extend from 1850 to 2005, i.e. from the beginning of the industrial revolution to the near present. The period 1901–2005 is evaluated since the observations are available from the beginning of the 20th Century. The 30-year hindcasts of decadal simulations start in 1961. We evaluate the 1961–2010 period by coupling the time series of the 30-year hindcasts initialized in 1961 and 1981. A common-time period 1971–2000 is selected to study climatological features (annual cycles and spatial distribution of annual means) of the present climate for both long-term and decadal simulations. This period is chosen as it is the last normal (30-year) period available on both simulations.

The observational data set of temperature and precipitation used in this study is the CRU TS 3.20 from the Climate Research Unit – University of East Anglia (Harris *et al.,* 2014). CRU TS 3.20 consists of monthly gridded data with a $0.5° \times 0.5°$ spacing extending from January 1901 to December 2011. Other gridded data sets are available, such as the Global Precipitation Climatology Centre data set (GPCC, Schneider *et al.,* 2011) and the University of Delaware Air Temperature and Precipitation database (UDEL; Matsuura and Willmott, 2009). All gridded products for southeastern South America face the problem of sparse gauge coverage at the beginning of the 20th century (e.g. Barreiro *et al.,* 2014). According to Lovino (2015), CRU TS 3.20 fits well observed monthly mean precipitation and temperature over the study region, although the most significant biases occur in the first decades of the 20th century.

2.3. Evaluation methodology

The models performance is assessed, first, contrasting mean annual cycles and spatial patterns against observations. Second, the GCMs' skill to simulate precipitation and temperature is evaluated with statistical metrics. The metrics used are the mean bias error (MBE), the mean absolute error (MAE), the root mean square error (RMSE), the centred root mean square error (CRMS), the Nash–Sutcliffe efficiency (NSE) and the Pearson correlation coefficient (*r*). A description of these metrics can be found in Déqué (2012). The NSE is presented in Nash and Sutcliffe (1970) and further discussed in Moriasi *et al.* (2007).

MBE indicates whether a GCM simulation overestimates or underestimates the observed data. MAE and RMSE represent the magnitude of the error. NSE varies between $-\infty$ and 1, 1 being the perfect score. Moriasi *et al.* (2007) suggested that model performance can be evaluated as 'satisfactory' if NSE > 0.5 and 'very good' if NSE > 0.75.

For the spatial evaluation, we compute the MBE, the RMSE and the spatial correlation between the mean annual fields of simulated and observed variables. To compute the spatial metrics, simulated data are scaled to a 1° longitude × 1° latitude grid through the inverse distance weighting (IDW; Shepard, 1968). The CRU data set is also used at 1° grid spacing. We choose 1° grid spacing as an intermediate resolution among the different resolutions of the GCMs and the observed data set (see Table 1).

The results are shown for the individual models and multi-model ensemble means. For long-term simulations, we compute the multi-model ensemble means for the 25 studied GCMs and for subsets of models selected according to the best performance to simulate the regional climate. For decadal simulations, we perform the multi-model ensemble means for the seven GCMs examined here.

Table 1. CMIP5 models studied and their attributes.

Model		Institution, country	Type	Atmospheric resolution (°lon. × °lat.) Model levels (L)	Reference
1	ACCESS 1.0	CSIRO-BOM, Australia	AOGCM	1.875 × 1.25 L38	Bi *et al.* (2013)
2	BCC-CSM1.1	BCC-CMA, China	ESM	2.81 × 2.81 L26	Xin *et al.* (2013)
3	BNU-ESM	CGCESS-BNU, China	AOGCM	2.81 × 2.81 L26	Ji *et al.* (2014)
4	CanESM2	CCCMA, Canada	ESM	2.81 × 2.81 L35	Arora *et al.* (2011)
5	CESM1-BGC	NSF-DOE-NCAR, USA	AOGCM	1.25 × 0.9424 L26	Long *et al.* (2013)
6	CESM1-FASTCHEM	NSF-DOE-NCAR, USA	ChemESM	1.25 × 0.9424 L26	Lamarque (2013)
7	CNRM-CM5.2	CNRM-CERFACS, France	AOGCM	1.41 × 1.41 L31	Voldoire *et al.* (2013)
8	CSIRO-Mk3.6.0	CSIRO-QCCCE, Australia	AOGCM	1.875 × 1.875 L18	Rotstayn *et al.* (2010)
9	GFDL-CM3	NOAA GFDL, USA	AOGCM	2.5 × 2 L48	Donner *et al.* (2011)
10	GFDL-ESM2M	NOAA GFDL, USA	ESM	2.5 × 2 L24	Dunne *et al.* (2012)
11	GISS-E2-R	NASA GISS, USA	ChemAO	2.5 × 2 L40	Miller *et al.* (2014)
12	HadGEM2-CC	MOHC, UK	ESM	1.875 × 1.25 L60	Collins *et al.* (2011)
13	HadGEM2-ES	MOHC, UK	ChemESM	1.875 × 1.25 L60	Collins *et al.* (2011)
14	INM-CM4	INM, Russia	AOGCM	2 × 1.5 L21	Volodin *et al.* (2010)
15	MIROC5	MIROC, Japan	AOGCM	1.41 × 1.41 L40	Watanabe *et al.* (2010)
16	MIROC-ESM-CHEM	MIROC, Japan	ChemESM	2.81 × 2.81 L80	Watanabe *et al.* (2010)
17	MPI-ESM-LR	MPI-M, Germany	ESM	1.875 × 1.875 L47	Zanchettin *et al.* (2013)
18	MPI-ESM-MR	MPI-M, Germany	ESM	1.875 × 1.875 L95	Zanchettin *et al.* (2013)
19	MRI-CGCM3	MRI, Japan	AOGCM	1.125 × 1.125 L48	Yukimoto *et al.* (2012)
20	NorESM1-M	NCC, Norway	ESM	2.5 × 1.875 L26	Bentsen *et al.* (2012)
21	CCSM4[a]	NCAR,USA	AOGCM	1.25 × 0.94 L26	Gent *et al.* (2011)
22	CMCC-CM[a]	CMCC, Italy	AOGCM	0.75 × 0.75 L31	Scoccimarro *et al.* (2011)
23	EC-EARTH[a]	EC-Earth Consortium	AOGCM	1.125 × 1.125 L62	Hazeleger *et al.* (2010)
24	FGOALS-g2[a]	LASG-CESS, China	AOGCM	2.81 × 2.81	Li *et al.* (2013)
25	MIROC4h[a]	MIROC, Japan	AOGCM	0.56 × 0.56 L56	Sakamoto *et al.* (2012)
26	CanCM4[b]	CCCMA, Canada	AOGCM	2.81 × 2.81 L35	Merryfield *et al.* (2013)
27	CFSc2-2011[b]	COLA and NCEP, USA	ESM	1 × 1 L64	Saha *et al.* (2010)

[a]Models evaluated with both decadal and historical long-term simulations. [b]Models only evaluated with decadal simulations.

3. Historical long-term simulations

3.1. Temporal analysis of temperature

The evaluation metrics for the CMIP5 models that performed historical long-term simulations (Table 2) indicate that the regional mean temperature is noticeably well simulated. All models have correlations above 0.92 and low errors (the averages of the individual models' errors are MAE = 2.3 °C and RMSE = 2.7 °C). The error values are in the range of those reported by Flato *et al.* (2013) for southeastern South America. GCMs that achieve the best performances were chosen for further analysis. The selection criterion was that they must satisfy a minimum NSE of 0.8. The selected threshold is in the range of 'very good values' as recommended by Moriasi *et al.* (2007) for general performance ratings. Nine models, shaded in Table 2, meet the requirement. The selected models have high correlations with observations and very low statistical errors (1.40 °C < MAE < 1.72 °C and 1.78 °C < RMSE < 2.10 °C). The 9-model ensemble achieved the best evaluation metrics, exceeding those of the individual members. Of the individual models, CESM1-BGC and CCSM4 models (rows 5 and 21) show the best skill scores, with an efficiency of 0.86, the lower errors, and correlations of 0.94.

Figure 2 presents the time series of areal-averaged annual and spring mean temperature for the ensemble of the 25 models and the ensemble of the subset of 9

models with the better performance. In both time series, the 25-model ensemble presents a systematic error, which is significantly reduced with the 9-model ensemble. These results are consistent with the reduction of the MBE from 0.68 to 0.02 °C (Table 2). Figure 2(a) shows that the ensembles depict lower variability than observations, but still reproduce the trend in regional annual mean temperature in the latter part of the period. This warming trend is captured by all the selected models, although with different magnitudes (not shown). It has been reported that northeastern Argentina registered the most significant positive trends in spring mean temperature (Lovino, 2015). Figure 2(b) shows that the 9-model ensemble correctly recognizes this trend. Similar results were reported by de Barros Soares *et al.* (2017) who showed that most CMIP5 historical simulations reproduce the observed annual and spring temperature trends in southeastern South America.

Scatterplots of the simulated and observed temperature of the two multi-model ensembles are presented in Figure 3. For reference, similar pictures are shown for the two single models that exhibit the best evaluation metrics, the CCSM4 and CESM1-BGC. Figures 3(a) and (b) show that the ensembles achieve a low dispersion around the linear regression line ($R^2 = 0.93$). The dispersion of the data is slightly higher in cold months than in warm months. The regression line in Figure 3(a) shows that the 25-model ensemble overestimates temperature on warm months.

Table 2. Statistical evaluation metrics between the areal-averaged time series of mean temperature observed and simulated by historical long-term simulations (1901–2005). The shaded rows correspond to the models that achieve the best evaluation metrics, performing the 9-model ensemble mean.

Model		MBE (°C/month)	MAE (°C/month)	RMSE (°C/month)	NSE	Correlation coefficient
1	ACCESS 1.0	0.87	1.70	2.10	0.81	0.94
2	BCC-CSM1.1	−1.40	2.01	2.51	0.71	0.94
3	BNU-ESM	0.25	2.46	2.97	0.61	0.94
4	CanESM2	−1.25	1.80	2.31	0.77	0.94
5	CESM1-BGC	0.31	1.41	1.78	0.86	0.94
6	CESM1-FASTCHEM	0.06	1.44	1.85	0.85	0.94
7	CNRM-CM5.2	0.47	1.74	2.18	0.79	0.93
8	CSIRO-Mk3.6.0	1.64	2.59	3.08	0.58	0.94
9	GFDL-CM3	1.54	2.30	2.82	0.65	0.93
10	GFDL-ESM2M	3.27	3.49	4.06	0.27	0.92
11	GISS-E2-R	3.49	3.76	4.42	0.14	0.93
12	HadGEM2-CC	−0.01	1.69	2.09	0.81	0.94
13	HadGEM2-ES	0.22	1.78	2.21	0.79	0.94
14	INM-CM4	−3.81	3.90	4.67	0.04	0.94
15	MIROC5	4.05	4.08	4.45	0.13	0.94
16	MIROC-ESM-CHEM	1.06	1.72	2.10	0.81	0.94
17	MPI-ESM-LR	1.74	2.23	2.73	0.67	0.93
18	MPI-ESM-MR	2.03	2.48	3.08	0.58	0.92
19	MRI-CGCM3	1.70	2.21	2.67	0.69	0.93
20	NorESM1-M	0.23	1.41	1.79	0.86	0.93
21	CCSM4	0.12	1.40	1.81	0.86	0.94
22	CMCC-CM	−0.26	1.50	1.91	0.84	0.93
23	EC-EARTH	0.21	1.57	2.00	0.82	0.92
24	FGOALS-g2	−3.12	3.16	3.57	0.44	0.94
25	MIROC4h	2.84	3.11	3.58	0.44	0.94
	25-model ensemble	0.68	1.40	1.70	0.87	0.94
	9-model ensemble	0.02	1.08	1.39	0.92	0.96

In the 9-model ensemble (Figure 3(b)), the regression line fits the 1 : 1 line better than the 25-model ensemble reaching a slope value $m = 1.03$. Although the ensembles fit well the data, their variability is reduced given that the ensembles diminish the dispersion of simulated temperatures keeping their monthly values in a narrow band when compared with observations. It produces a month-to-month discontinuity in the scatterplots mainly on transitions months between cold and warm seasons (for instance, see green dots representing April and October temperatures). The 9-model ensemble improves the representation of the monthly observed variability and reduces the discontinuity displayed by the scatterplot of the 25-model ensemble. Regarding the best-performance individual models, Figures 3(c) and (d) show that the two models are close to the 1 : 1 line with observations, with a very good dispersion given by the determination coefficient $R^2 \sim 0.88$. As the multi-model ensembles, the reference models fit better mean temperatures than cold or hot extremes.

