



Assessment of CORDEX simulations over South America: added value on seasonal climatology and resolution considerations

Magdalena Falco, Andrea F Carril, Claudio G Menéndez, Pablo G Zaninelli,
Laurent Li

► To cite this version:

Magdalena Falco, Andrea F Carril, Claudio G Menéndez, Pablo G Zaninelli, Laurent Li. Assessment of CORDEX simulations over South America: added value on seasonal climatology and resolution considerations. *Climate Dynamics*, 2019, 52 (7-8), pp.4771-4786. 10.1007/s00382-018-4412-z . hal-02187115

HAL Id: hal-02187115

<https://hal.sorbonne-universite.fr/hal-02187115>

Submitted on 17 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Assessment of CORDEX simulations over South America: added value on seasonal climatology and resolution considerations

Magdalena Falco · Andrea F. Carril · Claudio G. Menéndez · Pablo G. Zaninelli · Laurent Z.X. Li

Received: date / Accepted: date

Abstract A new set of CORDEX simulations over South America, together with their coarser-resolution driving Global Climate Models (GCMs) are used to investigate added value of Regional Climate Models (RCMs) in reproducing mean climate conditions over the continent. There are two types of simulations with different lateral boundary conditions: five hindcast simulations use re-analysis as boundary conditions, and five other historical simulations use GCMs outputs. Multi-model ensemble means and individual simulations are evaluated against two or three observation-based gridded datasets for 2-meter surface air temperature and total precipitation. The analysis is performed for summer and winter, over a common period from 1990 to 2004. Results indicate that added value of RCMs is dependent on driving fields, surface properties of the area, season and variable considered. A robust added value for RCMs driven by ERA-Interim is obtained in reproducing the summer climatology of surface air temperature over tropical and subtropical latitudes. Mixed results can be seen, however, for summer precipitation climatology in both hindcast and historical experiments. For winter, there is no noticeable improvement by the RCMs for the large-scale precipitation and surface air temperature climatology. To further understand the added value of RCMs, models deviations from observation are decomposed according to different terms that reflect the observational uncertainty, the representativeness error, the interpolation error, and the actual performance of the model. Regions where these errors are

Magdalena Falco (✉) · Andrea F. Carril · Claudio G. Menéndez · Pablo G. Zaninelli
Centro de Investigaciones del Mar y la Atmósfera (CIMA), CONICET-UBA, Ciudad Universitaria, Ciudad Autónoma de Buenos Aires, Int. Guiraldes 2160, Pabellón 2, Piso 2, C1428EGA Buenos Aires, Argentina
Tel.: +54 (11) 5285-8452
Fax: +54 (11) 4788-3572
E-mail: magdalena.falco@cima.fcen.uba.ar

Magdalena Falco · Andrea F. Carril · Claudio G. Menéndez · Pablo G. Zaninelli
Instituto Franco-Argentino sobre Estudios de Clima y sus Impactos (UMI3351-IFAECI/CNRS-CONICET-UBA), Buenos Aires, Argentina

Magdalena Falco · Claudio G. Menéndez · Pablo G. Zaninelli
Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Buenos Aires, Argentina

Magdalena Falco · Laurent Z.X. Li
Laboratoire de Météorologie Dynamique (LMD), CNRS, Sorbonne Université, École Normale Supérieure, École Polytechnique, Paris, France

Pablo G. Zaninelli
Universidad Nacional de La Plata, Facultad de Ciencias Astronómicas y Geofísicas, La Plata, Buenos Aires, Argentina

not negligible, such as in complex terrain regions, among others, can be identified. There is a clear need for complementary assessment to understand better the real value added by RCMs.

Keywords Model assessment · Seasonal climatology · Added value · South America · CORDEX

1 Introduction

Regional Climate Models (RCMs), also called limited area models, are one of the most common downscaling tools applied to Global Climate Models (GCMs) output or global reanalysis. The main purpose of using RCMs is to provide physically-consistent regional to local climate information in scales that cannot be represented in the coarser-resolution driving data. The fundamental hypothesis of this technique is that while the GCM provides large-scale information to the RCM through lateral boundary conditions, the RCM can reproduce sub-scale information within a limited domain (Laprise, 2008). One of the key debates associated with RCMs is whether they improve their driving GCMs. In this context, added value (AV) is defined as a measure of the extent to which downscaled climate is closer to the observation than the model from which the boundary conditions are obtained (IPCC, 2013).

When it is to identify the added value of an RCM, it is important to note that not all available metrics are optimal for its calculation. The methodology selected should not only assess model performance but should also consider in some way the difference between the resolution of the RCM and its driving data. For example, Gilleland et al. (2009) argue that most traditional verification metrics are based on a point-to-point validation without any consideration of the spatial information. Kanamitsu and DeHaan (2011) illustrate this problem by showing that the identification of high skill regions is lost when averaging over large areas. They also raised the idea of differentiating relevant terms in the deviations between RCM and its driving GCM, including terms related to model resolution or model error. They conclude that the deviations of the model from the validating observations can not only be associated with the model performance but also with model resolution, interpolation, and observational uncertainty. For these reasons, many authors proposed new AV metrics to assess model performance (Denis et al., 2002, Feser and von Storch, 2005, Kanamitsu and DeHaan, 2011, Di Luca et al., 2012, Wang et al., 2015, among others).

The growing interest of the Regional Climate Model Community on assessing the AV of RCMs is accompanied by an important production of scientific articles focused on this issue. General conclusions regarding AV of RCMs are not easy to obtain, as results depend on many factors, such as the study domain, the model evaluated, the reference used, the variable analyzed, the temporal scale, the specific application, the experiment configuration, along others (Torma et al., 2015, Giorgi and Gutowski, 2015, Di Luca et al., 2015). Nevertheless, a general consensus is that AV signal is expected for mesoscale atmosphere phenomena (Feser et al., 2011, Di Luca et al., 2012, Karmacharya et al., 2016, Lucas-Picher et al., 2016), over regions with complex surface forcing (Gao et al., 2006, De Haan et al., 2014, Di Luca et al., 2012, Prömmel et al., 2010, Feser and von Storch, 2005, Lenz et al., 2017), and also for extreme events (Coppola et al., 2014, Giorgi, 2002, Di Luca et al., 2012, Rummukainen, 2016, Prein et al., 2016). One of the main debates is whether RCMs can add value for large-scale mean values. While some works find an improvement on representing the mean large-scale features (Kerkhoff et al., 2014, Llopart et al., 2014, Nikulin et al., 2012, Torma et al., 2015, Fotso-Nguemo et al., 2017, Veljovic et al., 2010, Prein et al., 2016), others find a deterioration (Coppola et al., 2014, Feser, 2006, Giorgi et al., 2014, Roads et al., 2003, Mishra et al., 2017).

Over South America (SA), the limited computational resources currently available in local research institutions makes RCMs the most viable tool for the production of high-resolution climate

simulations. This has, in turn, prompted the need for understanding regional climate modeling strengths and weaknesses over the continent by establishing common experimental frameworks. Under these circumstances, Roads et al. (2003) were the first to analyze an ensemble of four RCMs driven by the NCEP/NCAR I global reanalysis (Kalnay et al., 1996), followed by CREAS regional program (Marengo et al., 2009). The CLARIS (Menéndez et al., 2010) and CLARIS-LPB Projects (Boulanger et al., 2016) succeeded to create a large ensemble of regional simulations, involving different RCMs driven by ERA-Interim reanalysis (Dee et al., 2011) and CMIP3 Global Climate Models (Meehl et al., 2007). The CLARIS-LPB Project comprehended two types of simulations for the recent past concerning the boundary forcing: the hindcast simulations and the historical simulations.

Simulating SAs climate is challenging due to the extension of the continent, and the local to regional climatic features associated to the South American Monsoon (SAM) System over tropical and subtropical latitudes, and also to frontal activities over mid-latitudes (Garreaud et al., 2009). While GCMs are generally capable of reproducing the main large-scale circulation characteristics, their performance deteriorates when evaluating regional aspects of climate (Gulizia et al., 2013, Rusticucci et al., 2010, Silvestri and Vera, 2008, Zazulie et al., 2017). For example, Gulizia et al. (2013) and Zazulie et al. (2017) found that CMIP5 GCM simulations (Taylor et al., 2012) do not reproduce adequately the moisture transport associated with the South American Low-Level Jet. Previous works have assessed regional climate simulations over SA within the context of CLARIS-LPB Project (Carril et al., 2012, 2016, López-Franca et al., 2016, Sánchez et al., 2015, Solman, 2016, Solman et al., 2013, Zaninelli et al., 2015, 2018, Sanabria and Carril, 2018, Menéndez et al., 2016), indicating an overall good performance of RCMs in reproducing different aspects of the observed climate. However, none of these works focus entirely on the AV issue, and only a few of them (Sánchez et al., 2015, Solman, 2016) make a general comparison of the RCM with its driving data.

Apart from CLARIS-LPB Project, the CREMA experiment (Coppola et al., 2014) succeeded to produce several simulations of RegCM4 RCM forced by different GCMs and reanalysis over SA. When comparing these simulations with their driving data, the results suggest an added value of RegCM4 in reproducing extreme events over SA (Coppola et al., 2014, Giorgi et al., 2014). Also, the model partly reduces the dry bias of precipitation in the Amazon Basin when comparing to different CMIP5 GCMs (Llopart et al., 2014, da Rocha et al., 2014). Some few more works have addressed the added value issue over SA, and there are no matching results: Roads et al. (2003) assessed the capabilities of RCMs to make skillful forecasts compared to its forcing data over SA, finding no noticeable improvement in precipitation when assessing the large-scale fields; Seth et al. (2007) studied the performance of RCMs forced by reanalysis for different regions on tropical and extratropical latitudes in SA. In their work they concluded that regions where remote influences are strong (Northeast SA, Southeast SA, and Amazon regions) the AV of RCMs is low; by contrast, where local processes dominate and remote influences are weak (maximum monsoonal precipitation region) there is some potential for the regional model to add value. Finally, De Sales and Xue (2011) assessed the role of resolution on the South American climate simulation by performing a sensitivity test with the Andean topography represented by different heights. They suggested that height representation of the Andes Mountain Range topography impact on low-level moisture flux transport and therefore on precipitation in La Plata Basin.

