Vol. 57: 201–220, 2013
doi: 10.3354/cr01165

CLIMATE RESEARCH
Clim Res

Published online October 8

FREE ACCESS

Mean fields and interannual variability in RCM simulations over Spain: the ESCENA project

P. Jiménez-Guerrero[1,*], J. P. Montávez[1], M. Domínguez[2], R. Romera[2] , L. Fita[3], J. Fernández[4], W. D. Cabos[4], G. Liguori[4], M. A. Gaertner[5]

[1]Department of Physics, Regional Campus of International Excellence 'Campus Mare Nostrum', University of Murcia, 30100 Murcia, Spain

[2]Environmental Science Institute, Universidad de Castilla la Mancha, 45071 Toledo, Spain

[3]Grupo de Meteorología, Dept. Applied Mathematics and Computer Science, Universidad de Cantabria, 39005 Santander, Spain

[4]Department of Physics, Climate Physics Group, Universidad de Alcalá de Henares, 28805 Madrid, Spain

[5]Environmental Science Faculty, Universidad de Castilla la Mancha, 45071 Toledo, Spain

ABSTRACT: The ESCENA (2008 to 2012) project is a Spanish initiative, which applies the dynamical downscaling technique to generate climate change scenarios based on an ensemble of Regional Climate Models (RCMs) consisting of PROMES, WRF, MM5 or REMO over Peninsular Spain and the Balearic and Canary Islands using a high resolution of 25 km. We describe the mean fields and interannual variability for temperature and precipitation in an ensemble of simulations forced by the high resolution ERA-Interim reanalysis (1990 to 2007) and compare them to the Spain02 observed data set. Maximum surface air temperature shows seasonal cold biases up to −2.5K in all models and it is clearly underestimated during the coldest seasons, but less so during summertime (JJA). Generally, there is a better agreement between observed and simulated minimum surface air temperature, which is slightly overestimated (up to +2K) especially during wintertime (DJF). Regarding precipitation, all models except PROMES tend to show low dry biases during all seasons, especially for autumn on the Mediterranean coast of the Iberian Peninsula. With respect to the interannual variability, the PROMES simulations overestimate the standard deviation of maximum surface air temperature, while the remaining models tend to slightly underestimate it, and most models tend to underestimate the standard deviation of precipitation. The results highlight the ability of these RCMs to reproduce the mean fields and the interannual variability in a very complex terrain such as the Iberian Peninsula, showing a great diversity of climatic behavior. The evaluation of the ensemble results indicates a great improvement in the temporal correlation and the representation of the spatial patterns of temperature and precipitation for all seasons with respect to the individual models.

KEY WORDS: Regional climate models · Spain · Interannual variability · Present climate · Spain02 gridded dataset

1. INTRODUCTION

The climate over regions characterized by complex topographical and land-ocean features exhibits fine scale structure that can be captured only by Regional Climate Models (RCMs) (Gao et al. 2006). Therefore, high-resolution climate information provided by RCMs is required for the assessment of the regional impacts of climate variability and change. This work focuses on the Iberian Peninsula (IP), as an example of a topographically complex region, and characterizes the ability of a new set of RCM simulations in reproducing the observed climate at a regional scale.

*Email: pedro.jimenezguerrero@um.es

The IP presents a large climate heterogeneity because of its position with respect to the North Atlantic circulation and its complex orography. The region exhibits a wide range of precipitation and temperature regimes. This heterogeneity necessitates the use of RCMs to simulate the regional climate details under present, past or future climate change conditions. Thus, the IP is an ideal test area for studying the accuracy of RCMs (see e.g. Jacob et al. 2007, Gómez-Navarro et al. 2010). There is also a great deal of RCM sensitivity experiments over the area. For instance, Fernández et al. (2007) showed the sensitivity of the climatology (mean temperature and precipitation) to the changes in the model physics. More recent studies have shown the influence of land-surface model choice on the reproduction of observed climate (Jerez et al. 2010). Most of the previous studies considered a single RCM. The relative performance of RCMs over the region has been compared only recently (e.g. Herrera et al. 2010).

The climate simulated by an RCM suffers from uncertainties arising from a variety of sources such as internal variability, different model formulations, and imperfections in the boundary conditions. Such uncertainties can be explored by running an ensemble of simulations, varying the source of the uncertainty. For example, the uncertainty associated with imperfections in the model formulation can be addressed by running a multi-model ensemble. Climate signals common to all models are usually given more confidence. This reasoning relies on the independence of the errors of the different models, even though this is not fully justified due to the common building blocks of the diverse models (Fernández et al. 2009, Knutti et al. 2010).