Figure 4 depicts the Taylor diagram of the regional mean temperature. Taylor diagrams summarize the degree of correspondence between the observed and simulated fields in terms of their correlation, their CMRS and the ratio of their variances (Taylor, 2001). The reference point 'a' in Figure 4 shows that observations have a standard deviation of 4.8 °C. The diagram shows graphically that the 9-model ensemble mean (point 2) presents the best scores and correctly represent the reference standard deviation. The 25-model ensemble mean (point 1) exhibits similar performance than the individual models with better skill scores. All analysed GCMs are in phase with observations, with correlation coefficients varying in the range 0.92–0.94. Also, most of the models form a cluster of points as they exhibit standard deviations between 5 and 7 °C and CRMS close to 2 °C. On this cluster, CCSM4 and CESM1-BGC (blue points) stand out as the closer models to the reference point. In general terms, the evaluation metrics computed in northeastern Argentina displayed by the Taylor diagrams are consistent with those found in most regions of the world (see Kumar *et al.*, 2014), with slightly lower skills than in North America or southeastern Asia.

So far we have seen that models tend to reproduce statistical features of the observations. Next, we explore basic aspects of the observed and simulated climatology, including the mean annual cycle and spatial patterns. Figure 5(a) shows that the 25-model ensemble mean correctly fits the annual cycle, slightly overestimating temperature from November to April, in agreement with Flato *et al.* (2013) for all southern South America. The inter-model range has an average amplitude of 8 °C increasing to 10 °C on
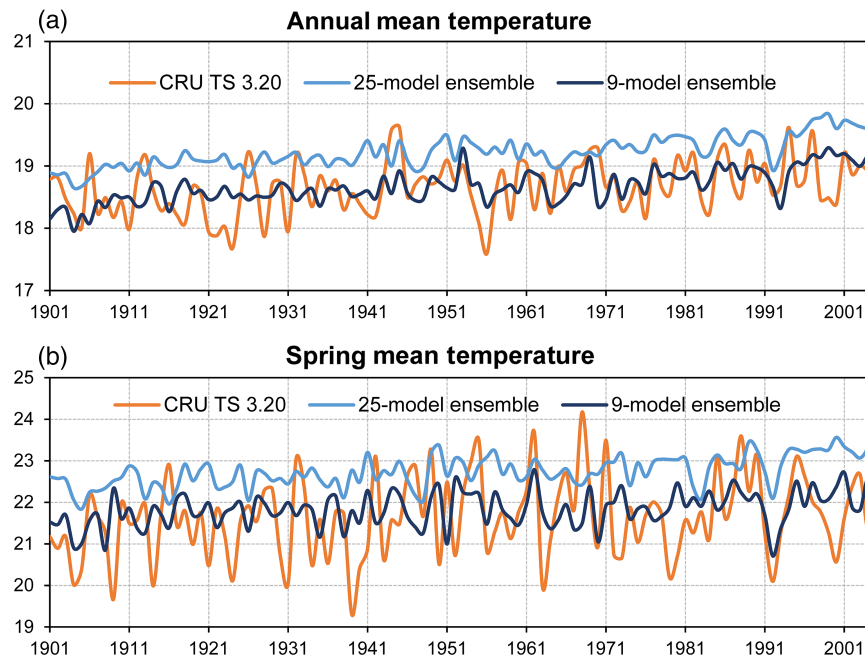
Figure 2. Areal-averaged time series of mean annual and spring temperature (in °C) observed and simulated by the multi-model ensembles of the historical long-term simulations.
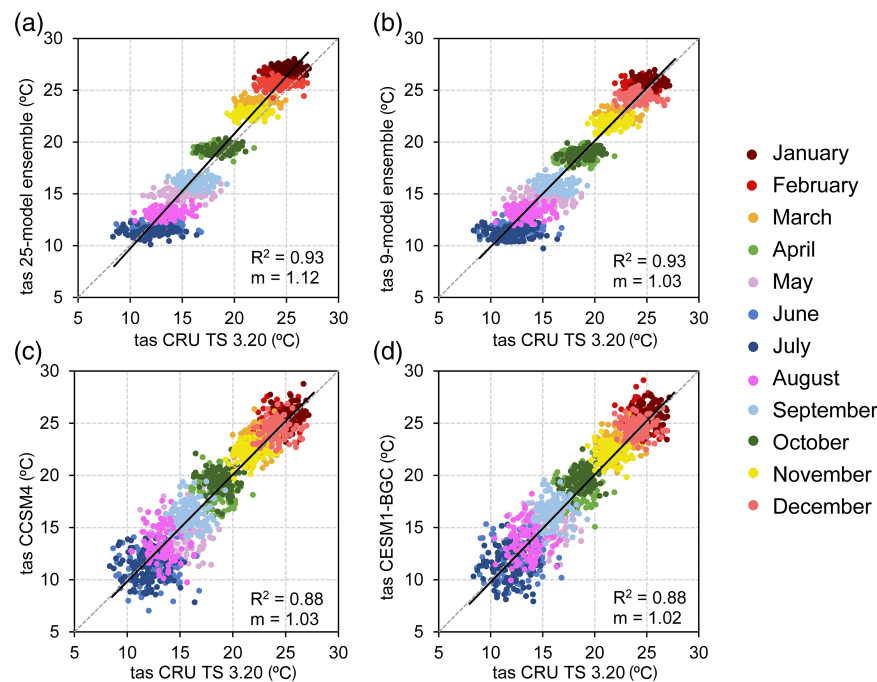

Figure 3. Scatterplots of simulated *versus* observed areal-averaged monthly mean temperature (period 1901–2005) of the multi-model ensembles (a, b) and models with the best evaluation metrics (CCSM4 in c; CESM1-BGC in d). The linear regression line is plotted and its slope m is presented together with the determination coefficient R^2 in the bottom-right corner of each scatterplot.

austral winter (June to September). Figure 5(c) indicates that the inter-model range of the nine selected GCMs is remarkably reduced to almost 2–3 °C. The 9-model ensemble mean improves the representation of the annual cycle with respect to the 25-model ensemble, mainly on warm months. Also, note that all the GCMs identify the observed annual cycle (Figures 5(b) and (d)) with different systematic errors that are quantified in Table 2 by

the MBE scores. Table 2 shows that 19 of the 25 models have positive MBEs, indicating their tendency to overestimate the average regional temperature. As Figure 5(b) shows, most of these models mainly overestimate the temperature in warm months. Figure 5(d) shows that most of the selected models reduce this error, accurately fitting observed mean values between March and December and only overestimate them in January and February. These
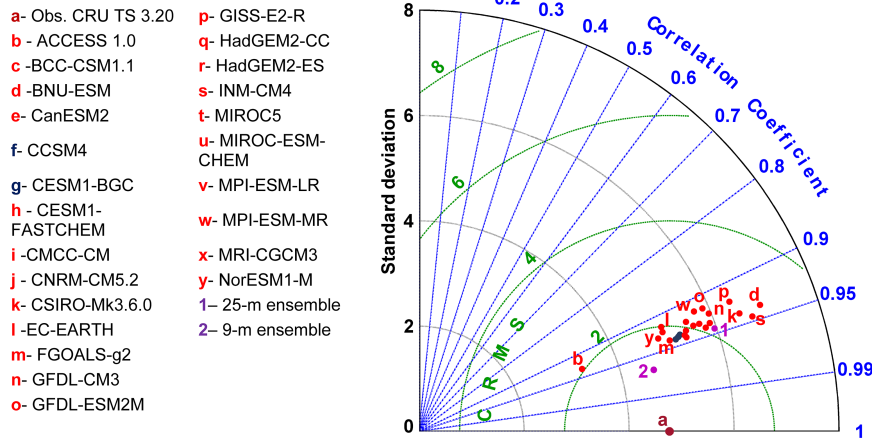
Figure 4. Taylor diagram of the mean temperature simulated by historical long-term simulations (period 1901–2005). Only the GCMs slightly distant from the central cluster of points are identified with letters.

results suggest that the nine models that achieve the best evaluation metrics and its multi-model-mean ensemble are able to properly simulate the mean annual cycle of regional temperature.

3.2. Spatial analysis of temperature

The mean observed and simulated temperature fields are presented in Figure 6. Observations (Figure 6(a)) reveal a mostly south–north gradient with a temperature range of about 6 °C. The 9-model ensemble mean (Figure 6(b)) exhibits a smoothed gradient resulting from a bias of about −2 °C towards the south and a bias of about +1 °C towards the north. Two of the models in the ensemble are shown in Figures 6(c) and (d). They are CCSM4 and CESM1-BGC, which were already shown to have the best temporal performance. The two models are able to simulate the observed field with only a slight overestimation towards the north.

The spatial evaluation metrics of the mean annual temperature fields for the nine GCMs and their ensemble are presented in Table 3. Spatial correlations vary between 0.62 and 0.96, with six models reaching values above 0.9. In general, the GCMs with higher horizontal resolution obtain the best skill scores. For example, the CCSM4 and CESM1-BGC models exhibit spatial correlations of about $r = 0.96$, with a slightly warm MBE between 0.2 and 0.4 °C, consistent with the overestimation of the spatial fields discussed above. In this case, the 9-model ensemble mean achieves similar spatial scores than the two models with the best performance. These results suggest that the 9-model ensemble does not improve the ability of the best performance individual models to reproduce the annual average temperature field.

3.3. Temporal analysis of precipitation

Metrics for the areal-averaged precipitation evaluation are presented in Table 4. The values are lower than those for temperature (for reference see Flato *et al.,* 2013). The correlation coefficients range from 0.31 to 0.64. The

RMSE and MAE metrics tend to be rather uniform among models, with RMSE varying between 42 and 53 mm and MAE between 32 and 42 mm. These errors are high but similar to those found in southeastern South America (Gulizia and Camilloni, 2015) and in other regions of the world (Kumar *et al.,* 2014). For instance, the RMSE values are lower than in Africa but greater than in North America (see also Sheffield *et al.,* 2013). The MBE metric suggests that most GCMs (18 out of 25) tend to underestimate precipitation, consistently with the results reported by Díaz and Vera (2017) for the La Plata Basin and Maenza *et al.* (2017) for the western Pampas of Argentina (south of the study region).

Statistical errors of precipitation are high, resulting in low Nash–Sutcliffe efficiencies. The correlation coefficient is more useful as a measure of the phase between time series, and the errors can be corrected using statistical methods (e.g. downscaling). Seven models with correlations higher than 0.6 are chosen for further analysis (see shaded rows in Table 4). Their statistical scores are remarkably better than for the rest of the models. This improvement is also reflected on the 7-model ensemble, which has a temporal correlation above 0.7, while the mean errors are reduced and the efficiency increases considerably.