Within the CORDEX framework (Giorgi and Gutowski, 2015), several simulations with updated model versions have been recently incorporated to the CORDEX SA domain (Fig. 1), enabling the availability of state-of-the-art nested RCMs at a horizontal resolution of $\sim 0.44^\circ \times \sim 0.44^\circ$. Therefore, motivated by the availability of new simulations and the lack of studies about AV over the continent, this work aims to assess the AV of CORDEX simulations in

South America. As a first step, we evaluate the capacity of RCMs to simulate seasonal mean precipitation and surface air temperature over SA through its comparison to their coarser-resolution driving data. We believe this is a fundamental step, not only for RCM evaluation but also due to the discrepancies in the literature whether RCMs can add value or not over large-scale mean values (Llopart et al., 2014, Roads et al., 2003).

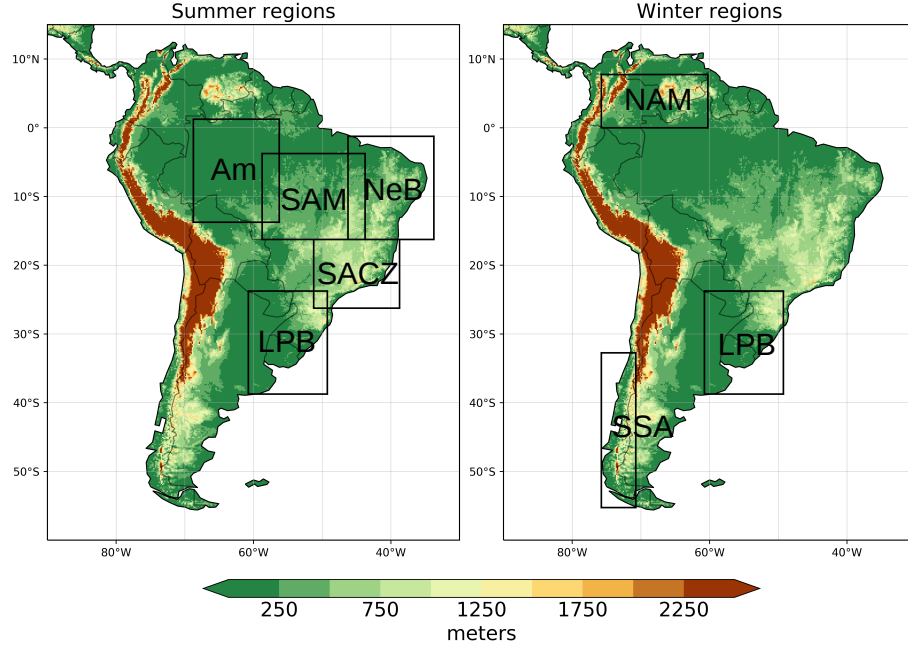


Fig. 1 CORDEX region 1 domain, with ETOPO2 (National Geophysical Data Center, 2006) terrain height (m) indicated in shading. Boxes in solid-black lines denote areas selected for area-mean assessment.

The paper is organized as follows: in section 2 we first describe the observed and simulated data used in this study, followed by the introduction of model errors and observed uncertainties. In section 3 we assess the general and regional performance of individual and multi-model ensemble data in simulating the seasonal spatial patterns of the mean state, including a regional assessment of model errors and observational uncertainties. Finally, in section 4, we present a discussion of the most relevant results together with our main conclusions.

2 Data and methods

2.1 CORDEX simulations and observed data

For our analysis, we selected the common time period of the integrations, covering from 1990 to 2004, to keep the same time-scale variability. The seasons considered were austral summer (December, January, and February) and winter (June, July, and August). All RCMs use roughly the same domain and horizontal resolution ($\sim 0.44^\circ \times \sim 0.44^\circ$), but different grid types. For a direct assessment and intercomparison of RCM simulations, we interpolated all original monthly

precipitation and 2-meter-temperature data onto a regularly spaced latitude/longitude grid of $0.5^\circ \times 0.5^\circ$. ERAi and CMIP5 GCM fields were also interpolated into a common regularly spaced grid of $1.5^\circ \times 1.5^\circ$. Both variables were interpolated using a bilinear interpolation procedure, and a height correction was applied to the surface air temperature fields. The height correction consisted of adding the difference between the model native height (interpolated into the regular grid) and the observed height, multiplied by a standard lapse rate of $-6.5^\circ/\text{km}$.

Table 1 Overview of Regional Climate Models used in the present study and their boundary conditions.

Experiment Acronym	RCM id / Reference	Resolution / Number of grid points	Boundary condition id / Reference	Boundary resolution
HadRM3P-fERAi	MOHC-HadRM3P Gordon et al. (2000)	$0.44^\circ \times 0.44^\circ$ 146x167	ECMWF-ERAINT Dee et al. (2011)	$1.5^\circ \times 1.5^\circ$
REMO2009-fERAi	MPI-CSC-REMO2009 Jacob et al. (2012)	$0.44^\circ \times 0.44^\circ$ 143x167	ECMWF-ERAINT	$1.5^\circ \times 1.5^\circ$
RCA4-fERAi	SMHI-RCA4 Kupiainen et al. (2014)	$0.44^\circ \times 0.44^\circ$ 146x167	ECMWF-ERAINT	$1.5^\circ \times 1.5^\circ$
RegCM4-fERAi	ICTP-RegCM4-3 Giorgi et al. (2012)	$0.44^\circ \times 0.44^\circ$ 189x199	ECMWF-ERAINT	$1.5^\circ \times 1.5^\circ$
WRF341I-fERAi	UCAN-WRF341I Skamarock et al. (2005)	$0.44^\circ \times 0.44^\circ$ 146x167	ECMWF-ERAINT	$0.75^\circ \times 0.75^\circ$
RCA4-fEC-EARTH	SMHI-RCA4 Kupiainen et al. (2014)	$0.44^\circ \times 0.44^\circ$ 146x167	EC-EARTH Hazeleger et al. (2010)	$1.25^\circ \times 1.25^\circ$
REMO2009-fMPI	MPI-CSC-REMO2009 Jacob et al. (2012)	$0.44^\circ \times 0.44^\circ$ 143x167	MPI-ESM-LR Stevens et al. (2013)	$1.875^\circ \times 1.875^\circ$
WRF341I-fCanESM2	UCAN-WRF341I Skamarock et al. (2005)	$0.44^\circ \times 0.44^\circ$ 146x167	CCCma-CanESM2 Chylek et al. (2011)	$1.875^\circ \times 1.875^\circ$
RegCM4-fHadGEM2	ICTP-RegCM4-3 Giorgi et al. (2012)	$0.44^\circ \times 0.44^\circ$ 189x199	MOHC-HadGEM2-ES Collins et al. (2011)	$1.875^\circ \times 1.25^\circ$
LMDZ4-fIPSL-CM5A	IPSL-LMDZ4 Hourdin et al. (2006)	$\sim 0.48^\circ \times \sim 0.48^\circ$ 184x180	IPSL-CM5A Hourdin et al. (2012)	$3.75^\circ \times 1.9^\circ$

Different gridded data for precipitation and surface air temperature were used to evaluate the simulations. With the goal of measuring the uncertainty among observation data sets, we consider three station-based monthly data for precipitation (CRU, UDEL and CPC-UNI), over land on a $0.5^\circ \times 0.5^\circ$ resolution. We also considered two of the above data sets (CRU and UDEL) for surface air temperature. Complementary to the interpolation of model simulations, the observed datasets were also interpolated using the same procedure to the same $0.5^\circ \times 0.5^\circ$ regular grid, and also, to the $1.5^\circ \times 1.5^\circ$ grid to allow GCM evaluation. Throughout the work, the set of observed datasets is denoted as Obs, and the reference field for the assessment is calculated as the average across all the observed datasets.

2.2 Model error and observational uncertainties

The main advantage of RCMs is they are integrated with a higher spatial resolution than their boundary-condition providers and, therefore, add to the latter information on a regional scale. Thus, one can distinguish between two different spatial scales: the small scales that can be reproduced only by the RCM and the large scales reproduced by both simulations. However, when it comes to consciously comparing an RCM with its driving data, there is no optimal methodology to follow (Prein et al., 2016). On the one hand, if we choose to interpolate the

RCM and the observations to the resolution of the GCM, we would benefit the GCM, since we are removing the small-scale variability of the RCM. But if we interpolate the global model to the resolution of the regional model and the observations, then the GCM is disfavored since the reduction of the grid will generate non-consistent physical information. In this paper, we performed the evaluation of the experiments in their resolution, that is, in a 0.5° grid for the regional simulations and a 1.5° grid for the global simulations. However, for the evaluation to be adequate, we considered other sources of error associated with grid interpolation, resolution and observational uncertainty.

Kanamitsu and DeHaan (2011) introduced a clear picture of model error or deviation from the observed reference by describing the whole process of model validation. Firstly, each grid point value in model represents an entire area of the field, and therefore the model will not be able to reproduce the sub-grid variability. This error or limitation can be called representativeness error and is a function of resolution of the model and variability of the field. Secondly, interpolation error can be identified as those changes of a field when interpolated from the model grid into a validation grid. The magnitude of this error is also influenced by the model's resolution and the variability of the field. Finally, the observations used for the evaluation are not perfect, and generally have errors associated with instrumental and data acquisition imperfections. Interpolation and representativeness errors should also be counted when using station-based gridded datasets. The concept of Kanamitsu and DeHaan (2011), expressed as an apparent deviation between the model and observation for a single location, can be qualitatively formulated as

$$Model\ Error = [\varepsilon_M] + [\varepsilon_R] + \varepsilon_I + \varepsilon_{obs}, \quad (1)$$

where $[\]$ is a spatial interpolation operator, ε_M is the model deviation from observations, ε_R is the representativeness error of model grid value, ε_I is the interpolation error, and ε_{obs} is the error associated with observation. Eq. (1) shows that model deviation seen in a validation operation is directly related to the performance of the model, i.e., the true model error, if and only if ε_R , ε_I , and ε_{obs} are negligible. In real world, they are generally not negligible, and they need an appropriate evaluation, which allows us to have just an appreciation for the true performance (or error) of a model.