A recent series of EU-funded projects exploited this multi-model ensemble approach to provide an estimation of the uncertainties of regional climate change over Europe. The latest projects in this series were PRUDENCE (Christensen & Christensen 2007) and ENSEMBLES (Van der Linden & Mitchell 2009). In PRUDENCE (2001 to 2004), 2 SRES emission scenarios (A2 and B2) were analyzed for the last 3 decades of this century (2071 to 2100) with an ensemble of 11 RCMs at a spatial resolution of 50 km. Most of the simulations were forced with the same General Circulation Model (GCM) (HadCM3). The main objective of the subsequent ENSEMBLES project (2004 to 2009) was to generate an objective probabilistic estimate of uncertainty in future climate scenarios. Several combinations of 13 RCMs and 7 GCMs were applied to the SRES emissions scenarios A1B and A2, with a finer horizontal resolution (25 km). All RCMs simulated a common period extending over the first decades of the century, up to 2050. Additionally, ENSEMBLES included RCM evaluation simulations nested into 'perfect' boundary conditions taken from the ERA-40 Reanalysis (Uppala et al. 2005).

Unfortunately, the Spanish territory is not well covered in these pan-European multi-model experiments. For this reason, the Spanish government funded a strategic action to generate high-resolution downscaled scenarios over Spain, complementing those produced within ENSEMBLES. The ESCENA project is in charge of generating downscaled scenarios using nested RCMs, including perfect-boundary evaluation simulations, which are analyzed in this study. The simulations produced within ESCENA and, in particular, those described in this work, are publicly available at http://proyectoescena.uclm.es.

Compared to ENSEMBLES, ESCENA uses improved (PROMES) or additional RCMs (MM5, WRF), new GCM/RCM combinations, and a larger set of emission scenarios (A1B, A2 and B1), with the target domain centered over the IP and covering parts of the Atlantic Ocean, including the Canary Islands. One of the RCMs (WRF) has been run with 2 different configurations of the boundary layer scheme. Additionally, the evaluation runs in ESCENA are forced by the higher resolution ERA-Interim reanalysis, instead of the ERA-40 reanalysis used in ENSEMBLES.

Several previous studies analyzed the performance of an ensemble of RCM simulations forced by perfect boundaries and focused on the IP or included the IP in their analyses. The main source for those analyses has been the ENSEMBLES RCM database. In particular, Herrera et al. (2010) validated the mean and extreme precipitation regimes simulated by the ENSEMBLES RCMs over the IP. They used the Spain02 gridded observational data set (Herrera et al. 2012), which is also used in our study. Our work extends that of Herrera et al. (2010) by including the analysis of maximum and minimum temperature and considering the ability of reproducing the interannual variability. We also performed a seasonal analysis, given that the climate in the IP exhibits a marked intra-anual variability.

A number of references provide a measure of current state-of-the-art RCM seasonal biases. For instance, Christensen et al. (2010) explored different metrics in order to weight the RCM simulations from ENSEMBLES. They showed seasonal temperature and precipitation climatologies all over Europe, but did not find compelling evidence of an improved description of mean climate states using performance-based weights in comparison to the use of equal

weights. Coppola et al. (2010) indicate a wide variability in the performance across models included in an ensemble of ENSEMBLES-project simulations for the European region, mostly deriving from the magnitude/sign of precipitation-based functional metrics. The authors indicate that weighting leads to an overall improvement of the performance of the ensemble especially over topographically complex regions. Lastly, Kjellström et al. (2011) use an ensemble of 16 regional climate model simulations to indicate that the spread of the results and the biases in the 1961–1990 period are strongly related to the representation of the large-scale circulation in the GCMs.

Thus, the objectives of the present study were to (1) characterize the ability of a multi-model ensemble of RCMs to reproduce the observed regional climate over peninsular Spain and the Balearic Islands; (2) quantify the performance of the RCMs, focusing on the mean seasonal climate. not only on the climatology, but particularly on the interannual variability, which usually receives little attention (Giorgi et al. 2004); (3) compare the performance of the ensemble mean with that of the individual models; and (4) compare the model-to-model variability with the intra-model variability induced by the use of a local and a non-local planetary boundary layer (PBL) scheme. The present work focuses on the mean climate and interannual variability, while a companion paper by Domínguez et al. (2013) characterizes the ability of this new set of model simulations in reproducing the extreme climatic situations over the region.