The annual and summer precipitation time series of the 25-model ensemble and the 7-model ensemble along with observations are presented in Figure 7. The figure shows that the multi-model ensembles have lower temporal variability than observed precipitation. In the case of annual precipitation (Figure 7(a)), the 25-model ensemble underestimates the observed precipitation by 12% on average. The 7-model ensemble reduces the underestimation to just 1%. In the case of summer precipitation (Figure 7(b)), the 25-model ensemble underestimates the observed precipitation by 7% whereas the 7-model ensemble overestimates it by 13% on averages. The percentages are estimated as $\left[\left(\overline{S} - \overline{O} \right) / \overline{O} \right] x\ 100$ being $\overline{S}$ and $\overline{O}$ the simulated and observed mean values of the time
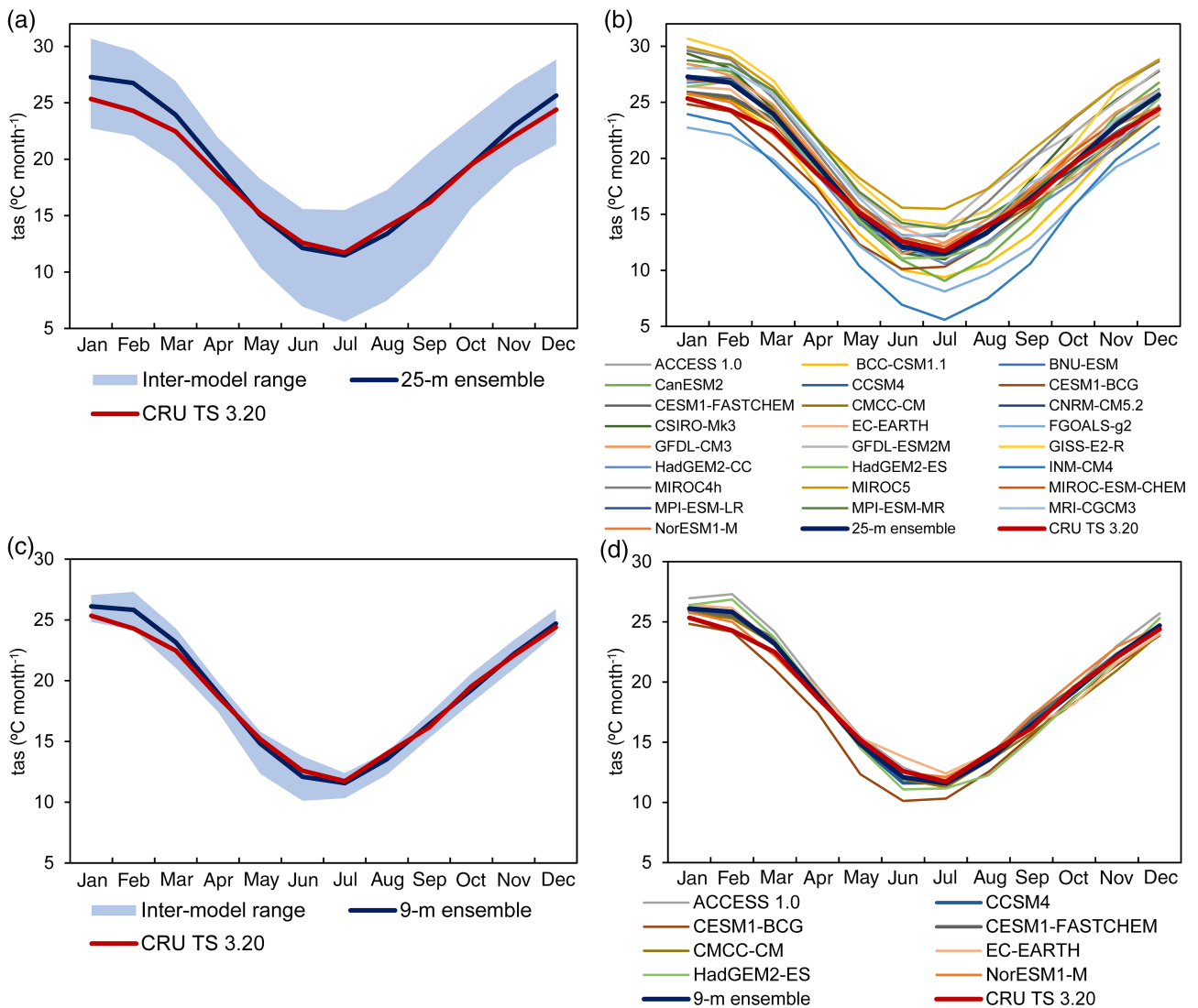
Figure 5. Annual cycle of mean temperature observed and simulated by historical long-term simulations of individual GCMs and the multi-model ensembles in the 1971−2000 period: (a) and (b) the 25 analysed GCMs; (c) and (d) the 9 GCMs with best evaluation metrics.

series, respectively. A common feature of both ensembles is that they do not reproduce accurately the positive trend in annual and summer precipitation, but weakly recognize the positive pattern of change. The seven individual models do present a trend, but it is smaller than that in the observations. As mention above, most of the historical simulations of the CMIP5 models can recognize the right sign of the summer precipitation changes in southeastern South America, although weaker than observed (Vera and Díaz, 2015). Our results suggest that the annual and summer mean trend in precipitation are weakly simulated also in the northeast region of Argentina.

Scatterplots of simulated precipitation *versus* the observed precipitation are presented in Figure 8. The scatterplot of the multi-model ensembles (Figures 8(a) and (b)) reflect the difficulties that some models have in representing observed precipitation. As in temperature, the ensembles reduce the temporal variability flattening the dispersion of the data around the simulated monthly

mean values. Thus, the scatterplot of the 25-model ensemble (Figure 8(a)) shows two quasi-horizontal bands and a severe dry bias for months with more than 100 mm of rain. Also, low precipitation monthly values (lesser than 20 mm of rain) are not simulated by the multi-model ensemble. These biases are reduced when considering the 7-model ensemble mean (Figure 8(b)) with a closer fit to the 1 : 1 reference line, mainly on rainy months. In general, the individual models in the 7-model ensemble underestimate the mean precipitation in rainy months and overestimate it in months with low rainfall, a feature also reproduced in most of the 25 studied models (not shown). This behaviour can be seen in Figures 8(c) and (d) with the scatterplots of two of the best performing models, CCSM4 and CanESM2. As the linear regression lines show, the two models underestimate precipitation higher than 150 mm and overestimate precipitation lower than 50 mm. The data dispersion is lower for low precipitation values and increases for high precipitation values,
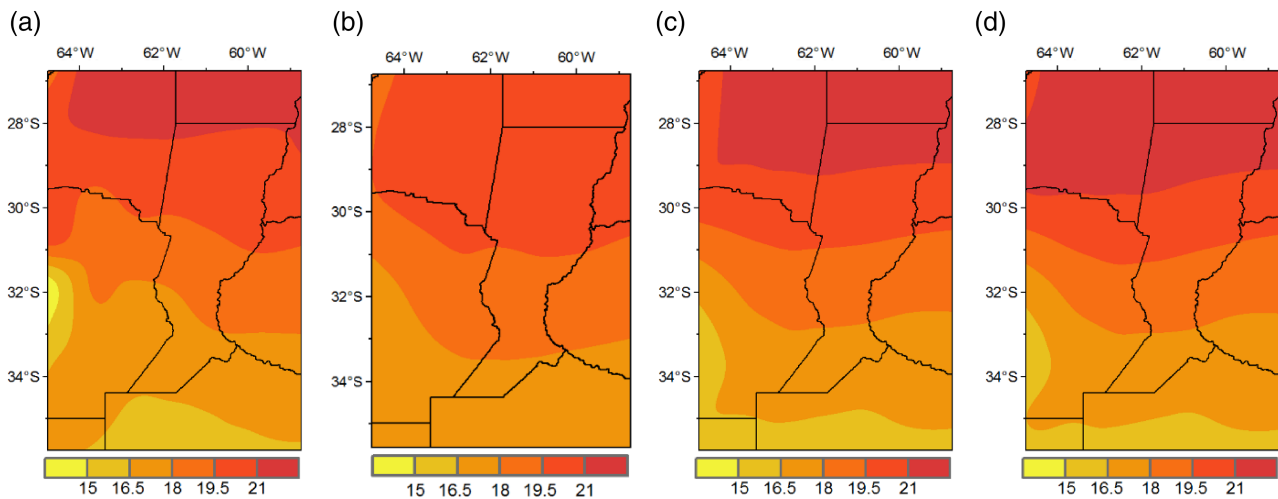
(a)

(b)

(c)

(d)

Figure 6. Spatial distribution of the annual mean temperature (in °C) observed and simulated by historical long-term simulations in the period 1971–2000 for (a) CRU TS 3.20, (b) 9-model ensemble mean, (c) CCSM4 and (d) CESM1-BGC.

Table 3. Statistical evaluation metrics between the mean fields of annual mean temperature observed and simulated by historical long-term simulations (analysed in the normal period 1971–2000).

Model	MBE	RMSE	Spatial correlation
ACCESS 1.0	0.72	1.51	0.78
CCSM4	0.23	0.65	0.96
CESM1-BGC	0.42	0.79	0.96
CESM1-FASTCHEM	0.13	0.64	0.96
CMCC-CM	−0.05	0.65	0.95
EC-EARTH	0.24	0.70	0.96
HadGEM2-CC	−0.37	1.60	0.64
MIROC-ESM-CHEM	0.68	1.81	0.62
NorESM1-M	−0.09	0.86	0.92
Ensemble mean	0.21	0.72	0.93

resulting in determination coefficients between 0.36 and 0.39. These results suggest that while individual GCMs are able to recognize the temporal variability of the observed series, they present difficulties in simulating both the rainy months and those months of low precipitation.

Figure 9 summarizes the performance of the models in a Taylor diagram. The 25-model ensemble mean (point 1) improves the CRMS and the correlation values of each GCM but underestimates the standard deviation of the reference value. The temporal variability is also reduced as discussed earlier. The 7-model ensemble mean (point 2) notably reduces the difference in the standard deviation while maintaining the performance in correlation and CRMS. The individual standard deviation values indicate that six out of seven selected models exceed the observed variability. The CCSM4 model (point f) fits best the observations (point a) with a slightly higher standard deviation and a correlation above 0.6.

The annual cycle of observed precipitation and all model simulations are plotted in Figure 10. Figure 10(a) presents the inter-model range of the 25 selected GCMs and the 25-model ensemble mean. The 25-model ensemble underestimates observations during much of the year, with exception of the austral winter. Penalba and Rivera (2016) found a similar result in the eastern portion of southern South America, reporting that CMIP5 models reproduce the shape of the annual cycle of precipitation but underestimate monthly totals during all year. A similar finding was reported by Maenza *et al.* (2017) in the western Pampas of Argentina. It is also evident in Figure 10(b) that most individual members underestimate the observed annual cycle in northeastern Argentina. The dry bias and the high dispersion are largely reduced by the 7-model ensemble (see Figures 10(c) and (d)). Figure 10(c) indicates that the 7-model ensemble mean properly reproduces the mean annual cycle of precipitation in the study region. The relative precipitation maximum in March, which is a particular feature of the region, is not simulated by the 7-model ensemble but by some individuals models (see Figure 10(b)).