In our work, we revisit Eq. (1) and adapt it to our experimental design. Our reference observations for evaluation are station-based gridded datasets, so the grid spacing of the observational dataset defines the finest scale on which the comparisons are meaningful (Prein et al., 2016). All observed datasets have a resolution of $0.5^\circ \times 0.5^\circ$. This entails that it is impossible to quantify the sub-grid representativeness error of RCMs, as they have a finer resolution. Following Eq. (1), the deviation of both RCM experiments from the reference field will be affected by the following terms:

$$\varepsilon_{RCM}^m = f(\varepsilon_M^m, \varepsilon_R^{0.5^\circ}, \varepsilon_I^{0.5^\circ}, \varepsilon_{obs}), \quad (2)$$

where m can be RCMs-fERAi or RCMs-fCMIP5, $\varepsilon_R^{0.5^\circ}$ is the representativeness error at a resolution of 0.5° , that is neglected hereafter since its resolution is close to that of the observations, and $\varepsilon_I^{0.5^\circ}$ is the interpolation error at a resolution of 0.5° . In the same way, the errors associated with GCMs will include the same terms but for a coarser resolution grid. Therefore, the deviation of both GCM experiments from the reference field will be affected by the following terms:

$$\varepsilon_{GCM}^m = f(\varepsilon_M^m, \varepsilon_R^{1.5^\circ}, \varepsilon_I^{1.5^\circ}, \varepsilon_{obs}), \quad (3)$$

again, m can be ERAi or CMIP5, $\varepsilon_I^{1.5^\circ}$ is the interpolation error at a resolution of 1.5° , and $\varepsilon_R^{1.5^\circ}$ is the representativeness error between the resolutions of 0.5° and 1.5° . In this case, the term

associated with the representativeness error is not negligible, as degrading the observed field into the GCM resolution for evaluation implies losing details in the process of filtering.

In this work, we present a methodology that accounts for the different terms involved in the process of model evaluation. Despite strong constraints and difficulties in calculating precisely each term in Eqs. (2) and (3), we try to estimate their order of magnitude. We aim to find regions with significant contributions in order to determine if the model's deviations from observation are genuinely associated with model performance or not.

Starting with the model error, ε_M , we cannot execute a temporal validation as historical simulations are not synchronized with the observed variability. Consequently, we perform a multi-model evaluation of the seasonal climatology by using the root-mean square (RMS) error metric. We define ε_M as the RMS deviations of all individual members of an ensemble m from the reference data, that is

$$\varepsilon_M^m = \sqrt{\frac{1}{N_m} \sum_{i=1}^{N_m} ([F_{m,i}]^{res} - [F_{ref}]^{res})^2} \quad (4)$$

with N_m the number of associated simulations, and $[]^{res}$ the interpolation of the simulation and station-based observed datasets to a regular grid of resolution res (0.5° for both RCMs and 1.5° for both GCMs). It is worthy to note that the reference field is calculated as an average among available datasets in the resolution of each experiment.

The degradation of the observed field from 0.5° to 1.5° is done by aggregating the reference field at 0.5° into a coarser-resolution grid of $1.5^\circ \times 1.5^\circ$. The aggregation is performed by averaging all original sub-grid values inside each of the coarser grid boxes (Di Luca et al., 2012). The aggregated values are then re-distributed into the original resolution, which will give an estimation of the details lost in the process of filtering. The error ε_R is calculated as the RMS difference between the degraded and reconstructed fields

$$\varepsilon_R^{1.5^\circ} = \sqrt{(\overline{F_{ref}^{0.5^\circ}} - \overline{F_{ref}^{1.5^\circ}})^2} = | \overline{F_{ref}^{0.5^\circ}} - \overline{F_{ref}^{1.5^\circ}} |. \quad (5)$$

Eq. (6) permits to measure the information contained in the range of resolutions from 0.5° to 1.5° .

Unlike Kanamitsu and DeHaan (2011), we won't neglect ε_I and ε_{obs} as both terms can be significant over some regions. While there is no easy way to calculate the interpolation error, we estimate it by interpolating the mean observed climatological field into a regularly-spaced but staggered grid, and then interpolating it back to the original grid. For the $0.5^\circ \times 0.5^\circ$ resolution we consider a displacement of 0.25° in both latitude and longitude directions and for the $1.5^\circ \times 1.5^\circ$ a displacement of 0.75° . The interpolation error is then estimated as half the difference between the double interpolated field and the original field

$$\varepsilon_I^{res} = \frac{\sqrt{([\overline{F_{ref}^{res}}] - \overline{F_{ref}^{res}})^2}}{2} = \frac{| [\overline{F_{ref}^{res}}] - \overline{F_{ref}^{res}} |}{2}. \quad (6)$$

Finally, we need to point out that a main part of observation errors comes from the representativeness error and interpolation error when station-based information is converted into a regular grid. But its estimation is out of the scope of our study. Nevertheless, we can estimate the observational uncertainty as the spreading among individual datasets. For that, we substitute m by Obs and res by 0.5° in Eq. (4), and obtain

$$\varepsilon_{obs} = \sqrt{\frac{1}{N_{Obs}} \sum_{i=1}^{N_{Obs}} ([F_{Obs,i}]^{0.5} - [F_{ref}]^{0.5})^2}. \quad (7)$$

3 Results

3.1 Seasonal spatial patterns of the mean state

The challenge for regional climate modelers in representing the South American climate comes from the extent of the computational domain and the complexity of the physical phenomena involved. During the austral summer, precipitation maximizes over tropical and subtropical latitudes and is largely associated to the SAM System (Fig. 2a). All multi-model ensemble climatologies can reproduce large-scale features of the summer precipitation, but the presence of deviations from the reference indicates some systematic biases (Figs. 2b-2e). There are some common biases in all experiments, such as an overestimation of precipitation intensity along the Andes Cordillera, a dry bias in northern and central Brazil and a wet bias in Northeast Brazil (NeB). This last behavior is accentuated in historical experiments, especially in RCMs-fCMIP5. All the three ensembles, RCMs-fERAi, RCMs-fCMIP5 and CMIP5, show a wet bias over the South Atlantic Convergence Zone (SACZ) and a dry bias over La Plata Basin (LPB). On the contrary, ERAi has opposite biases in these two regions. The spatial distribution of surface air temperature biases (Figs. 2g-2j) reveals a cold bias over most of the continent in all experiments. This is not true over the Amazon basin, where RCMs-fERAi and CMIP5 ensembles show small positive values. Also, over LPB, most ensembles except ERAi product have a warm bias.

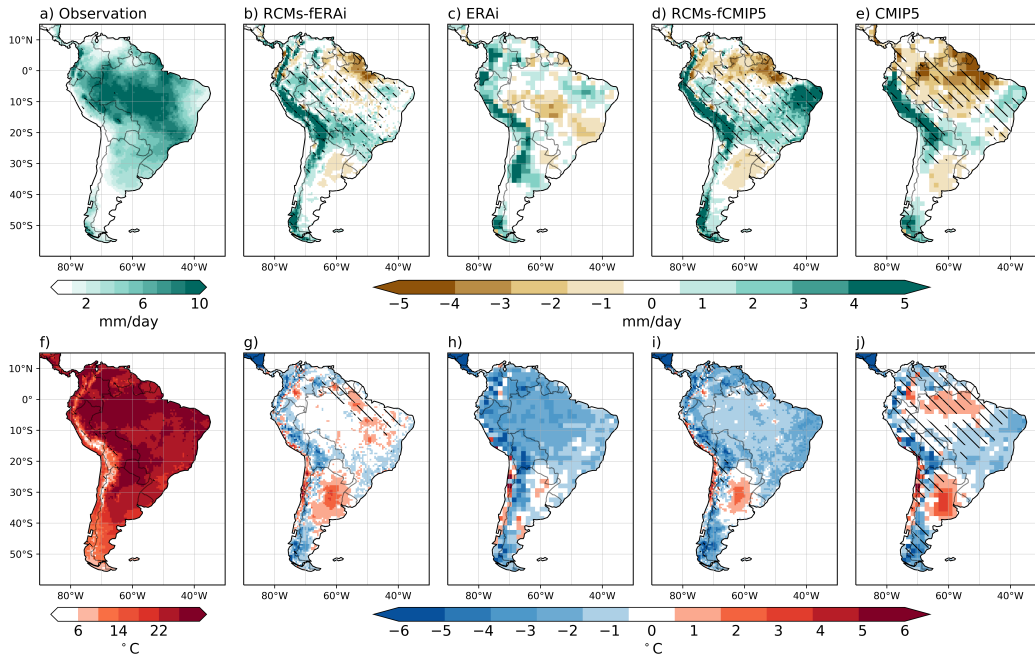


Fig. 2 Summer mean precipitation (mm/day; top panels) and 2-meter surface air temperature (°C; bottom panels) of the observed fields and the multi-model ensembles. The first column is the average of observed datasets and the second, third, forth and fifth columns are the climatologies of RCMs-fERAi, ERAi, RCMs-fCMIP5, and CMIP5, respectively. Hatched areas indicate regions where the inter-model standard deviation is larger than the natural variability, represented by the mean interannual standard deviation of all simulations involved.

Given that results are presented as ensemble mean, it is important to assess the multi-model dispersion over each ensemble. Following Rinke et al. (2006) and Solman et al. (2013), we can highlight those areas where the dispersion among datasets is larger than the interannual variability of the field (hatched areas in Fig. 2). In the case of summer precipitation, all three ensemble mean climatologies have large inter-model spread over most of tropical and subtropical SA (Fig. 2b, 2d, and 2e), suggesting that the ensemble climatology is not very representative of individual climatologies. The absence of hatched areas for observed climatologies (Fig. 2a) indicates that the inter-dataset spread is small over most of SA. Regarding surface air temperature, only CMIP5 ensemble has large hatched areas that denote a large dispersion among simulations over most of the continent.