2. METHODOLOGY AND DATA

2.1. RCM description

This study involved 4 different RCMs: PROMES, WRF, MM5 and REMO. The WRF model was run with 2 different physical parameterization configurations in order to compare the model-to-model variability with the variability induced by changing the physical scheme within the same RCM. All simulations cover a present-day 18 yr period (1990 to 2007) and were driven by ECMWF/ERA-Interim reanalysis (Section 2.2). The simulations cover most of Europe, with domains centered in the IP (Fig. 1). Some model domains (PROMES, WRF and REMO) were rotated in order to include the area of the Canary Islands (for the ESCENA project) with a domain smaller than the unrotated one used by MM5. The different domain sizes and locations may bring additional uncertainty to the model results. All models were run with a horizontal resolution of ~25 km and set the top vertical layer at 10 hPa using a different number of vertical levels (see Table 1). A lateral boundary relaxation zone of 5 to 10 additional grid points was used by the different models. A brief description of each model configuration is provided in the following sections and is summarized in Table 1.

2.1.1. PROMES

The regional climate model PROMES (Castro et al. 1993) has been developed by MOMAC (Modelización para el Medio Ambiente y el Clima) research group at the Complutense University of Madrid (UCM) and the University of Castilla-La Mancha (UCLM). PROMES has been applied over many different simulation domains: IP (Gaertner et al. 2001, Arribas et al. 2003), Europe (Gaertner et al. 2007, Sánchez et al. 2009), West Africa (Domínguez et al. 2010, Gaertner et al. 2010) and South America (Boulanger et al. 2010). The PROMES version used in this work includes several improvements to the physical parameterizations which have been introduced during the last years. The resolved-scale cloud formation and its associated precipitation processes are modeled according to Hong et al. (2004). This scheme includes ice microphysics processes. The sub-grid scale convective clouds and their precipitation are simulated with the Kain-Fritsch parameterization (Kain & Fritsch 1993). The shortwave

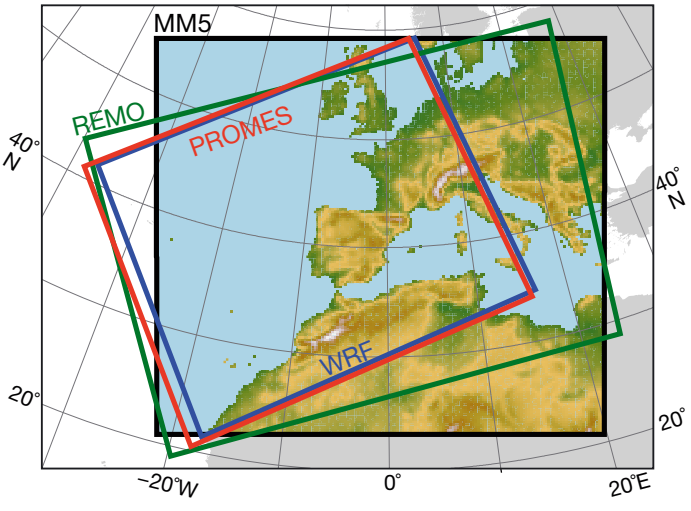

Fig. 1. Simulation domains used in the models used in the ESCENA project. Color shades: topography of domain. Relaxation zone, where the regional climate model output is relaxed towards the reanalysis, is excluded

Table 1. Configurations used in each of the models run in this study. Columns: model name, geographical projection (Geo. proj), number of vertical levels (Vert. level), horizontal resolution (Horiz. resol.), and physical parameterizations for convective clouds and precipitation (Cumulus), resolved cloud microphysics, radiation, planetary boundary layer (PBL) and land surface models (LSM). See Section 2.1 for details; Fig. 1 shows domain covered by each model

Model	Geo. proj.	Vert. level	Horiz. resol.	Physics parameterizations				
				Microphysics	Cumulus	Radiation	PBL	LSM
PROMES	Lambert	37	25 km	Includes ice processes (Hong et al. 2004)	Kain & Fritsch (1993)	ECMWF (2004) with fractional cloud cover (Chaboureau & Bechtold 2002, 2005)	Cuxart et al. (2000)	ORCHIDEE LSM (Krinner et al. 2005)
WRF-A	Lambert	33	25 km	WSM5	GD	CAM	MYJ	Noah
WRF-B	Lambert	33	25 km	WSM5	GD	CAM	YSU	Noah
MM5	Lambert	30	25 km	Simple ice	Grell	RRTM	MRF	Noah
REMO	Rotated lat-lon	31	0.22°	Roeckner et al. (1996)	Tiedkte (1989) with modifications after Nordeng (1994)	Morcrette et al. (1986) with modifications after Giorgetta & Wild (1995)	Higher-order closure scheme (Brinkop & Roeckner 1995)	Bucket scheme for hydrology; 5 layers for thermal processes