3.4. Spatial analysis of precipitation

The observed and 7-model ensemble mean precipitation fields are presented in Figures 11(a) and (b) along with the two best performing models (CanESM2 and CCSM4) in Figures 11(c) and (d). Observations show a W-E gradient and slight increase towards the northeast (Figure 11(a)). The 7-model ensemble, as well as two of the best performing GCMs, tend to show an SW−NE gradient, mostly due to a negative bias towards the SW. The 7-model ensemble mean (Figure 11(b)) underestimates the regional precipitation field and reduces the spatial variability. The CanESM2 model (Figure 11(c)) achieves the best individual representation of the observed field as it reproduces the observed spatial variability, while the CCSM4 model smooths the gradient with positive bias in the whole region (Figure 11(d)).

The spatial skill scores of the annual simulated precipitation fields for the seven best performing models are

Table 4. Statistical evaluation metrics between the areal-averaged time series of monthly precipitation observed and simulated by historical long-term simulations (1901–2005). The shaded rows correspond to the models that achieve the best evaluation metrics, performing the 7-model ensemble mean.

|	Model	MBE (mm month^{-1})	MAE (mm month^{-1})	RMSE (mm month^{-1})	NSE	Correlation coefficient
1	ACCESS 1.0	3.5	34.4	45.3	−0.03	0.47
2	BCC-CSM1.1	−10.9	32.3	42.3	0.10	0.47
3	BNU-ESM	−21.1	35.9	46.5	−0.09	0.44
4	CanESM2	7.5	36.0	47.8	−0.15	0.60
5	CESM1-BGC	14.6	34.8	45.3	−0.03	0.60
6	CESM1-FASTCHEM	14.8	34.0	45.4	−0.04	0.62
7	CNRM-CM5.2	−19.5	32.6	43.5	0.05	0.54
8	CSIRO-Mk3.6.0	−22.3	37.7	48.2	−0.17	0.56
9	GFDL-CM3	−12.2	34.0	44.0	0.03	0.52
10	GFDL-ESM2M	−19.8	36.5	47.6	−0.14	0.47
11	GISS-E2-R	−29.5	38.3	51.0	−0.31	0.43
12	HadGEM2-CC	11.9	37.4	47.2	−0.12	0.45
13	HadGEM2-ES	9.0	36.0	46.1	−0.07	0.46
14	INM-CM4	−20.6	32.2	42.0	0.11	0.62
15	MIROC5	−27.7	36.6	47.8	−0.15	0.52
16	MIROC-ESM-CHEM	−29.4	36.0	47.5	−0.13	0.56
17	MPI-ESM-LR	−13.6	35.0	45.9	−0.06	0.37
18	MPI-ESM-MR	−15.9	36.9	48.5	−0.18	0.31
19	MRI-CGCM3	−5.3	35.8	46.5	−0.09	0.40
20	NorESM1-M	5.1	33.4	43.6	0.04	0.64
21	CCSM4	14.5	33.9	44.1	0.02	0.62
22	CMCC-CM	8.6	35.7	47.8	−0.15	0.48
23	EC-EARTH	−2.7	33.0	42.5	0.09	0.46
24	FGOALS-g2	−23.8	35.7	47.2	−0.13	0.41
25	MIROC4h	−22.7	35.3	46.1	0.08	0.62
	25-Model ensemble	−9.0	24.3	32.5	0.47	0.72
	7-Model ensemble	−0.4	24.9	32.3	0.48	0.71
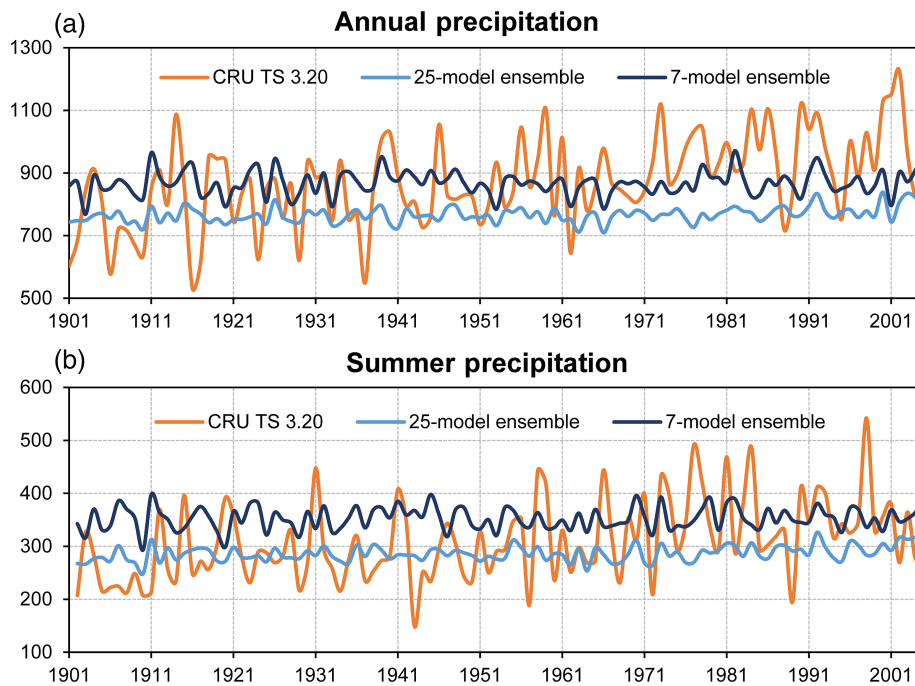
Figure 7. Areal-averaged time series of annual and summer precipitation (mm year^{-1} or mm season^{-1}) observed and simulated by the multi-model ensembles of the historical long-term simulations.
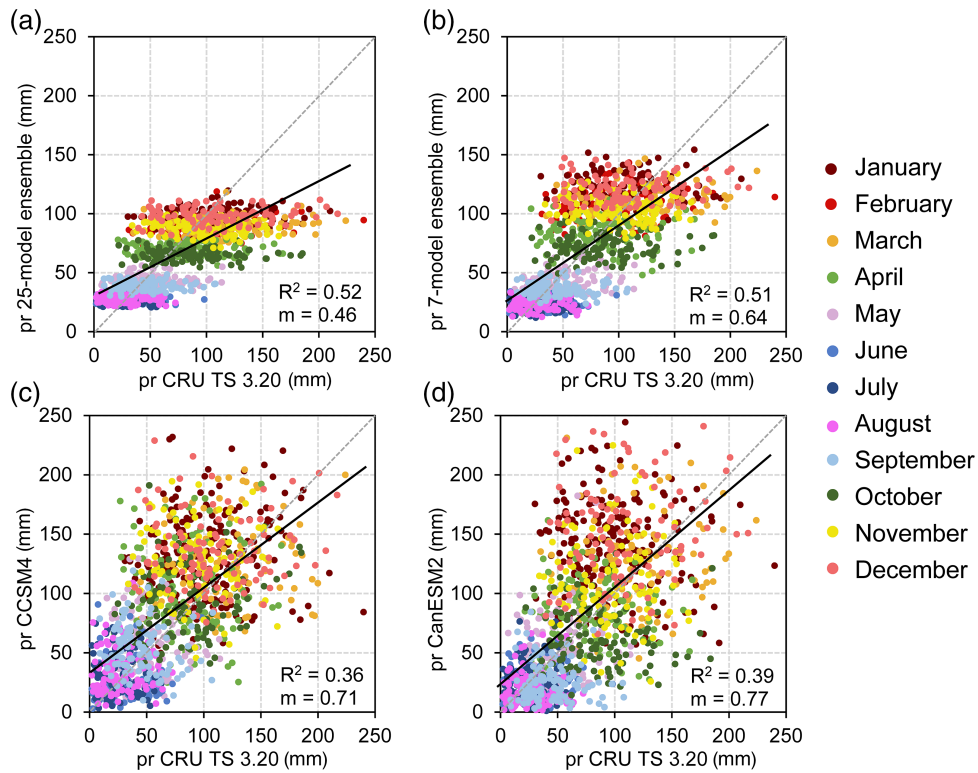
Figure 8. Scatterplots of simulated *versus* observed areal-averaged monthly precipitation (period 1901−2005) of the multi-model ensembles (a, b) and models with the best evaluation metrics (CCSM4 in c, CanESM2 in d). The linear regression line is plotted and its slope m is presented together with the determination coefficient R^2 in the bottom-right corner of each scatterplot.
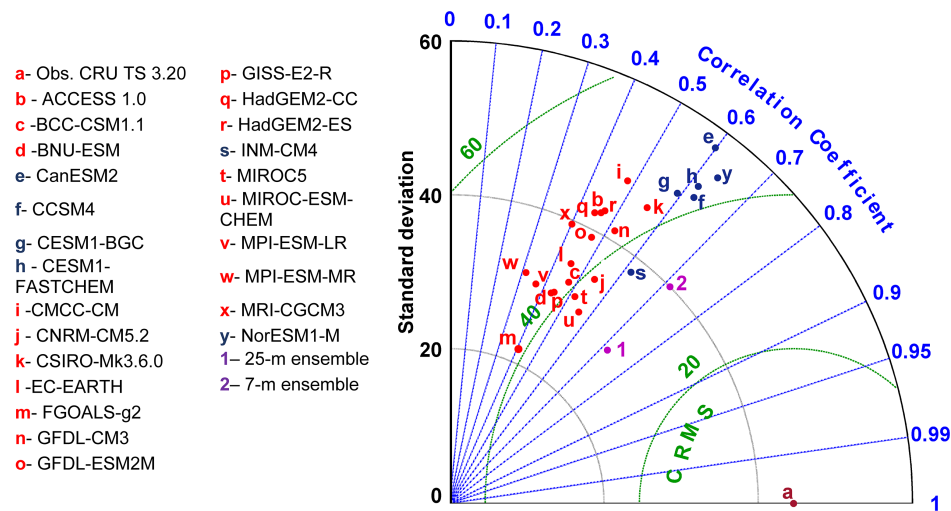

Figure 9. Taylor diagram of the regional precipitation simulated by historical long-term simulations (period 1901−2005). The blue dots correspond to the models that achieve the best evaluation metrics and perform the 7-model ensemble. CRMS: centred root mean square error.

presented in Table 5. Most of the GCMs present correlations in the range 0.6−0.7. The exception is the MIROC4h that exhibits the best correlation ($r = 0.8$) probably due to its high resolution ($0.56° \times 0.56°$) despite underestimating the annual precipitation by about 30%. The 7-model ensemble reduces the statistical errors of each individual GCM but it does not improve their spatial correlations. The CanESM2 model exhibits high correlation and an MBE of about −50 mm year^{-1}, which represents a 5%