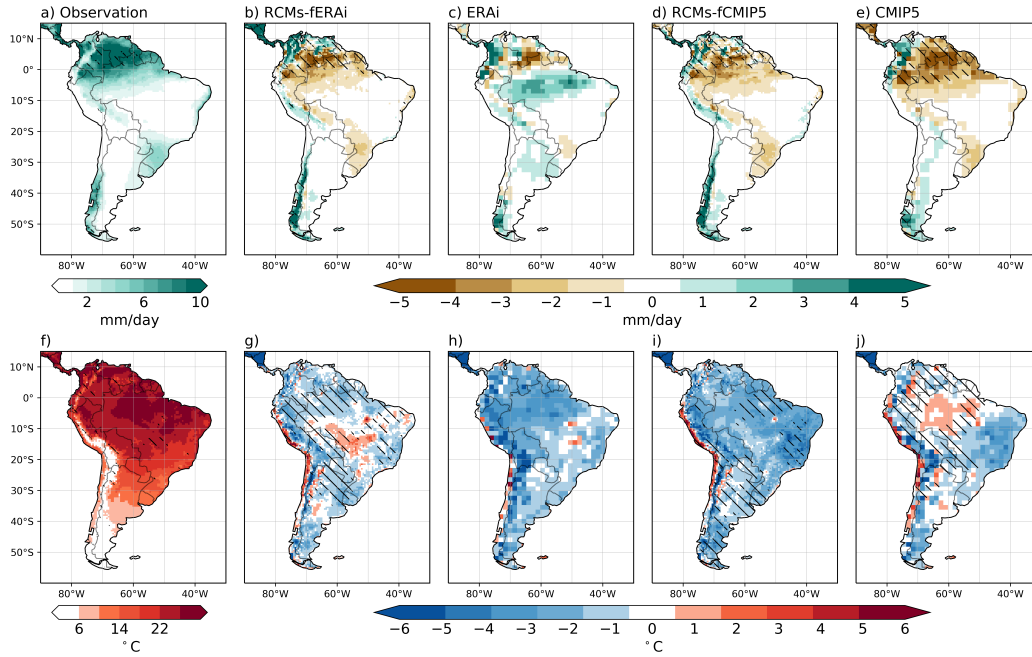


Fig. 3 As in Fig. 2 but for the winter season.

Winter precipitation climatology (Fig. 3a) exhibits two maxima: the first over the northern part of the continent, north of the equator, associated to the North American Monsoon (NAM) System; and the other one over the meridional Andes, south of 40°S, associated to baroclinic activities. While all experiments show the northern precipitation maximum associated to the NAM System (not shown), they fail to capture the observed intensity by largely underestimating it (Figs. 3b-3e). Even ERAi fails to simulate the southward extension of the NAM influence, which is noticed by the wet bias at the southern edge of the region. All ensembles present a significant multi-model spread over the tropics, indicating a wide variation in the representation of the NAM. In the extratropics, all multi-model ensembles overestimate precipitation over the western slope of the Andes. The bias is maximal in RCMs, as they also present an unrealistic extension over northern Chile. Regarding 2-meter temperature (Fig. 3, bottom panels), we note again that most experiments have a cold bias over most of the continent, except RCMs-fERAi multi-model ensemble over the southern Amazon basin and CMIP5 multi-model ensemble over

northern Brazil and LPB. During this cold season, unlike in summer, both RCM ensembles have large inter-model spread for surface-air temperature climatology over the tropical and subtropical latitudes, and also over Patagonia in the case of RCMs-fCMIP5. CMIP5 simulations continue to show large inter-model spread over the same latitudes as in summer.

Taylor diagrams (Taylor, 2001) are multi-parameter-based diagrams that allow intercomparison of models performance in reproducing a reference pattern. They are a useful and visual tool to summarize how well climate models reproduce the observed climate. We first assess models performance over the South American low-level terrain (lower than 1500m) in Fig. 4. Results suggest that there is a clear distinction between models ability to reproduce 2-meter temperature climatology and precipitation climatology (Fig. 4). The spatial correlation between observed and modeled climatology is always higher for 2-meter temperature (ranging from 0.9 to 0.99) than for precipitation (from 0.5 to 0.9). It is interesting to note that in all cases the ensemble mean improves the field when comparing to each individual simulation, as suggested by other authors (Carril et al., 2012, López-Franca et al., 2016). Surface air temperature tends to be closer from each other in the Taylor diagrams compared to precipitation points, which indicates a larger inter-model spread for precipitation.

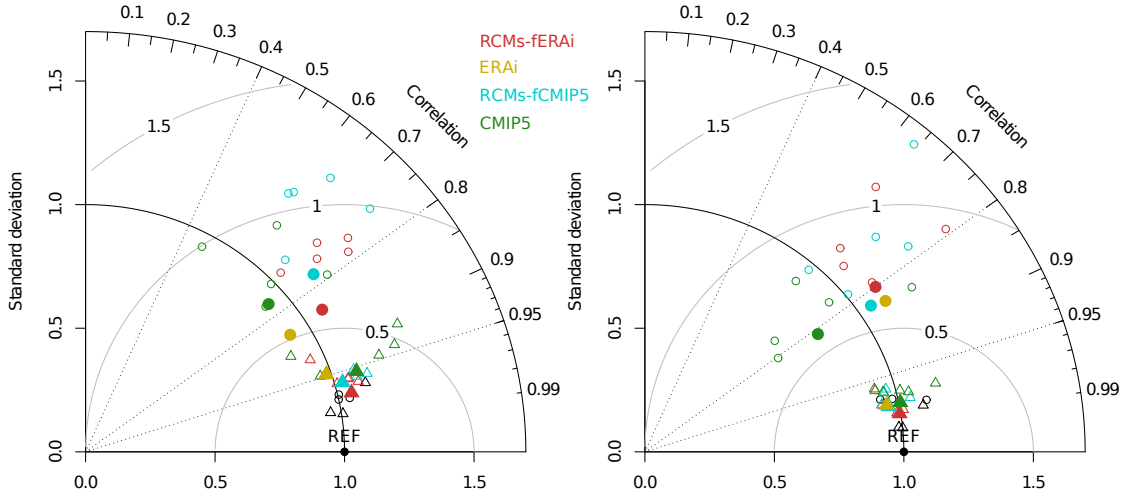


Fig. 4 Taylor diagrams showing the correlation, normalized standard deviation and centered RMSE for individual simulations (empty markers) and for multi-model ensembles (full markers). The left panel corresponds to summer and the right one to winter. The observation ensemble is considered as the reference data. Red, yellow, light-blue and green colors represent the RCMs forced by reanalysis, the reanalysis itself, RCMs forced by CMIP5 and CMIP5 models themselves, respectively. Black markers represent individual observed datasets. Circles correspond to seasonal precipitation climatology and triangles to seasonal surface air temperature climatology. The calculation excludes grid points with topography higher than 1500m.

During summer, ERAi reanalysis and RCMs-fERAi multi-model ensemble have both the best performance in reproducing the observed summer precipitation climatology (Fig. 4a). Their overall performance is similar in terms of correlation but not in terms of normalized standard deviation (nSTD). When compared with the reference, ERAi has lower spatial variability (nSTD<1) and RCMs has higher spatial variability (nSTD>1). These characteristics are also valid when evaluating the ensemble of RCMs-fCMIP5 and their driving data ensemble. During this season, all RCMs-fCMIP5 individual simulations show worse performance, compared to RCMs-fERAi individual simulations, probably due to the quality of the different boundary conditions. Nevertheless, all RCM individual simulations have larger nSTD when compared to their driving data,

suggesting there is an excessive spatial variability. Regarding surface air temperature, the ensemble of RCMs-fERAi is the one showing the best match with the observed climatology pattern by presenting the highest spatial correlation coefficient. In both experiments, there is an added value from RCM ensembles in reproducing the referenced surface air temperature, each individual simulation being also the case. Again, as noted in Fig. 2, CMIP5 simulations tend to have a significant spread during this season.

RCMs-fCMIP5 ensemble is the most skillful multi-model ensemble in simulating the winter precipitation climatology (Fig. 4b). RCMs-fERAi has similar nSTD than its driving data but lower spatial correlation. In the historical experiments, RCMs-fCMIP5 has higher correlation and a spatial variability nearer the observed when comparing to its driving data. Alike summer, RCMs-fERAi continue having the best performance in reproducing the winter 2-meter temperature climatology, but all ensemble and individual simulations have high correlation values and near to 1 nSTD, suggesting an overall good performance.

3.2 Regional performance of individual simulations and multi-model ensembles

Regional-scale climate assessment is fundamental for understanding the potential added value of RCMs since AV is expected only over those regions where stationary climatic features are associated to small-spatial scale processes (Di Luca et al., 2012). For this reason, we emphasize on the evaluation of regional performance by selecting several key areas over the whole continent (Fig. 1). It is clear that RCMs ensembles can add value to the simulated climate when compared to their driving data (Figs. 2 and 3), but it is worthy to note that there is a large spread among simulations (Figs. 2, 3 and 4). We need to evaluate the performance of individual RCM simulations against their driving data, which can reveal the robustness of results. To do so, we use an AV index defined as the ratio between the absolute error (AE) of the RCM and the AE of the GCM. The AE is calculated as the absolute difference between the simulation climatology with the reference, followed by an average over each individual region. Hence, if the index is smaller than one, the RCM has a lower AE than its driving data, suggesting a better representation of the mean climatology over a region. The index is also calculated for the multi-model ensemble climatology to see how representative it is.

Table 2 Added Value Index (AVI) of individual regional climate simulations against their driving data for austral summer (DJF) and for each study region. The index is calculated as the ratio of the area-averaged absolute error (AE) of RCM over that of GCM. Bold figures highlight cases with AVI smaller than 1, indicating added value in RCM.

Experiment Acronym	Am		SAM		NeB		SACZ		LPB	
	pr	tas	pr	tas	pr	tas	pr	tas	pr	tas
RCMs-fERAi	0.80	0.10	0.52	0.21	0.84	0.37	0.81	0.65	0.83	2.06
HadRM3P-fERAi	1.55	0.45	1.42	0.45	1.08	0.41	0.71	0.44	0.61	3.91
REMO2009-fERAi	1.67	0.25	0.97	0.92	1.45	1.59	0.97	0.53	1.12	1.47
RCA4-fERAi	1.44	0.13	1.47	0.27	3.04	0.98	1.13	0.99	1.91	3.16
RegCM4-fERAi	1.36	0.21	0.93	0.36	1.21	0.36	1.35	0.90	1.64	2.90
WRF341I-fERAi	3.19	0.67	0.98	0.75	1.07	1.02	1.12	1.17	1.15	1.77
RCMs-fCMIP5	0.56	1.75	1.23	1.63	2.46	1.44	1.82	1.37	1.15	0.72
RCA4-fEC-EARTH	1.05	0.60	2.49	0.70	3.31	0.92	1.56	1.05	2.17	1.17
REMO2009-fMPI	0.74	0.89	2.08	1.26	2.21	1.23	1.85	0.99	0.95	0.52
WRF341I-fCanESM2	0.58	0.19	0.79	1.16	2.44	3.44	3.72	2.80	0.84	1.43
RegCM4-fHadGEM2	1.51	1.67	1.22	0.80	0.96	1.26	1.22	1.06	1.22	0.92
LMDZ4-fIPSL-CM5A	0.81	0.94	0.89	1.32	1.07	1.66	0.72	1.50	0.43	0.61

When analyzing the AV index over all summer regions (table 2), we find that even though the ensemble of RCMs-fERAi represents the mean value of precipitation better than its driver does, individual simulations show different statistics. Over the five regions, only three of them (SAM, SACZ and LPB) show added value within at least one individual simulation: three simulations (over five) in the case of SAM region, two in the case of SACZ and only one in LPB. For RCMs driven by CMIP5, the RCMs add value over the CMIP5 driver only in Am region, coincident with results in Fig. 9, since 3 of 5 individual simulations have lower bias than its driver. Other regions where individual simulations add value in precipitation are SAM with two simulations, NeB with one simulation, SACZ with one simulation and LPB with three simulations.