of the areal mean precipitation. This result is consistent with its ability to simulate the annual mean field of precipitation discussed above. The CCSM4 model overestimates the areal-averaged mean precipitation in about 150 mm year^{-1}, which represents the 15% of the observed field. These results suggest that the multi-model ensemble does not improve the ability of individual models to simulate spatial precipitation fields. The best-performance individual models can recognize spatial features that do not
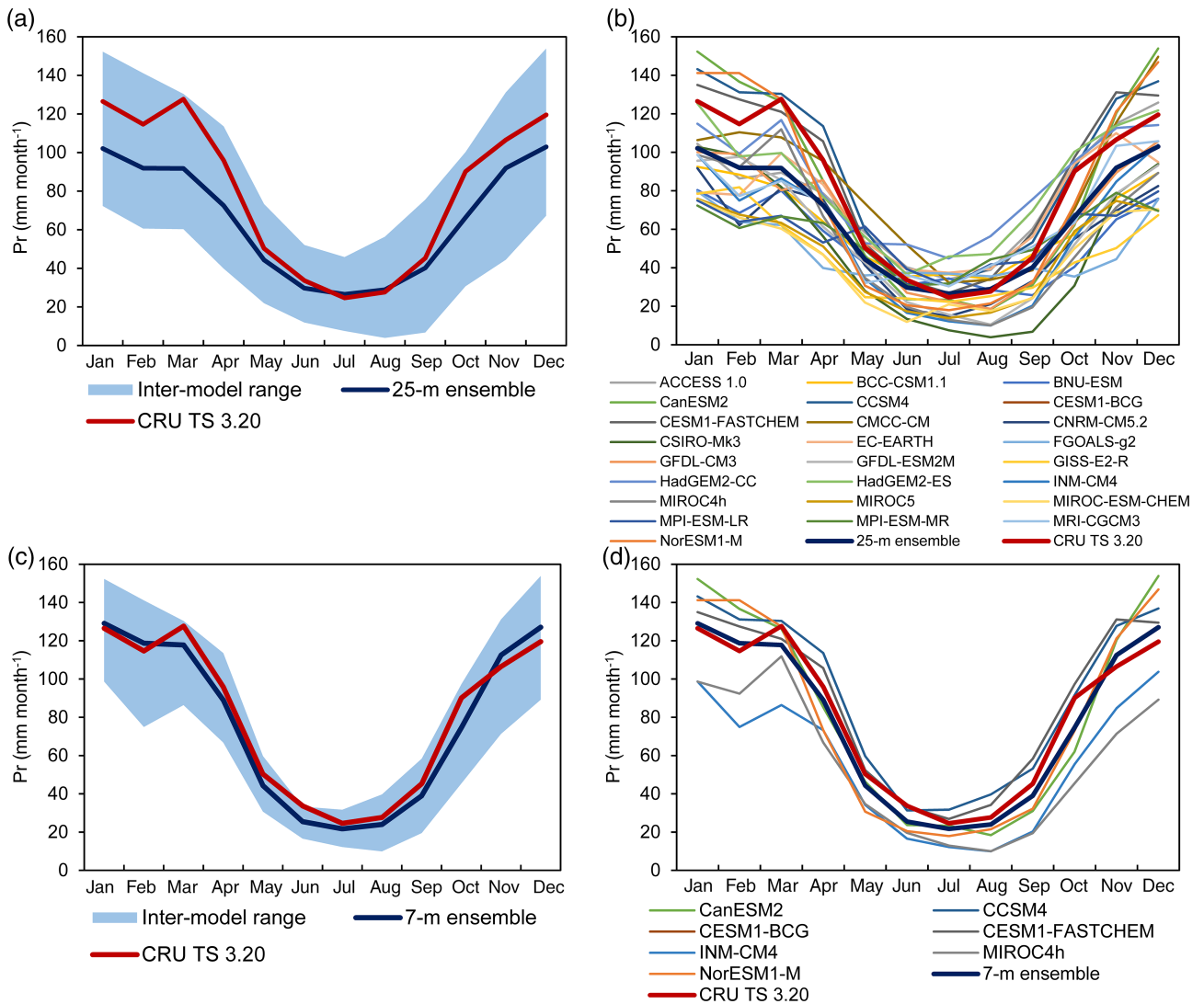
Figure 10. Annual cycle of the regional precipitation observed and simulated by historical long-term simulations of individual GCMs and the multi-model ensembles in the period 1971–2000: (a) and (b) the 25 analysed GCMs; (c) and (d) the 7 GCMs with best evaluation metrics.

simulate the multi-model ensemble, mainly the high spatial variability of the precipitation field.

4. Decadal simulations

Table 6 presents the GCMs with decadal simulations assessed in this study, including their evaluation metrics. The evaluation focuses on those GCMs that show greater ability to simulate the regional climate in the historical long-term experiment. Only five of the subset of best models present decadal simulations (indicated with 'a' in Table 1). Table 6 shows that the evaluation metrics of the decadal simulations do not show great differences from those found in the evaluation of long-term simulations (Tables 2 and 4).

In the case of the precipitation, the NSE shows a slight improvement with mostly positive values between 0.1 and 0.25; although the statistical errors and the correlations remain in the range of the historical long-term

simulations. Other studies have shown that the decadal simulations have been an advance in the modelling and predictability of the large-scale SST patterns that force the region's climate (e.g. Gonzalez and Goddard, 2016; Meehl *et al.,* 2016). While it is well known that the study region is strongly forced by large-scale SST patterns (Seager *et al.,* 2010; Barreiro *et al.,* 2014), our results suggest that the performance of decadal simulations is similar to that of the historical long-term simulations.

In the case of temperature, the evaluation metrics present values very similar to those of the historical long-term simulations as in precipitation. However, noted that the maximum errors of decadal simulations (MAE = 2.99 °C and RMSE = 3.43 °C; Table 6) are lower than the maximum errors of the historical long-term simulations (MAE = 4.08 °C and RMSE = 4.67 °C; Table 2). Thus, the magnitude of the errors is slightly reduced in decadal simulations.
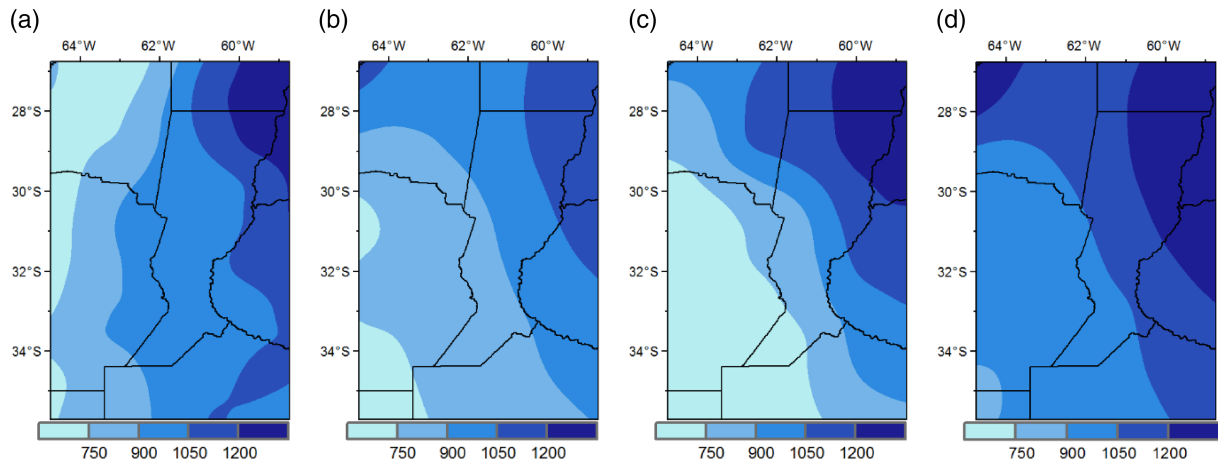
(a) (b) (c) (d)

Figure 11. Spatial distribution of the annual precipitation (in mm) observed and simulated by historical long-term simulations in the period 1971–2000 for (a) CRU TS 3.20, (b) 7-model ensemble mean, (c) CanESM2 and (d) CCSM4.

Table 5. Statistical evaluation metrics between the mean fields of annual precipitation observed and simulated by historical long-term simulations (analysed in the normal period 1971–2000).

Model	MBE	RMSE	Spatial correlation
CanESM2	−49.5	178.8	0.70
CCSM4	152.0	211.3	0.60
CESM1-BGC	134.3	190.8	0.66
CESM1-FASTCHEM	160.0	221.2	0.58
INM-CM4	−277.6	311.5	0.68
MIROC4h	−302.8	325.2	0.80
NorESM1-M	−11.6	200.7	0.24
Ensemble	−27.9	137.8	0.67

Table 6. Statistical evaluation metrics between the areal-averaged time series of monthly mean temperature and precipitation observed and simulated by decadal simulations (1961–2010).

Model	MBE	MAE	RMSE	NSE	Correlation coefficient
Temperature					
CanCM4	−1.69	2.22	2.88	0.63	0.94
CCSM4	0.76	1.53	1.90	0.84	0.94
CFSv2-2011	0.17	1.48	1.87	0.85	0.93
CMCC-CM	0.17	1.48	1.87	0.85	0.92
EC-EARTH	0.23	1.54	1.99	0.83	0.92
FGOALS-g2	−2.61	2.73	3.19	0.55	0.93
MIROC4h	2.73	2.99	3.43	0.48	0.95
Ensemble mean	−0.13	1.13	1.46	0.91	0.96
Precipitation					
CanCM4	4.23	34.52	44.97	0.16	0.65
CCSM4	11.32	33.10	43.48	0.22	0.66
CFSv2-2011	−3.44	34.23	45.46	0.11	0.48
CMCC-CM	6.97	37.16	48.25	0.03	0.54
EC-EARTH	−8.53	34.28	45.29	0.15	0.48
FGOALS-g2	−29.55	38.95	51.74	−0.11	0.51
MIROC4h	−19.42	34.49	44.94	0.16	0.63
Ensemble mean	−5.52	25.57	33.80	0.53	0.74

The units of the mean bias error (MBE), mean absolute error (MAE) and root mean square error (RMSE) are $°C\,month^{-1}$ for temperature and $mm\,month^{-1}$ for precipitation. The Nash–Sutcliffe efficiency (NSE) is dimensionless.

As in the historical long-term simulations, the multi-model ensemble improves the evaluation metrics of precipitation and temperature, over those of the individual GCMs (see Table 6). According to the Table 6, the CCSM4 model again shows the greater ability to simulate the regional mean temperature and precipitation. The CFSv2-2011 and the CMCC-CM models also exhibit good performance for temperature, while the CanCM4 model is among the best for precipitation.

The Taylor diagram of mean temperature (Figure 12(a)) is similar to that obtained for the long-term simulations (Figure 4): all GCMs simulate temperature with standard deviations greater than the reference value, correlation coefficients in the range 0.92–0.95 and CRMS close to 2 °C. The multi-model ensemble (point I) approaches the reference point as it improves both the correlation and CRMS values. Different standard deviations are found in the Taylor diagram of the precipitation (Figure 12(b)); certain models considerably underestimate the reference value of 49 mm (point A) while others overestimate it. The models with the highest standard deviation, also exhibit the highest correlation coefficients (points B and C, r ~ 0.65). The multi-model ensemble (point I) improves the correlation and the CRMS of each GCM, although

it widely underestimates the standard deviation of the reference value.

The mean annual cycle of temperature and precipitation from observations and simulations are shown in Figure 13. Figure 13(a) shows that temperature of the multi-model ensemble fits very well the observed annual cycle and is consistent with a low MBE of only −0.1 °C (see Table 6). Most of the GCMs (Figure 13(b)) reproduce the annual cycle but with systematic errors that are quantified by the MBE values in Table 6. Five out of the seven analysed models present positive MBEs less than 1 °C. Only two models have negative errors, but greater than 1.6 °C
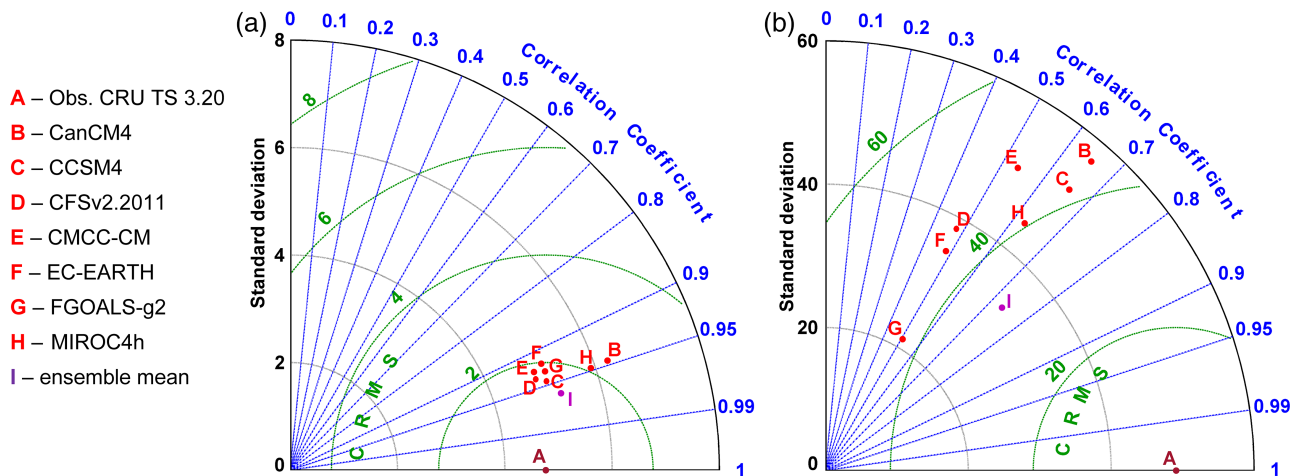
Figure 12. Taylor diagrams for (a) temperature and (b) precipitation simulated by decadal simulations (1961–2010). CRMS: centred root mean square error.

in magnitude. The multi-model ensemble precipitation (Figure 13(c)) underestimates observed precipitation during the rainy season (between September and April) while slightly overestimates the precipitation of the dry season. The annual cycle of precipitation is simulated by each GCM with dissimilar success; while some GCMs adequately reproduce the annual cycle (e.g. CCSM4 and CanCM4) others are far removed from the observed distribution (e.g. FGOALS-g2, CMCC-CM).