Summer surface air temperature shows a more positive outlook for the RCMs-fERAi ensemble, as in Am (5 simulations), SAM (5 simulations), NeB (3 simulations) and SACZ (4 simulations) almost all simulations, together with the ensemble mean, show a lower AE than ERAi. RCMs-fCMIP5 adds value over two regions: Am (4 simulations) and LPB (3 simulations). For Am, the ensemble mean did not represent individual performance. This counter-intuitive result is probably caused by the large spreading among CMIP5 models over this region (Fig. 2j).

Table 3 Same as in Table 2 but for austral winter (JJA).

Experiment Acronym	NAM		LPB		SSA	
	pr	tas	pr	tas	pr	tas
RCMs-fERAi	1.05	0.41	1.25	1.06	1.69	0.94
HadRM3P-fERAi	1.26	0.37	0.70	0.70	2.25	1.00
REMO2009-fERAi	0.77	0.27	0.93	0.65	2.34	1.03
RCA4-fERAi	1.61	0.50	2.73	1.42	1.64	1.00
RegCM4-fERAi	1.24	0.41	1.08	1.20	1.26	1.27
WRF341I-fERAi	1.64	1.08	1.58	2.68	1.58	1.17
RCMs-fCMIP5	0.71	1.73	1.08	2.08	2.23	1.08
RCA4-fEC-EARTH	1.08	0.84	1.94	1.98	1.62	1.04
REMO2009-fMPI	0.65	0.71	1.11	1.37	2.31	0.84
WRF341I-fCanESM2	0.69	1.20	0.74	3.89	1.55	1.35
RegCM4-fHadGEM2	1.73	1.04	1.53	1.43	0.82	0.91
LMDZ4-fIPSL-CM5A	1.10	0.94	0.70	2.09	2.43	1.18

For winter precipitation, Table 3 shows that AV can only be found in some individual simulations over NAM region (2 simulations of CMIP5-driving RCMs and one of ERAi-driving RCMs), four simulations in LPB (two in each experiment) and only one simulation in SSA. Regarding surface air temperature, most models show an improvement over NAM region for both experiments (5 in ERAi-driving RCMs and 3 in CMIP5-driving RCMs), but in the case of CMIP5-driving RCMs the higher skill is not represented by the ensemble mean. For RCMs-fERAi, there are two individual simulations showing added value in LPB and also two simulations in SSA with an AV index of 1.

Box-and-whisker plots are an excellent non-parametric tool for evaluating quartile distribution, allowing us to visually analyze and compare various L-estimators beyond the mean value. Assessing the distribution of rainfall and surface air temperature allows us not only to assess the biases in the mean climate but also to evaluate the interannual variability of the simulations. Generally speaking, models are far from correctly reproducing the observed variability (Figs. 5 and 6). However, in most regions, data sets broadly reproduce the observed statistical distribution by intersecting at least part of the inter-quartile range (IQR; that is 25th-75th quartile distance) with the corresponding observed one.

Most of the experiments tend to have a systematic cold bias over most regions during summer (Fig. 5). This is especially true for ERAi reanalysis and RCMs-fCMIP5 ensemble. Precipitation is

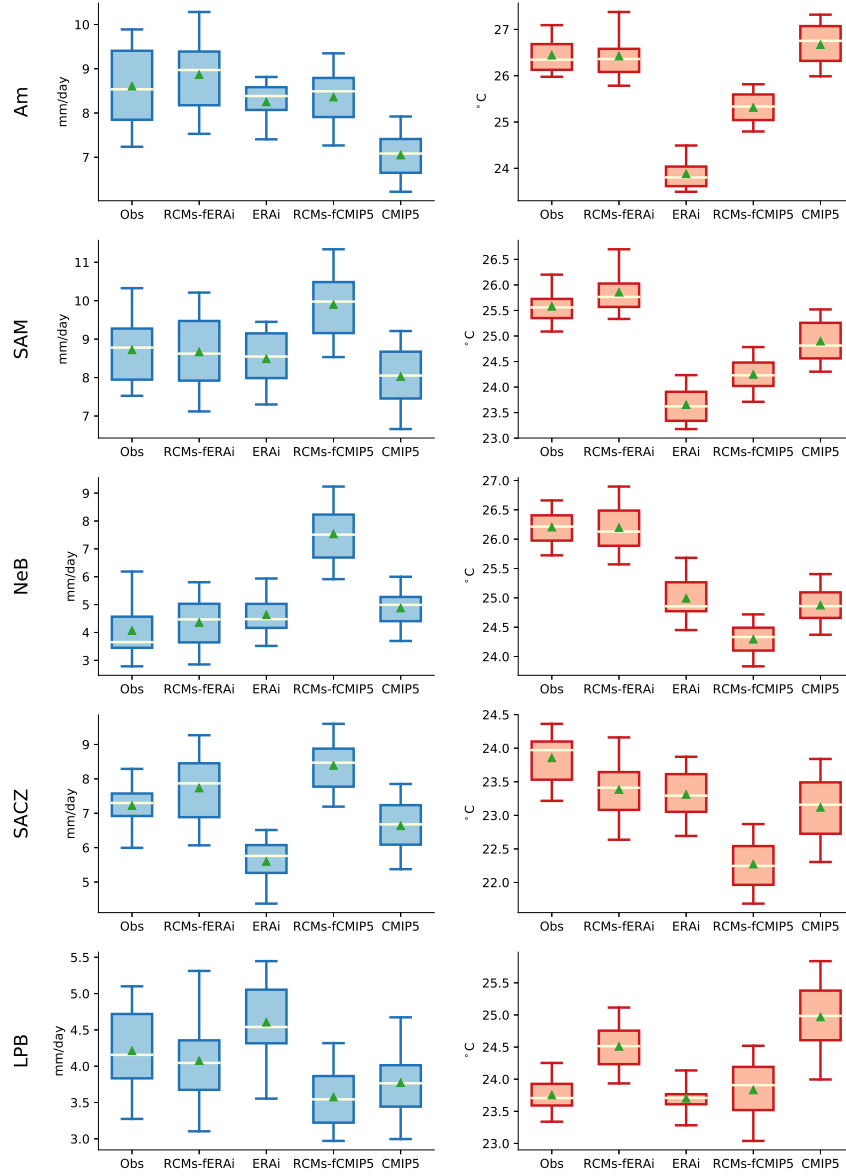


Fig. 5 Box-and-whisker plots showing the mean temporal summer distribution statistics in each study region. Every box represents the inter-quartile range (25th-75th percentile), the median is indicated with a white line and the mean with a green triangle. The whiskers indicate the 95th and 5th percentiles. The regional time series of each simulation is obtained by spatially averaging all values within the region in every time step. Also, all percentiles of the multi-model ensembles are calculated as the average of individual percentiles. The left panels correspond to precipitation (mm/day; blue) and the right panels to 2-meter temperature (°C; red).

well simulated in all experiments and over most regions, an exception is RCMs-fCMIP5 ensemble which has a systematic wet bias over SAM, NeB and SACZ. Concerning the AV of RCMs in precipitation, there is a little improvement compared to its driving data. RCMs-fERAi improves ERAi by better reproducing the observed median value and the interannual variability over Am region. There are also some improvements over LPB and SACZ for the median value but not so much for the interannual variability. RCMs-fCMIP5 only slightly improve its driving data for precipitation distribution over LPB. Regarding surface air temperature, the added value is more evident, especially for the case of hindcast simulations. RCMs-fERAi improves the reproduction of the observed distribution over Am, SAM, NeB and SACZ region. On the other hand, RCMs-fCMIP5 ensemble does not seem to reduce the bias of CMIP5 simulations, but they rather improve the representation of the interannual variability; this is the case for Am, SAM, SACZ and LPB regions.

During winter, there is a general disability of climate models in reproducing the observed statistics over NAM, by systematically underestimating both temperature and precipitation (Fig. 6). AV hardly manifests, but RCMs-fERAi reduces the large cold bias that ERAi has over the NAM region and the same region sees an improvement in RCMs-fCMIP5 which reduces the dry bias of CMIP5 simulations. Winter precipitation statistics over SSA indicate a general systematic overestimation of rainfall by all models. This is especially true for both RCM ensembles, which lowers its performance when comparing to their driving data.

3.3 Assessment on model errors and observational uncertainties

The assessment performed in the previous sections combines different types of error, which makes difficult to understand how the RCM simulations add value (or not), compared to its coarser-resolution driver. We follow the resolution-aware methodology described in the methods section to identify the type of error that should not be neglected during the assessment. A regional comparison of all the terms presented in Eqs. (2) and (3) is performed over each region for both temperature and precipitation. We present the average value of each term over each region for DJF and JJA respectively (Fig. 7). The observational uncertainty, ε_{obs} , is considered as a reference value, as no assessment can be done when errors are smaller than this term. Results show that model error (ε_M) is usually the largest one. We recall that it generally represents the model performance that we actually want to evaluate. However, ε_R , and ε_I are not always negligible. The spatial distribution of these errors (see supplementary material) shows they are particularly large over complex terrain regions, especially for surface air temperature. This result is also evident in the regional diagram (Fig. 7), where the coarse-resolution errors, i.e., $\varepsilon_I^{1.5}$ and $\varepsilon_R^{1.5}$, are relatively large over SACZ in summer and NAM and SSA in winter. ε_R is usually larger than ε_I , being the latter the one that presents the smallest values when comparing to the other terms. ε_I is also dependent of resolution, being the interpolation error associated to a grid of 1.5° resolution ($\varepsilon_I^{1.5}$) larger than the one from a grid of 0.5° resolution ($\varepsilon_I^{0.5}$).