5. Conclusions

This paper evaluated the ability of CMIP5 GCMs to simulate the observed spatio-temporal behaviour of the climate in northeastern Argentina. The GCMs were selected from the two main core sets of CMIP5 simulations: long-term (1901–2005) and decadal (1961–2010). Monthly simulations of precipitation and temperature were evaluated with the observed CRU TS 3.20 data set. The simulations of each GCM and the multi-model ensembles were evaluated through the MBE, the MAE, the RMSE, the coefficient of NSE, and the Pearson correlation. These metrics were computed for both areal-averaged time series and spatial fields. We also determined the ability of GCMs to simulate the annual mean fields and the mean annual cycle of temperature and precipitation for the normal 30-year period 1971–2000. Inspection of the results allowed selection and evaluation of model subsets that represent better the regional climate.

The temperature observations show that the region: (1) experienced a warming trend in annual mean temperature, with a strong signal of positive change during austral spring; (2) presents a marked annual cycle of temperature with mean values of 25 °C in summer and 12 °C in winter; and (3) has a spatial south–north gradient that ranges from 15 to 21 °C. These three features are correctly simulated by all GCMs that performed historical long-term simulations. All models reach correlations in the range 0.92–0.94 for the 1901–2005 time series. Among them,

the CCSM4 and CESM1-BGC models have the highest ability to simulate regional temperature as they achieve the best evaluation metrics and reproduce properly the features of the climatology. Furthermore, the ensemble of the nine most skilful models improves the performance of the individual GCMs. The 9-model ensemble reaches a correlation of 0.96, reduces more than 20% the positive systematic errors shown by the all-model ensemble in the 1901–2005 time series, and fits well the climatological annual cycle.

The precipitation observations show that the region: (1) experienced a positive trend in annual and summer precipitation after the 1960s (Barros *et al.,* 2008; Lovino *et al.,* 2014); (2) has a marked intra-annual variability, with the warm season (October to April) accounting for more than 80% of the annual precipitation; and (3) presents a spatial west–east gradient. The set of CMIP5 historical long-term simulations reproduce the regional precipitation features with diverse degrees of success. The ensemble performed with the seven more skilful (in terms of precipitation) models reveals better skill than the individual GCMs, reducing all-model biases in more than 25%. On the other hand, the 7-model ensemble accurately reproduces the annual precipitation cycle improving the performance of all the rest of the GCMs that underestimate rainfall in the rainy season. However, the 7-model ensemble does not clearly identify the positive trend in annual and summer precipitation. Spatially, the best individual models and the 7-model ensemble mean correctly reproduce the large scale features of the observed annual precipitation pattern but fail to replicate the smaller scales spatial variability.

The seven models with better precipitation skill had already been identified as those that best resemble the South American monsoon with the most accurate quantitative estimations of the summer rainfall in the monsoon area (Hsu *et al.,* 2013). Furthermore, Jones and Carvalho (2013) showed that most of the selected GCMs ably reproduce the large-scale features like seasonal amplitude, onset or duration of the South America Monsoon System. Thus,
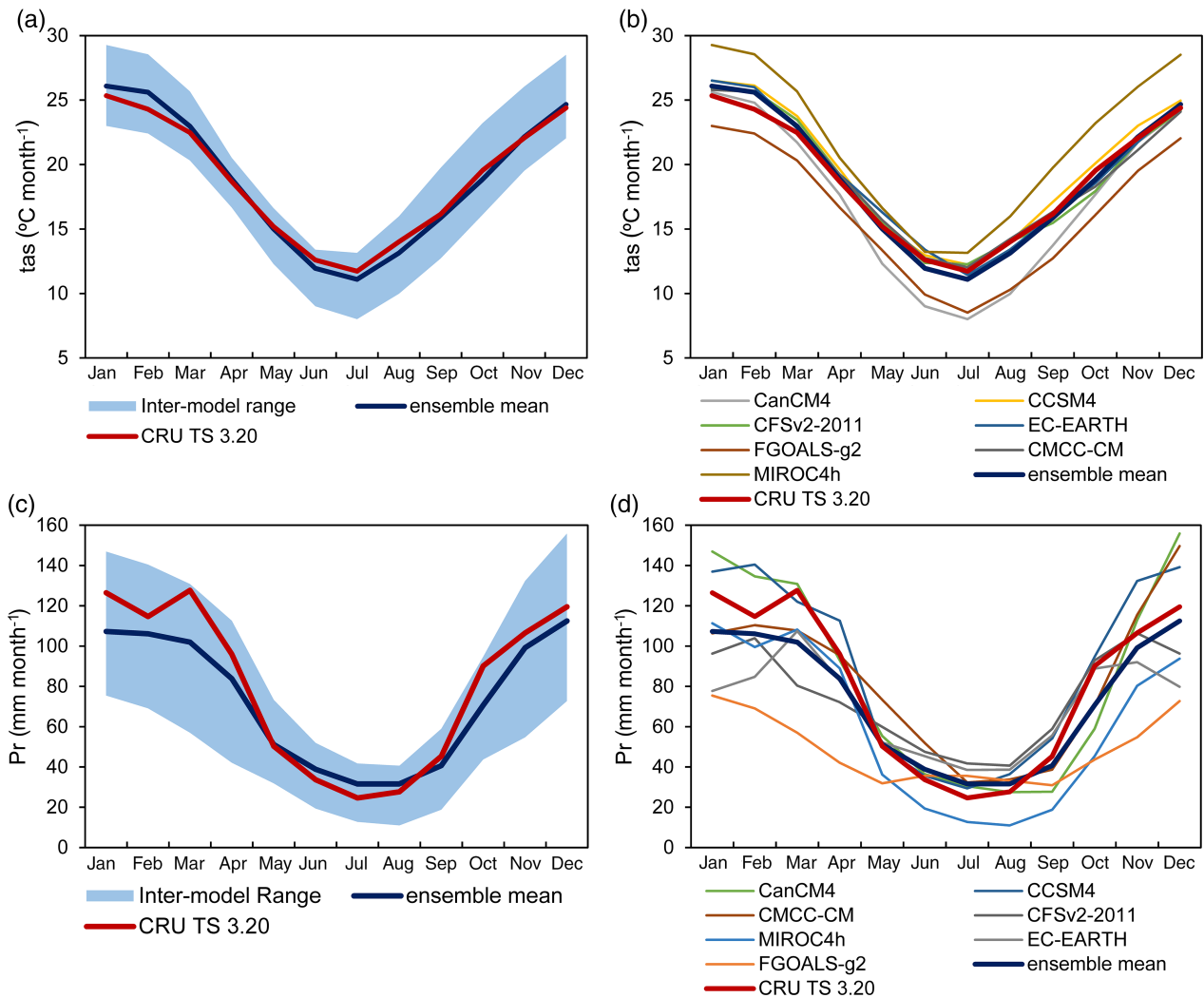
Figure 13. Annual cycle of temperature (a, b) and precipitation (c, d) observed and simulated by decadal simulations of individual GCMs and their multi-model ensembles in the period 1971–2000.

the ability of the selected GCMs to reproduce the monsoon features may explain their high performance to simulate the precipitation in northeastern Argentina. Four of these seven selected models for precipitation also achieve the best evaluation metrics for temperature: CCSM4, CESM1-BGC, CESM1-FASTCHEM, and NorESM1-M.

It is widely known that the decadal variability, strongly forced by large-scale SST patterns, plays a key role in modulating the climate of northeastern Argentina (e.g. Barreiro *et al.*, 2014 and references therein). Therefore, it would be expected that the inclusion of this external forcing in the decadal simulations would lead to improvements in skill. Nevertheless, our results suggest that decadal simulations do not improve the evaluation metrics achieved by the long-term simulations, suggesting that the initialization of decadal simulations with climate observations does not lead necessarily to a better representation of regional climate. Still, some individual models do have success. The CCSM4 model achieves the best evaluation metrics for both temperature and precipitation, demonstrating the greater ability to simulate the regional climate on decadal

time scales. This model also presents one of the best performances in the evaluation of long-term simulations.

Reliable projections of future climate require models that adequately represent the characteristics of the regional climate system. The evaluation of the long-term simulations allowed us to select a subset of 9 GCMs that best reproduce the mean temperature fields and a subset of 7 GCMs that most skilfully represents regional precipitation. The multi-model ensemble means computed using the selected GCMs achieve the largest skill to represent historical and present climate as they reached the best evaluation metrics, exceeding the performance of each GCM and the multi-model ensemble performed with all the GCMs. It is expected that these models with a greater capacity of representing the regional historical climate, will be more appropriate for simulating the future climate.

Acknowledgements

References

Agrofy News. 2017. *Special Reports of the Rosario Stock Exchange, 2014–2015 Campaign: Maize and Soybeans* (in Spanish). http://news.agrofy.com.ar/especiales/soja14-15/siembra and http://news.agrofy.com.ar/especiales/maiz14-15/siembra-maiz (accessed 1 June 2017).

Arora VK, Scinocca JF, Boer GJ, Christian JR, Denman KL, Flato GM, Kharin VV, Lee WG, Merryfield WJ. 2011. Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. *Geophys. Res. Lett.* **38**: L05805. https://doi.org/10.1029/2010GL046270.

Baethgen WE, Goddard L. 2013. Latin American perspectives on adaptation of agricultural systems to climate variability and change. In *Handbook of Climate Change and Agroecosystems: Global and Regional Aspects and Implications*. ICP Series on Climate Change Impacts, Adaptation, and Mitigation, Vol. **2**, Hillel D, Rosenzweig C (eds). Imperial College Press: London, UK, 57–72.

Barreiro M, Díaz N, Renom M. 2014. Role of the global oceans and land–atmosphere interaction on summertime interdecadal variability over northern Argentina. *Clim. Dyn.* **42**(7–8): 1733–1753. https://doi.org/10.1007/s00382-014-2088-6.

de Barros Soares D, Lee H, Loikith PC, Barkhordarian A, Mechoso CR. 2017. Can significant trends be detected in surface air temperature and precipitation over South America in recent decades? *Int. J. Climatol.* **37**: 1483–1493. https://doi.org/10.1002/joc.4792.

Barros V, Doyle M, Camilloni I. 2008. Precipitation trends in southeastern South America: relationship with ENSO phases and with low-level circulation. *Theor. Appl. Climatol.* **93**(1–2): 19–33. https://doi.org/10.1007/s00704-007-0329-x.

Bentsen M, Bethke I, Debernard J, Iversen T, Kirkevåg A, Seland Ø, Drange H, Roelandt C, Seierstad I, Hoose CKristjánsson J. 2012. The Norwegian earth system model, NorESM1-M – Part 1: Description and basic evaluation. *Geosci. Model Dev. Discus.* **5**(3): 2843–2931. https://doi.org/10.5194/gmdd-5-2843-2012.

Berbery E, Barros V. 2002. The hydrologic cycle of the La Plata basin in South America. *J. Hydrometeorol.* **3**: 630–645. https://doi.org/10.1175/1525-7541(2002)003<0630:THCOTL>2.0.CO;2.

Bi D, Dix M, Marsland SJ, O'Farrell S, Rashid HA, Uotila P, Hirst AC, Kowalczyk E, Golebiewski M, Sullivan A, Yan H, Hannah N, Franklin C, Sun Z, Vohralik P, Watterson I, Zhou X, Fiedler R, Collier M, Ma Y, Noonan J, Stevens L, Uhe P, Zhu H, Griffies SM, Hill R, Harris C, Puri K. 2013. The ACCESS coupled model: description, control climate and evaluation. *Austr. Meteorol. Oceanogr. J.* **63**: 41–64. https://doi.org/10.22499/2.6301.004.

Boulanger J, Carril A, Sanchez E. 2016. CLARIS-La Plata Basin: regional hydroclimate variability, uncertainties and climate scenarios. *Clim. Res.* **68**: 93–94. https://doi.org/10.3354/cr01392.

Caffera RM, Berbery EH. 2006. La Plata basin climatology. In *Climatic Change in the La Plata Basin*, Barros V, Clarke R, Diaz P (eds). National Scientific and Technical Research Council (CONICET): Argentina, 16–32.

Carril AF, Cavalcanti IFA, Menéndez CG, Sörensson A, López-Franca N, Rivera J, Robledo F, Zaninelli P, Ambrizzi T, Penalba O, da Rocha R, Sánchez E, Bettolli M, Pessacg N, Renom M, Ruscica R, Solman S, Tencer B, Grimm A, Rusticucci M, Cherchi A, Tedeschi R, Zamboni L. 2016. Extreme events in the La Plata basin: a retrospective analysis of what we have learned during CLARIS-LPB project. *Clim. Res.* **68**: 95–116. https://doi.org/10.3354/cr01374.

Cavalcanti IFA, Carril AF, Penalba OC, Grimm AM, Menéndez CG, Sanchez E, Chechi A, Sörensson A, Robledo F, Rivera J, Pántano V, Betolli LM, Zaninelli P, Zamboni L, Tedeschi RG, Dominguez M, Ruscica R, Flach R. 2015. Precipitation extremes over La Plata Basin – review and new results from observations and climate simulations. *J. Hydrol.* **523**: 211–230–211–230. https://doi.org/10.1016/j.jhydrol.2015.01.028.

Collins WJ, Bellouin N, Doutriaux-Boucher M, Gedney N, Halloran P, Hinton T, Hughes J, Jones CD, Joshi M, Liddicoat S, Martin G, O'Connor F, Rae J, Senior C, Sitch S, Totterdell I, Wiltshire A, Woodward S. 2011. Development and evaluation of an earth-system model – HadGEM2. *Geosci. Model Dev.* **4**: 1051, 2011–1075. https://doi.org/10.5194/gmd-4-1051-2011.

Déqué M. 2012. Continuous variables. In *Forecast Verification – A Practitioner's Guide in Atmospheric Science*, 2nd edn, Jolliffe IT, Stephenson DV (eds). John Wiley and Sons: Chichester, UK, 97–120.

Díaz LB, Vera CS. 2017. Austral summer precipitation interannual variability and trends over southeastern South America in CMIP5 models. *Int. J. Climatol.* **37**: 681–695. https://doi.org/10.1002/joc.5031.

Donner L, Wyman B, Hemler R, Horowitz L, Ming Y, Zhao M, Golaz J, Ginoux P, Lin S, Schwarzkopf M, Austin J, Alaka G, Cooke W, Delworth T, Freidenreich S, Gordon C, Griffies S, Held I, Hurlin W, Klein S, Knutson T, Langenhorst A, Lee H, Lin Y, Magi B, Malyshev S, Milly P, Naik V, Nath M, Pincus R, Ploshay J, Ramaswamy V, Seman C, Shevliakova E, Sirutis J, Stern W, Stouffer R, Wilson R, Winton M, Wittenberg A, Zeng F. 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**(13): 3484–3519. https://doi.org/10.1175/2011jcli3955.1.

Dunne J, John J, Adcroft A, Griffies S, Hallberg R, Shevliakova E, Stouffer R, Cooke W, Dunne K, Harrison M, Krasting J, Malyshev S, Milly P, Phillips P, Sentman L, Samuels B, Spelman M, Winton M, Wittenberg A, Zadeh N. 2012. GFDL's ESM2 global coupled climate–carbon earth system models. Part I: physical formulation and baseline simulation characteristics. *J. Clim.* **25**(19): 6646–6665. https://doi.org/10.1175/jcli-d-11-00560.1.

Fischer RA, Byerlee D, Edmeades GO. 2014. *Crop Yields and Global Food Security: Will Yield Increase Continue to Feed the World? ACIAR Monograph No. 158*. Australian Centre for International Agricultural Research: Canberra, Australia.

Flato G, Marotzke J, Abiodun B, Braconnot P, Chou SC, Collins W, Cox P, Driouech F, Emori S, Eyring V, Forest C, Gleckler P, Guilyardi E, Jakob C, Kattsov V, Reason C, Rummukainen M. 2013. Evaluation of climate models. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds). Cambridge University Press: Cambridge, UK/New York, NY.

Gallardo C, Gil V, Tejeda C, Sánchez E, Gaertner MA. 2016. Köppen-Trewartha classification used to assess climate changes simulated by a regional climate model ensemble over South America. *Clim. Res.* **68**: 137–149. https://doi.org/10.3354/cr01340.

Gent PR, Danabasoglu G, Donner LJ, Holland MM, Hunke EC, Jayne SR, Lawrence DM, Neale RB, Rasch PJ, Vertenstein M, Worley PH, Yang Z, Zhang M. 2011. The Community Climate System Model Version 4. *J. Clim.* **24**: 4973–4991. https://doi.org/10.1175/2011JCLI4083.1.

Gonzalez PLM, Goddard L. 2016. Long-lead ENSO predictability from CMIP5 decadal hindcasts. *Clim. Dyn.* **46**: 3127–3147. https://doi.org/10.1007/s00382-015-2757-0.

Grimm A, Saboia J. 2015. Interdecadal variability of the south American precipitation in the monsoon season. *J. Clim.* **28**(2): 755–775. https://doi.org/10.1175/jcli-d-14-00046.1.

Gulizia C, Camilloni I. 2015. Comparative analysis of the ability of a set of CMIP3 and CMIP5 global climate models to represent precipitation in South America. *Int. J. Climatol.* **35**: 583–595. https://doi.org/10.1002/joc.4005.

Harris I, Jones PD, Osborn TJ, Lister DH. 2014. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 dataset. *Int. J. Climatol.* **34**: 623–642. https://doi.org/10.1002/joc.3711.

Hazeleger W, Severijns C, Semmler T, Ştefănescu S, Yang S, Wang X, Wyser K, Dutra E, Baldasano J, Bintanja R, Bougeault P, Caballero R, Ekman A, Christensen J, van den Hurk B, Jimenez P, Jones C, Kållberg P, Koenigk T, McGrath R, Miranda P, Van Noije T, Palmer T, Parodi J, Schmith T, Selten F, Storelvmo T, Sterl A, Tapamo H, Vancoppenolle M, Viterbo PWillén U. 2010. EC-earth: a seamless earth-system prediction approach in action. *Bull. Am. Meteorol. Soc.* **91**(10): 1357–1363. https://doi.org/10.1175/2010bams2877.1.

Hsu P, Li T, Murakami H, Kitoh A. 2013. Future change of the global monsoon revealed from 19 CMIP5 models. *J. Geophys. Res. Atmos.* **118**(3): 1247–1260. https://doi.org/10.1002/jgrd.50145.

Jacob D, Elizalde A, Haensler A, Hagemann S, Kumar P, Podzun R, Rechid D, Remedio AR, Saeed F, Sieck K, Teichmann C, Wilhelm

C. 2012. Assessing the transferability of the regional climate model REMO to different COordinated regional climate downscaling EXperiment (CORDEX) regions. *Atmosphere* **3**(4): 181–199. https://doi.org/10.3390/atmos3010181.

Ji D, Wang L, Feng J, Wu Q, Cheng H, Zhang Q, Yang J, Dong W, Dai Y, Gong D, Zhang R, Wang X, Liu J, Moore J, Chen D, Zhou M. 2014. Description and basic evaluation of Beijing Normal University earth system model (BNU-ESM) version 1. *Geosci. Model Dev.* **7**(5): 2039–2064. https://doi.org/10.5194/gmd-7-2039-2014.

Jones C, Carvalho LMV. 2013. Climate change in the south American monsoon system: present climate and CMIP5 projections. *J. Clim.* **26**(17): 6660–6678. https://doi.org/10.1175/jcli-d-12-00412.1.

Kumar D, Kodra E, Ganguly A. 2014. Regional and seasonal intercomparison of CMIP3 and CMIP5 climate model ensembles for temperature and precipitation. *Clim. Dyn.* **43**(9–10): 2491–2518. https://doi.org/10.1007/s00382-014-2070-3.

Lamarque J. 2013. *CESM1-FASTCHEM Model Output Prepared for CMIP5 Historical, Served by ESGF*. WDCC at DKRZ: Hamburg, Germany. https://doi.org/10.1594/WDCC/CMIP5.NFCFhi.

Li L, Lin P, Yu Y, Wang B, Zhou T, Liu L, Liu J, Bao Q, Xu S, Huang W, Xia K, Pu Y, Dong L, Shen S, Liu Y, Hu N, Liu M, Sun W, Shi X, Zheng W, Wu B, Song M, Liu H, Zhang X, Wu G, Xue W, Huang X, Yang G, Song ZQiao F. 2013. The flexible global ocean-atmosphere-land system model, grid-point version 2: FGOALS-g2. *Adv. Atmos. Sci.* **30**(3): 543–560. https://doi.org/10.1007/s00376-012-2140-6.

Long M, Lindsay K, Peacock S, Moore JK, Doney S. 2013. Twentieth-century oceanic carbon uptake and storage in CESM1(BGC). *J. Clim.* **26**(18): 6775–6800. https://doi.org/10.1175/jcli-d-12-00184.1.

Lovino MA. 2015. *Environmental Impact of Climate Variability and Extreme Events in the Province of Santa Fe (Argentina), Within the Climate Change Context* (in Spanish). PhD thesis, Universidad Nacional del Litoral, Argentina, 238 pp. http://bibliotecavirtual.unl.edu.ar:8080/tesis/handle/11185/707 (accessed 1 June 2017)

Lovino MA, García NO, Baethgen WE. 2014. Spatiotemporal analysis of extreme precipitation events in the northeast region of Argentina (NEA*). J. Hydrol. Reg. Stud.* **2**: 140–158. https://doi.org/10.1016/j.ejrh.2014.09.001.

Maenza RA, Agosta EA, Bettolli ML. 2017. Climate change and precipitation variability over the western 'pampas' in Argentina. *Int. J. Climatol.* **37**: 445–463. https://doi.org/10.1002/joc.5014.