In summer, all $\varepsilon_R^{1.5}$ and $\varepsilon_I^{1.5}$ are of a similar order of magnitude, close to ε_{obs} for both temperature and precipitation. $\varepsilon_R^{1.5}$ often exceeds ε_{obs} value, as is the case in NeB and SACZ for precipitation and in NeB, SACZ and LPB for surface air temperature (Fig. 7, left panels). When analyzing the regional performance of models, ERAi has the best performance in reproducing regional precipitation climatology, since ε_M^{ERAi} has the lowest values compared to $\varepsilon_M^{RCMs-fERAi}$, $\varepsilon_M^{RCMs-fCMIP5}$ and ε_M^{CMIP5} over most regions. This is the case over Am, SAM, NeB and LPB regions. For surface air temperature, the experiment with lower values is $\varepsilon_M^{RCMs-fERAi}$, mainly over Am, SAM, NeB and SACZ region. The results are not very encouraging for RCM historical simulation ensemble for precipitation, but improve for surface air temperature: $\varepsilon_M^{RCMs-fCMIP5}$

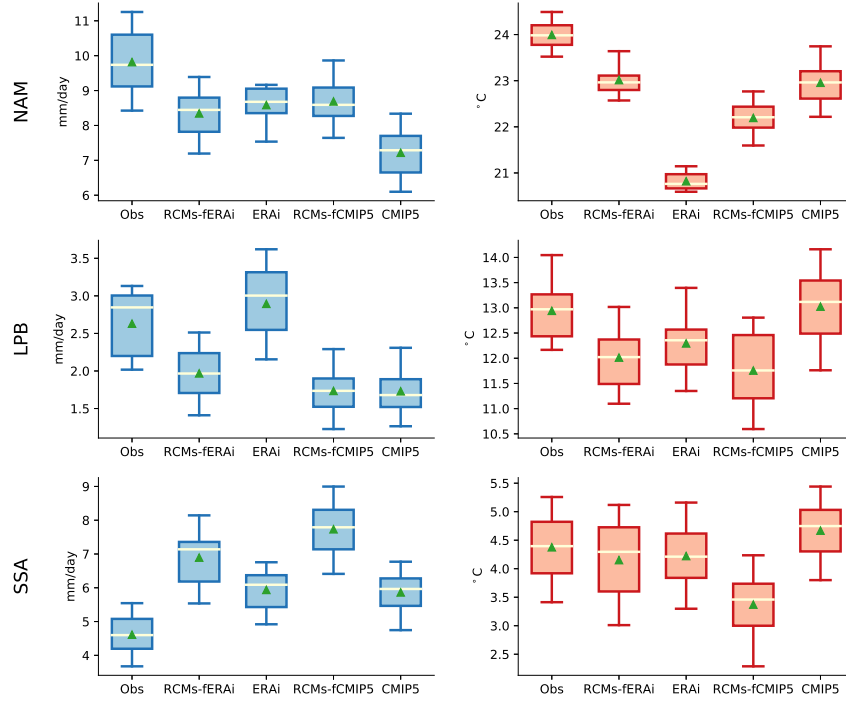


Fig. 6 As in Fig. 5 but for the winter season.

is lower than ε_M^{CMIP5} only in Am region for precipitation and in Am, SAM, and LPB for surface air temperature. In addition to this, $\varepsilon_M^{RCMs-fCMIP5}$ presents an outlier in precipitation over NeB, with a value of 4.7 mm/day.

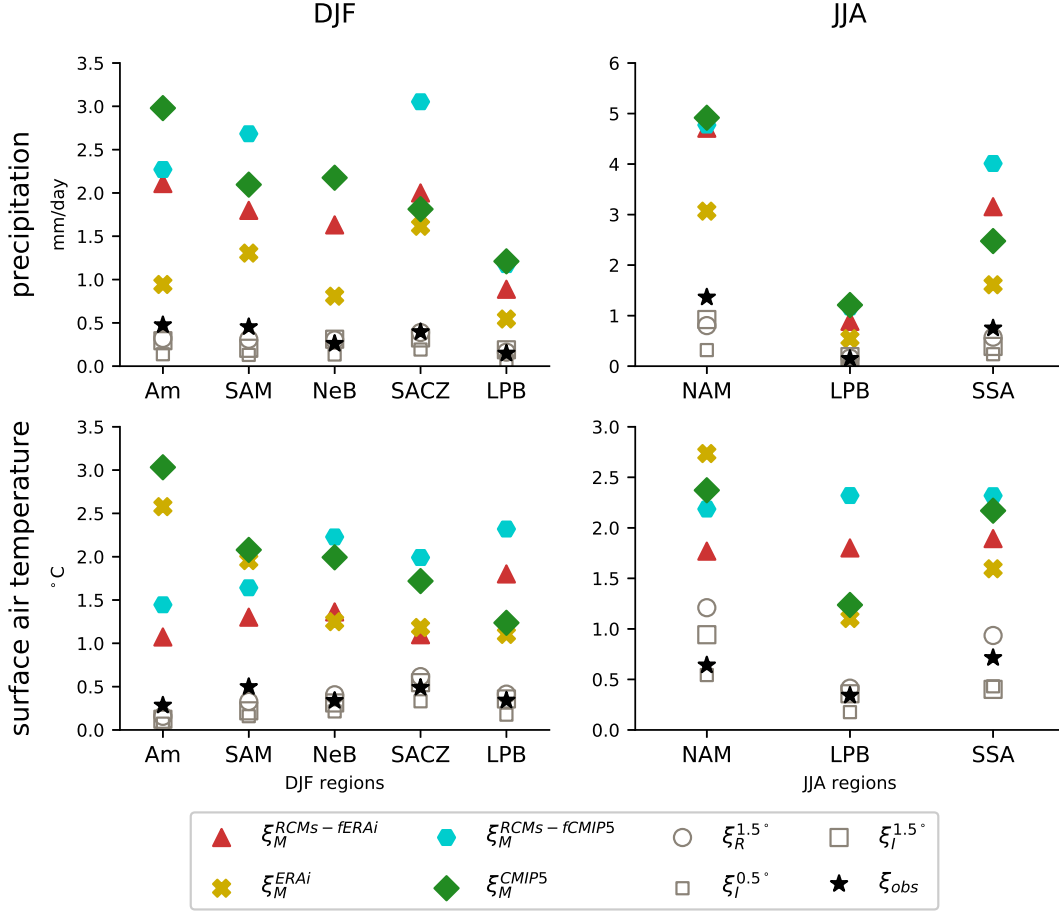


Fig. 7 Diagrams summarizing different errors or deviations obtained with our best estimation for each of our regions of interest. Left panels are for austral summer DJF, and right ones for austral winter JJA. Values plotted in the form of different symbols represent the spatial average of different terms in Eqs. (2) and (3). Top panels are for precipitation (mm/day), and bottom panels for 2-meter temperature (°C). Red, yellow, light-blue and green colors represent our estimations in RCMs forced by reanalysis, reanalysis, RCMs forced by CMIP5 and CMIP5, respectively. Black markers represent our estimation for observational uncertainty. Grey empty markers correspond to interpolation (square markers) and representativeness (circle markers) errors.

During winter, ϵ_{obs} is large for precipitation over the NAM region (Fig. 7, right panels), as was already seen with the spread among datasets in Fig. 3a. The uncertainty is also relatively large (larger than 0.5 °C) over NAM and SSA for surface air temperature, probably because of the selection of domains placed over complex terrain regions. $\epsilon_R^{1.5^\circ}$ is larger than the observational uncertainty for surface air temperature, by exceeding it over all three winter regions (i.e., NAM, LPB and SSA). $\epsilon_I^{1.5^\circ}$ is also larger than ϵ_{obs} in NAM for surface air temperature. Regarding ϵ_M , RCMs usually have similar or larger RMSE in precipitation than its driving data; except for the case of NAM region where $\epsilon_M^{RCMs-fCMIP5}$ has a lower value than ϵ_M^{CMIP5} . The aforementioned also represents the case for 2-meter temperature, except for RCMs-fERAi over LPB.

4 Conclusions and discussion

Although many CORDEX simulations over SA became available recently, there was a general lack of works investigating the AV of RCMs over the continent. It is thus very relevant for us to assess these regional simulations, which can provide a useful guide for researchers or end users of these simulations. The particularity of our approach was in the joint evaluation of RCM simulations and their coarser-resolution driving data. We first evaluated their capability in reproducing mean climate conditions over the continent. We used two types of RCM experiments with different nature in lateral boundary conditions: hindcast simulations with RCMs driven by ERA-Interim re-analysis and historical simulations with RCMs driven by CMIP5 global models. Our database included five hindcast simulations and five other historical simulations, as well as their driving data. The multi-model ensembles and individual simulations were evaluated against two or three observed gridded datasets for 2-meter temperature and total precipitation, respectively. We performed the analysis on a seasonal basis, for austral summer and winter, over a common period ranging from 1990 to 2004.

Generally speaking, all multi-model ensembles can reproduce main features of summer and winter climatologies. However, individual simulations exhibit a large inter-model spread, especially for precipitation, suggesting the need for assessing models individually. The ensembles also have some common model biases, as it is the overestimation of precipitation intensity along the Andes Cordillera. An important bias is found during the warm season since the ensembles differ substantially in their representation of location and intensity of the maximum monsoon rainfall. Regarding surface air temperature, models tend to simulate lower temperatures than the observed, as shown by in the bias fields and the box-and-whisker plots. This cold bias is maximized by ERAi and RCMs-fCMIP5 over the Amazon Basin region. Over the tropical and subtropical latitudes, CMIP5 models show a large spread among simulations in reproducing the surface air temperature climatology for both winter and summer seasons.

Solman et al. (2013) assessed the capability of a set of 7 hindcast RCMs, performed in the framework of the CLARIS-LPB Project, in reproducing the mean climate conditions over the South American continent. If we compare the CLARIS-LPB simulations with the ones shown here for CORDEX, we find that the distribution and intensity of biases are very similar (see Figs. 2 and 3 from Solman et al.). However, some improvements are found for CORDEX such as a reduction of the precipitation bias over LPB during winter, and a reduction of the warm bias over central SA in the same season. Also, our results are consistent with Solman et al. (2013) in terms of spread among simulations, both showing that the inter-model uncertainty is larger than the observational uncertainty.

We performed a regional assessment in several key areas of the whole continent in summer and winter. In summer, RCMs-fERAi ensemble shows a better reproduction of observed precipitation and surface air temperature when compared to ERAi. However, some opposite results can be found for precipitation when conducting the comparison individually. There is a clear warning message on the practice of only analyzing the multi-model ensemble, especially over regions with large inter-model spread. Surface air temperature exhibits some encouraging results, with a general improvement when compared to ERAi. Regarding historical simulations with RCMs driven by CMIP5, results are inconclusive, since added value is only found in certain regions and in limited simulations. It is important to highlight that the complexity to be considered is increased when assessing the AV of RCMs in their historical simulations, as it should include GCM/RCM physical compatibility. We need to consider different GCM/RCM configurations to understand if biases are inherited from GCM through boundary conditions. During winter, RCMs do not noticeably improve the representation of precipitation or surface air temperature. As a matter of fact, they even tend to degrade the winter climatology.

We are aware that the direct comparison between the modeled field and the observed field include some additional sources of error independent of model performance. Thus, model error is associated to model performance only in those regions where ε_R , ε_I , and ε_{obs} are negligible. In general, model error, ε_M , is the largest one over full SA continent, allowing a direct evaluation of climate models. However, we have identified some areas where other sources of error are not at all negligible. ε_R has its highest values over regions with strong surface forcing, such as the Andes Cordillera. ε_I also showed the same behavior. ε_{obs} is also large over complex-terrain regions, especially for surface air temperature. This last result is probably due to the scarcity of surface meteorological stations over isolated mountainous regions and to different interpolation techniques used in each dataset.

There is a general disagreement over literature regarding added value of RCMs for large-scale mean values, as some works find an improvement on representing the mean climatic features (Kerkhoff et al., 2014, Llopart et al., 2014, Nikulin et al., 2012, Torma et al., 2015, Fotso-Nguemo et al., 2017, Veljovic et al., 2010, Prein et al., 2016), while others find a deterioration (Coppola et al., 2014, Feser, 2006, Giorgi et al., 2014, Roads et al., 2003, Mishra et al., 2017). Our results are consistent with this general uncertainty since they suggest that added value of RCMs depends on the different driving fields (i.e., hindcast or historical), the surface properties of the area, the season and the variable analyzed. The dependency of AV to diverse factors was already noted by Torma et al. (2015). The methodology used throughout this paper leaves an important message about comparing RCMs with its coarser-resolution driving data, as the model assessment should not only consider model performance but also the resolution of the model and the uncertainty in observations. As stated in the Introduction, the present work is a first step to assessing the AV of state-of-the-art RCMs over SA. However, to understand the potential added value of RCMs, further investigation should include the analysis of high-frequency temporal scales and small-scale spatial scales.

Acknowledgements This work was supported by the Argentinean agencies CONICET (PIP 11220150100402CO) and ANPCyT (PICT 2014-0887, PICT-2015-3097), and by the French National Program LEFE/INSU (AO2015-876370). The first author was supported by CONICET and by the Saint Exupry Program (MED-MAEDI). The authors would like to thank the two anonymous reviewers for comments that helped improve the manuscript.

References

- Boulanger, J., Carril, A., and Sanchez, E. (2016). CLARIS-La Plata Basin: regional hydroclimate variability, uncertainties and climate change scenarios. *Climate Research*, 68(2-3):93–94.
- Carril, A., Cavalcanti, I., Menéndez, C., 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., and Zamboni, L. (2016). Extreme events in the La Plata basin: a retrospective analysis of what we have learned during CLARIS-LPB project. *Climate Research*, 68(2-3):95–116.
- Carril, A. F., Menéndez, C. G., Remedio, A. R. C., Robledo, F., Sörensson, A., Tencer, B., Boulanger, J.-P., de Castro, M., Jacob, D., Le Treut, H., Li, L. Z. X., Penalba, O., Pfeifer, S., Rusticucci, M., Salio, P., Samuelsson, P., Sanchez, E., and Zaninelli, P. (2012). Performance of a multi-RCM ensemble for South Eastern South America. *Climate Dynamics*, 39(12):2747–2768.

- Chylek, P., Li, J., Dubey, M. K., Wang, M., and Lesins, G. (2011). Observed and model simulated 20th century Arctic temperature variability: Canadian Earth System Model CanESM2. *Atmospheric Chemistry and Physics Discussions*, 11(8):22893–22907.
- Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C. D., Joshi, M., Liddicoat, S., Martin, G., O'connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A., and Woodward, S. (2011). Geoscientific Model Development Development and evaluation of an Earth-System model HadGEM2. *Geosci. Model Dev*, 4:1051–1075.
- Coppola, E., Giorgi, F., Raffaele, F., Fuentes-Franco, R., Giuliani, G., Llopart-Pereira, M., Mamgain, A., Mariotti, L., Diro, G. T., and Torma, C. (2014). Present and future climatologies in the phase I CREMA experiment. *Climatic Change*, 125(1):23–38.
- da Rocha, R. P., Reboita, M. S., Dutra, L. M. M., Llopart, M. P., and Coppola, E. (2014). Inter-annual variability associated with ENSO: present and future climate projections of RegCM4 for South America-CORDEX domain. *Climatic Change*, 125(1):95–109.
- De Haan, L. L., Kanamitsu, M., De Sales, F., and Sun, L. (2014). An evaluation of the seasonal added value of downscaling over the United States using new verification measures. *Theoretical and Applied Climatology*, 122(1-2):47–57.
- De Sales, F. and Xue, Y. (2011). Assessing the dynamic-downscaling ability over South America using the intensity-scale verification technique. *International Journal of Climatology*, 31(8):1205–1221.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656):553–597.
- Denis, B., Côté, J., Laprise, R., and Olume, V. (2002). Spectral Decomposition of Two-Dimensional Atmospheric Fields on Limited-Area Domains Using the Discrete Cosine Transform (DCT). *Monthly Weather Review*, 130(7):1812–1829.
- Di Luca, A., de Elía, R., and Laprise, R. (2012). Potential for added value in precipitation simulated by high-resolution nested Regional Climate Models and observations. *Climate Dynamics*, 38(5-6):1229–1247.
- Di Luca, A., de Elía, R., and Laprise, R. (2015). Challenges in the Quest for Added Value of Regional Climate Dynamical Downscaling. *Current Climate Change Reports*, 1(1):10–21.
- Feser, F. (2006). Enhanced Detectability of Added Value in Limited-Area Model Results Separated into Different Spatial Scales. *Monthly Weather Review*, 134(8):2180–2190.
- Feser, F., Rockel, B., von Storch, H., Winterfeldt, J., and Zahn, M. (2011). Regional Climate Models Add Value to Global Model Data: A Review and Selected Examples. *Bulletin of the American Meteorological Society*, 92(9):1181–1192.
- Feser, F. and von Storch, H. (2005). A Spatial Two-Dimensional Discrete Filter for Limited-Area-Model Evaluation Purposes. *Monthly Weather Review*, 133(6):1774–1786.
- Fotso-Nguemo, T. C., Vondou, D. A., Pokam, W. M., Djomou, Z. Y., Diallo, I., Haensler, A., Tchotchou, L. A. D., Kamsu-Tamo, P. H., Gaye, A. T., and Tchawoua, C. (2017). On the added value of the regional climate model REMO in the assessment of climate change signal over Central Africa. *Climate Dynamics*, 49(11-12):3813–3838.
- Gao, X., Xu, Y., Zhao, Z., Pal, J. S., and Giorgi, F. (2006). On the role of resolution and topography in the simulation of East Asia precipitation. *Theoretical and Applied Climatology*,

- 86(1-4):173–185.
- Garreaud, R. D., Vuille, M., Compagnucci, R., and Marengo, J. (2009). Present-day South American climate. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 281(3-4):180–195.
- Gilleland, E., Ahijevych, D., Brown, B. G., Casati, B., Ebert, E. E., Gilleland, E., Ahijevych, D., Brown, B. G., Casati, B., and Ebert, E. E. (2009). Intercomparison of Spatial Forecast Verification Methods. *Weather and Forecasting*, 24(5):1416–1430.
- Giorgi, F. (2002). Dependence of the surface climate interannual variability on spatial scale. *Geophysical Research Letters*, 29(23):2101.
- Giorgi, F., Coppola, E., Raffaele, F., Diro, G. T., Fuentes-Franco, R., Giuliani, G., Mamgain, A., Llopart, M. P., Mariotti, L., and Torma, C. (2014). Changes in extremes and hydroclimatic regimes in the CREMA ensemble projections. *Climatic Change*, 125(1):39–51.
- Giorgi, F., Coppola, E., Solmon, F., Mariotti, L., Sylla, M., Bi, X., Elguindi, N., Diro, G., Nair, V., Giuliani, G., Turuncoglu, U., Cozzini, S., Güttler, I., OBrien, T., Tawfik, A., Shalaby, A., Zakey, A., Steiner, A., Stordal, F., Sloan, L., and Brankovic, C. (2012). RegCM4: model description and preliminary tests over multiple CORDEX domains. *Climate Research*, 52:7–29.
- Giorgi, F. and Gutowski, W. J. (2015). Regional Dynamical Downscaling and the CORDEX Initiative. *Annual Review of Environment and Resources*, 40(1):467–490.
- Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F. B., and Wood, R. A. (2000). The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics*, 16(2-3):147–168.
- Gulizia, C., Camilloni, I., and Doyle, M. (2013). Identification of the principal patterns of summer moisture transport in South America and their representation by WCRP/CMIP3 global climate models. *Theoretical and Applied Climatology*, 112(1-2):227–241.
- Hazeleger, W., Severijns, C., Semmler, T., Ștefănescu, S., Yang, S., Wang, X., Wyser, K., Dutra, E., Baldasano, J. M., and ... (2010). EC-Earth: A Seamless Earth-System Prediction Approach in Action. *Bulletin of the American Meteorological Society*, 91(10):1357–1363.
- Hourdin, F., Grandpeix, J.-Y., Rio, C., Bony, S., Jam, A., Cheruy, F., Rochetin, N., Fairhead, L., Idelkadi, A., Musat, I., Dufresne, J.-L., Lahellec, A., Lefebvre, M.-P., and Roehrig, R. (2012). LMDZ5B: the atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection. *Climate Dynamics*, 40(9-10):2193–2222.
- Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J.-L., Fairhead, L., Filiberti, M.-A., Friedlingstein, P., Grandpeix, J.-Y., Krinner, G., LeVan, P., Li, Z.-X., and Lott, F. (2006). The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection. *Climate Dynamics*, 27(7-8):787–813.
- IPCC (2013). 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, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia. *Cambridge University Press*, page 1535 pp.
- Jacob, D., Elizalde, A., Haensler, A., Hagemann, S., Kumar, P., Podzun, R., Rechid, D., Remedio, A. R., Saeed, F., Sieck, K., Teichmann, C., and 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.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Jenne, R., and Joseph, D. (1996). The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, 77(3):437–471.