Magrin GO, Marengo JA, Boulanger J-P, Buckeridge MS, Castellanos E, Poveda G, Scarano FR, Vicuña S. 2014. Central and South America. In *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects*. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Barros VR, Field CB, Dokken DJ, Mastrandrea MD, Mach KJ, Bilir TE, Chatterjee M, Ebi KL, Estrada YO, Genova RC, Girma B, Kissel ES, Levy AN, MacCracken S, Mastrandrea PR, White LL (eds). Cambridge University Press: Cambridge, UK/New York, NY, 1499–1566.

Matsuura K, Willmott CJ. 2009. *Terrestrial Air Temperature: 1900–2008 Gridded Monthly Time Series (version 2.01)*. Center for Climatic Research, Department of Geography, University of Delaware: Newark http://climate.geog.udel.edu/*climate/ (accessed 28 October 2017).

Meehl G, Covey C, Taylor K, Delworth T, Stouffer R, Latif M, McAvaney B, Mitchell J. 2007. The WCRP CMIP3 multimodel dataset: a new era in climate change research. *Bull. Am. Meteorol. Soc.* **88**(9): 1383–1394. https://doi.org/10.1175/bams-88-9-1383.

Meehl G, Hu A, Teng H. 2016. Initialized decadal prediction for transition to positive phase of the Interdecadal Pacific Oscillation. *Nat. Commun.* **7**: 11718. https://doi.org/10.1038/ncomms11718.

Merryfield WJ, Lee W, Boer G, Kharin V, Scinocca J, Flato G, Ajayamohan R, Fyfe J, Tang Y, Polavarapu S. 2013. The Canadian seasonal to interannual prediction system. Part I: models and initialization. *Mon. Weather Rev.* **141**(8): 2910–2945. https://doi.org/10.1175/mwr-d-12-00216.1.

Miller R, Schmidt G, Nazarenko L, Tausnev N, Bauer S, DelGenio A, Kelley M, Lo K, Ruedy R, Shindell D, Aleinov I, Bauer M, Bleck R, Canuto V, Chen Y, Cheng Y, Clune T, Faluvegi G, Hansen J, Healy R, Kiang N, Koch D, Lacis A, LeGrande A, Lerner J, Menon S, Oinas V, Pérez García-Pando C, Perlwitz J, Puma M, Rind D, Romanou A, Russell G, Sato M, Sun S, Tsigaridis K, Unger N, Voulgarakis A, Yao MZhang J. 2014. CMIP5 historical simulations (1850-2012) with GISS ModelE2. *J. Adv. Model. Earth Syst.* **6**(2): 441–478. https://doi.org/10.1002/2013ms000266.

Ministry of Agroindustry. 2017. *Open Data of Agroindustry* (in Spanish). https://datos.magyp.gob.ar/reportes.php?reporte=Estimaciones (accessed June 2017).

Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. Am. Soc. Agric. Biol. Eng.* **50**(3): 885–900. https://doi.org/10.13031/2013.23153.

Nash JE, Sutcliffe JV. 1970. River flow forecasting through conceptual models. Part I. A discussion of principles. *J. Hydrol.* **10**(3): 282–290. https://doi.org/10.1016/0022-1694(70)90255-6.

Penalba O, Rivera J. 2013. Future changes in drought characteristics over southern South America projected by a CMIP5 multi-model ensemble. *Am. J. Clim. Chang.* **2**: 173–182. https://doi.org/10.4236/ajcc.2013.23017.

Penalba O, Rivera J. 2016. Regional aspects of future precipitation and meteorological drought characteristics over southern South America projected by a CMIP5 multi-model ensemble. *Int. J. Climatol.* **36**(2): 974–986. https://doi.org/10.1002/joc.4398.

Rotstayn LD, Collier MA, Dix MR, Feng Y, Gordon H, O'Farrell S, Smith I, Syktus J. 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. https://doi.org/10.1002/joc.1952.

Saha S, Moorthi S, Pan H, Wu X, Wang J, Nadiga S, Tripp P, Kistler R, Woollen J, Behringer D, Liu H, Stokes D, Grumbine R, Gayno G, Wang J, Hou Y, Chuang H, Juang H, Sela J, Iredell M, Treadon R, Kleist D, Van Delst P, Keyser D, Derber J, Ek M, Meng J, Wei H, Yang R, Lord S, Van Den Dool H, Kumar A, Wang W, Long C, Chelliah M, Xue Y, Huang B, Schemm J, Ebisuzaki W, Lin R, Xie P, Chen M, Zhou S, Higgins W, Zou C, Liu Q, Chen Y, Han Y, Cucurull L, Reynolds R, Rutledge Goldberg M. 2010. The NCEP climate forecast system reanalysis. *Bull. Am. Meteorol. Soc.* **91**(8): 1015–1057. https://doi.org/10.1175/2010bams3001.1.

Sakamoto TT, Komuro Y, Nishimura T, Ishii M, Tatebe H, Shiogama H, Hasegawa A, Toyoda T, Mori M, Suzuki, Imada Y, Nozawa T, Takata K, Mochizuki T, Ogochi K, Emori S, Hasumi H, Kimoto M. 2012. MIROC4h – a new high-resolution atmosphere-ocean coupled general circulation model. *J. Meteorol. Soc. Jpn.* **90**: 325–359. https://doi.org/10.2151/jmsj.2012-301.

Sánchez E, Solman S, Remedio ARC, Berbery EH, Samuelsson P, Da Rocha RP, Mourão C, Li L, Marengo J, de Castro M, Jacob D. 2015. Regional climate modelling in CLARIS-LPB: a concerted approach towards twenty first century projections of regional temperature and precipitation over South America. *Clim. Dyn.* **45**: 2193–2212. https://doi.org/10.1007/s00382-014-2466-0.

Schneider U, Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Ziese M 2011. *GPCC Full Data Reanalysis Version 7.0 at 0.5°: Monthly Land-Surface Precipitation from Rain Gauges built on GTS-based and Historic Data*. Global Precipitation Climatology Centre, Deutscher Wetterdienst: Offenbach, Germany. https://doi.org/10.5676/DWD_GPCC/FD_M_V6_050.

Scoccimarro E, Gualdi S, Bellucci A, Sanna A, Giuseppe Fogli P, Manzini E, Vichi M, Oddo P, Navarra A. 2011. Effects of tropical cyclones on ocean heat transport in a high-resolution coupled general circulation model. *J. Clim.* **24**(16): 4368–4384. https://doi.org/10.1175/2011jcli4104.1.

Seager R, Naik N, Baethgen W, Robertson A, Kushnir Y, Nakamura J, Jurburg S. 2010. Tropical Oceanic Causes of Interannual to Multidecadal Precipitation Variability in Southeast South America over the Past Century. *J. Clim.* **23**: 5517–5539. https://doi.org/10.1175/2010JCLI3578.1.

Seneviratne SI, Nicholls N, Easterling D, Goodess CM, Kanae S, Kossin J, Luo Y, Marengo J, McInnes K, Rahimi M, Reichstein M, Sorteberg A, Vera C, Zhang X. 2012. Changes in climate extremes and their impacts on the natural physical environment. In *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, Field CB, Barros V, Stocker TF, Qin D, Dokken DJ, Ebi KL, Mastrandrea MD, Mach KJ, Plattner G, Allen SK, Tignor M, Midgley PM, A Special Report of Working Groups I and II of the IPCC, Field CB, Barros V, Stocker TF, Qin D, Dokken DJ, Ebi KL, Mastrandrea MD, Mach KJ, Plattner G-K, Allen SK, Tignor M, Midgley PM (eds). Cambridge University Press: Cambridge, UK/New York, NY, 109–230.

Sheffield J, Barrett AP, Colle B, Nelun Fernando DD, Fu R, Geil KL, Hu Q, Kinter J, Kumar S, Langenbrunner B, Lombardo K, Long LN, Maloney E, Mariotti A, Meyerson JE, Mo KC, David Neelin JJ, Nigam S, Pan Z, Ren T, Ruiz-Barradas A, Serra YL, Seth A, Thibeault JM, Stroeve JC, Yang Z, Yin L. 2013. North American climate in CMIP5

experiments. Part I. Evaluation of historical simulations of continental and regional climatology. *J. Clim.* **26**: 9209–9245. https://doi.org/10.1175/JCLI-D-12-00592.1.

Shepard D. 1968. A two-dimensional interpolation function for irregularly spaced data. In *ACM National Conference*. Association for Computing Machinery: New York, 517–523.

Solman SA, Sanchez E, Samuelsson P, da Rocha R, Li L, Marengo J, Pessacg N, Remedio A, Chou S, Berbery EH, Le Treut H, de Castro MJacob D. 2013. Evaluation of an ensemble of regional climate model simulations over South America driven by the ERA-interim reanalysis: model performance and uncertainties. *Clim. Dyn.* **41**: 1139–1157. https://doi.org/10.1007/s00382-013-1667-2.

Taylor K. 2001. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **106**(D7): 7183–7192. https://doi.org/10.1029/2000jd900719.

Taylor K, Stouffer RJ, Meehl G. 2012a. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* **93**: 485–498. https://doi.org/10.1175/bams-d-11-00094.1.

Taylor KE, Balaj V, Hankin S, Juckes M, Lawrence B., Pascoe S. 2012b. *CMIP5 Data Reference Syntax (DRS) and Controlled Vocabularies*. Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory: Livermore, CA. http://cmip-pcmdi.llnl.gov/cmip5/docs/cmip5_data_reference_syntax.pdf (accessed 27 October 2017).

Vera CS, Díaz L. 2015. Anthropogenic influence on summer precipitation trends over South America in CMIP5 models. *Int. J. Climatol.* **35**: 3172–3177. https://doi.org/10.1002/joc.4153.

Voldoire A, Sanchez-Gomez E, Salas y Mélia D, Decharme B, Cassou C, Sénési S, Valcke S, Beau I, Alias A, Chevallier M, Déqué M,

Deshayes J, Douville H, Fernandez E, Madec G, Maisonnave E, Moine M, Planton S, Saint-Martin D, Szopa S, Tyteca S, Alkama R, Belamari S, Braun A, Coquart L, Chauvin F. 2013. The CNRM-CM5.1 global climate model: description and basic evaluation. *Clim. Dyn.* **40**: 2091–2121. https://doi.org/10.1007/s00382-011-1259-y.

Volodin EM, Diansky NA, Gusev AV. 2010. Simulating present-day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations. *Izvestiya Atmos. Ocean. Phys.* **46**: 414–431. https://doi.org/10.1134/s000143381004002x.

Watanabe M, Suzuki T, O'ishi R, Komuro Y, Watanabe S, Emori S, Takemura T, Chikira M, Ogura T, Sekiguchi M, Takata K, Yamazaki D, Yokohata T, Nozawa T, Hasumi H, Tatebe HKimoto M. 2010. Improved climate simulation by MIROC5: mean states, variability, and climate sensitivity. *J. Clim.* **23**: 6312–6335. https://doi.org/10.1175/2010jcli3679.1.

Xin X, Wu T, Zhang J. 2013. Introduction of CMIP5 experiments carried out with the climate system models of Beijing climate center. *Adv. Clim. Change Res.* **4**: 41–49. https://doi.org/10.3724/sp.j.1248.2013.00041.

Yukimoto S, Adachi Y, Hosaka M, Sakami T, Yoshimura H, Hirabara M, Tanaka T, Shindo E, Tsujino H, Deushi M, Mizuta R, Yabu S, Obata A, Nakano H, Koshiro T, Ose TA. 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. https://doi.org/10.2151/jmsj.2012-a02.

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. https://doi.org/10.1007/s00382-012-1361-9.