- Kanamitsu, M. and DeHaan, L. (2011). The Added Value Index: A new metric to quantify the added value of regional models. *Journal of Geophysical Research*, 116(D11):D11106.
- Karmacharya, J., New, M., Jones, R., and Levine, R. (2016). Added value of a high-resolution regional climate model in simulation of intraseasonal variability of the South Asian summer monsoon. *International Journal of Climatology*.
- Kerkhoff, C., Künsch, H. R., and Schär, C. (2014). Assessment of Bias Assumptions for Climate Models. *Journal of Climate*, 27(17):6799–6818.
- Kupiainen, M., Jansson, C., Samuelsson, P., and Jones, C. (2014). Rossby Centre regional atmospheric model, RCA4, Rossby Center News Letter.
- Laprise, R. (2008). Regional climate modelling. *Journal of Computational Physics*, 227(7):3641–3666.
- Lenz, C.-J., Früh, B., and Adalatpanah, F. D. (2017). Is there potential added value in COS-MOCLM forced by ERA reanalysis data? *Climate Dynamics*, 49(11-12):4061–4074.
- Llopart, M., Coppola, E., Giorgi, F., da Rocha, R. P., and Cuadra, S. V. (2014). Climate change impact on precipitation for the Amazon and La Plata basins. *Climatic Change*, 125(1):111–125.
- López-Franca, N., Zaninelli, P., Carril, A., Menéndez, C., and Sánchez, E. (2016). Changes in temperature extremes for 21st century scenarios over South America derived from a multi-model ensemble of regional climate models. *Climate Research*, 68(2-3):151–167.
- Lucas-Picher, P., Laprise, R., and Winger, K. (2016). Evidence of added value in North American regional climate model hindcast simulations using ever-increasing horizontal resolutions. *Climate Dynamics*, pages 1–23.
- Marengo, J. a., Ambrizzi, T., da Rocha, R. P., Alves, L. M., Cuadra, S. V., Valverde, M. C., Torres, R. R., Santos, D. C., and Ferraz, S. E. T. (2009). Future change of climate in South America in the late twenty-first century: intercomparison of scenarios from three regional climate models. *Climate Dynamics*, 35(6):1073–1097.
- Meehl, G. A., Covey, C., Taylor, K. E., Delworth, T., Stouffer, R. J., Latif, M., McAvaney, B., Mitchell, J. F. B., Meehl, G. A., Covey, C., Taylor, K. E., Delworth, T., Stouffer, R. J., Latif, M., McAvaney, B., and Mitchell, J. F. B. (2007). THE WCRP CMIP3 Multimodel Dataset: A New Era in Climate Change Research. *Bulletin of the American Meteorological Society*, 88(9):1383–1394.
- Menéndez, C., Zaninelli, P., Carril, A., and Sánchez, E. (2016). Hydrological cycle, temperature, and land surface-atmosphere interaction in the La Plata Basin during summer: response to climate change. *Climate Research*, 68(2-3):231–241.
- Menéndez, C. G., de Castro, M., Sorensson, A., and Boulanger, J. P. (2010). CLARIS Project: towards climate downscaling in South America. *Meteorologische Zeitschrift*, 19(4):357–362.
- Mishra, S. K., Sahany, S., and Salunke, P. (2017). CMIP5 vs. CORDEX over the Indian region: how much do we benefit from dynamical downscaling? *Theoretical and Applied Climatology*, pages 1–9.
- National Geophysical Data Center (2006). 2-minute Gridded Global Relief Data (ETOPO2) v2.
- Nikulin, G., Jones, C., Giorgi, F., Asrar, G., Büchner, M., Cerezo-Mota, R., Christensen, O. B., Déqué, M., Fernandez, J., Hänsler, A., van Meijgaard, E., Samuelsson, P., Sylla, M. B., Sushama, L., Nikulin, G., Jones, C., Giorgi, F., Asrar, G., Büchner, M., Cerezo-Mota, R., Christensen, O. B., Déqué, M., Fernandez, J., Hänsler, A., Meijgaard, E. v., Samuelsson, P., Sylla, M. B., and Sushama, L. (2012). Precipitation Climatology in an Ensemble of CORDEX-Africa Regional Climate Simulations. *Journal of Climate*, 25(18):6057–6078.
- Prein, A. F., Gobiet, A., Truhetz, H., Keuler, K., Goergen, K., Teichmann, C., Fox Maule, C., van Meijgaard, E., Déqué, M., Nikulin, G., Vautard, R., Colette, A., Kjellström, E., and Jacob, D. (2016). Precipitation in the EURO-CORDEX 0.11° and 0.44° simulations: high resolution, high benefits? *Climate Dynamics*, 46(1-2):383–412.

- Prömmel, K., Geyer, B., Jones, J. M., and Widmann, M. (2010). Evaluation of the skill and added value of a reanalysis-driven regional simulation for Alpine temperature. *International Journal of Climatology*, 30(5):760–773.
- Rinke, A., Dethloff, K., Cassano, J. J., Christensen, J. H., Curry, J. A., Du, P., Girard, E., Haugen, J.-E., Jacob, D., Jones, C. G., Költzow, M., Laprise, R., Lynch, A., Pfeifer, S., Serreze, M. C., Shaw, M. J., Tjernström, M., Wyser, K., and Žagar, M. (2006). Evaluation of an ensemble of Arctic regional climate models: spatiotemporal fields during the SHEBA year. *Climate Dynamics*, 26(5):459–472.
- Roads, J., Chen, S., Cocke, S., Druyan, L., Fulakeza, M., LaRow, T., Lonergan, P., Qian, J., and Zebiak, S. (2003). International Research Institute/Applied Research Centers (IRI/ARCs) regional model intercomparison over South America. *Journal of Geophysical Research*, 108(D14):4425.
- Rummukainen, M. (2016). Added value in regional climate modeling. *Wiley Interdisciplinary Reviews: Climate Change*, 7(1):145–159.
- Rusticucci, M., Marengo, J., Penalba, O., and Renom, M. (2010). An intercomparison of model-simulated in extreme rainfall and temperature events during the last half of the twentieth century. Part 1: mean values and variability. *Climatic Change*, 98(3-4):493–508.
- Sanabria, A. L. and Carril, A. F. (2018). Maps of wind hazard over South Eastern South America considering climate change. *Climatic Change*, pages 1–13.
- Sánchez, E., Solman, S., Remedio, A. R. C., Berbery, H., Samuelsson, P., Da Rocha, R. P., Mourão, C., Li, L., Marengo, J., de Castro, M., and Jacob, D. (2015). Regional climate modelling in CLARIS-LPB: a concerted approach towards twentyfirst century projections of regional temperature and precipitation over South America. *Climate Dynamics*, 45(7-8):2193–2212.
- Seth, A., Rauscher, S. a., Camargo, S. J., Qian, J.-H., and Pal, J. S. (2007). RegCM3 regional climatologies for South America using reanalysis and ECHAM global model driving fields. *Climate Dynamics*, 28(5):461–480.
- Silvestri, G. and Vera, C. (2008). Evaluation of the WCRP-CMIP3 model simulations in the La Plata basin. *Meteorological Applications*, 15(4):497–502.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G. (2005). A Description of the Advanced Research WRF Version 2.
- Solman, S. A. (2016). Systematic temperature and precipitation biases in the CLARIS-LPB ensemble simulations over South America and possible implications for climate projections. *Climate Research*, 68(2-3):117–136.
- Solman, S. A., Sanchez, E., Samuelsson, P., da Rocha, R. P., Li, L., Marengo, J., Pessacg, N. L., Remedio, A. R. C., Chou, S. C., Berbery, H., Le Treut, H., de Castro, M., and Jacob, D. (2013). Evaluation of an ensemble of regional climate model simulations over South America driven by the ERA-Interim reanalysis: model performance and uncertainties. *Climate Dynamics*, 41(5-6):1139–1157.
- Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J. J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornblueh, L., Lohmann, U., Pincus, R., Reichler, T., and Roeckner, E. (2013). Atmospheric component of the MPI-M Earth System Model: ECHAM6. *Journal of Advances in Modeling Earth Systems*, 5(2):146–172.
- Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106(D7):7183–7192.
- Taylor, K. E., Stouffer, R. J., Meehl, G. A., Taylor, K. E., Stouffer, R. J., and Meehl, G. A. (2012). An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, 93(4):485–498.

- Torma, C., Giorgi, F., and Coppola, E. (2015). Added value of regional climate modeling over areas characterized by complex terrain-Precipitation over the Alps. *Journal of Geophysical Research: Atmospheres*, 120(9):3957–3972.
- Veljovic, K., Rajkovic, B., Fennessy, M. J., Altshuler, E. L., and Mesinger, F. (2010). Regional climate modeling: Should one attempt improving on the large scales? Lateral boundary condition scheme: Any impact? *Meteorologische Zeitschrift*, 19(3):237–246.
- Wang, J., Swati, F. N. U., Stein, M. L., and Kotamarthi, V. R. (2015). Model performance in spatiotemporal patterns of precipitation: New methods for identifying value added by a regional climate model. *Journal of Geophysical Research: Atmospheres*, 120(4):1239–1259.
- Zaninelli, P. G., Carril, A. F., and Menéndez, C. G. (2015). Explorando temperaturas máximas y mínimas en diferentes reanálisis. Parte 1: campos medios estacionales. *Meteorologica*, 40(1):43–58.
- Zaninelli, P. G., Menéndez, C. G., Falco, M., López-Franca, N., and Carril, A. F. (2018). Future hydroclimatological changes in South America based on an ensemble of regional climate models. *Climate Dynamics*.
- Zazulie, N., Rusticucci, M., and Raga, G. B. (2017). Regional climate of the subtropical central Andes using high-resolution CMIP5 models part I: past performance (1980-2005). *Climate Dynamics*, 49(11-12):3937–3957.