and longwave radiation parameterizations described in ECMWF (2004) are used, and the clouds-radiation interaction is simulated with a fractional cloud cover parameterization (Chaboureau & Bechtold 2002, 2005). The turbulent vertical exchange in the PBL is modeled following Cuxart et al. (2000). PROMES has been coupled to the land surface model ORCHIDEE (Organizing Carbon and Hydrology in Dynamic EcosystEms, Krinner et al. 2005) with the aim of improving the land surface-atmosphere coupling. A contour band of 10 points is used to relax the model variables following Davies (1976). The method for the vertical interpolation of the large scale variables to model levels is described in Gaertner & Castro (1996).

2.1.2. WRF

The Weather Research and Forecasting (WRF) model is a state-of-the-art limited area model developed in a collaboration between the National Center for Atmospheric Research (NCAR; Boulder, CO, USA) and a number of research institutions in the United States. In this study we used the Advanced Research WRF (ARW) core (version 3.1.1), which is the research version of the model and incorporates the latest advances in the physics schemes. The ARW solver (Klemp et al. 2007, Skamarock et al. 2008) integrates the non-hydrostatic fully compressible Euler equations in flux form, using a conservative scheme over a staggered (Arakawa-C) horizontal grid and a vertical mass coordinate. The WRF model simulations were run by the Santander Meteorology Group (SMG) of the University of Cantabria (Fita et al. 2010) through the WRF4G execution workflow (Fernández-Quiruelas et al. 2010), and represent the first WRF simulations available over the IP, together with the work of Argüeso et al. (2012). The main physical schemes used were the Grell & Devenyi (2002; GD) cumulus ensemble scheme, the WRF single-moment 5-class microphysics (WSM5; Hong & Lin 2006) and the Community Atmospheric Model (CAM) 3.0 radiation scheme (Collins et al. 2006). The simulations used 2 different configurations (labeled WRF-A and WRF-B in Table 1), which only differ in the PBL scheme. WRF-A uses a local scheme (Mellor-Yamada-Janjic; MYJ), whereas WRF-B uses a non-local scheme (Yonsei University; YSU). The non-local scheme treats vertical mixing of the entire PBL, meanwhile in the local scheme, vertical mixing is computed successively between contiguous vertical layers. As with the MM5 model, the Noah LSM was used to solve the soil processes on 4 layers to a depth of 2 m (Chen & Dudhia 2001a,b).

2.1.3. MM5

MM5 consists of a climatic version developed at the University of Murcia (Gómez-Navarro et al. 2011, Jerez et al. 2012a, 2013) of the Fifth-Generation Pennsylvania State University-NCAR Mesoscale Model (Dudhia 1993, Grell et al. 1994). MM5 has been extensively used in a number of regional cli-

mate simulations under different configurations (e.g. Boo et al. 2006, Nunez et al. 2009, Gómez-Navarro et al. 2010, among others). The physical configuration has been chosen in order to minimize the computational cost, since none of the tested configurations provides the best performance for all kinds of synoptic events and regions (Fernández et al. 2007, Jerez et al. 2012a). The physical options used were Grell cumulus parameterization (Grell 1993), Simple Ice for microphysics (Dudhia 1989), rapid radiative transfer model (RRTM) radiation scheme (Mlawer et al. 1997) and medium-range forecast scheme (MRF) for planetary boundary layer (Hong & Pan 1996). The Noah Land-Surface model (Chen & Dudhia 2001a,b) has been used in summer over the southern part of the IP, as it simulates the climate more accurately, especially in dry areas (Jerez et al. 2012b).

2.1.4. REMO

REMO is a hydrostatic, three-dimensional regional climate atmospheric model developed at the Max-Planck-Institute for Meteorology in Hamburg. It is based on the Europa Model, a former numerical weather prediction model of the German Weather Service and is described in Jacob et al. (2001). REMO uses the physical package of the global circulation model ECHAM4 (Roeckner et al. 1996). In the vertical, variations of the prognostic variables are represented by a hybrid vertical coordinate system (Simmons & Burridge 1981). The relaxation scheme according to Davies (1976) is applied.

2.2. Driving data

All the simulations were forced at the boundaries by the latest reanalysis product from the ECMWF: ERA-Interim (Dee et al. 2011). ERA-Interim has several differences with respect to ERA-40 (Uppala et al. 2005), such as variational bias correction of satellite radiance data, improvements in model physics and new humidity analysis, among others. The ERA-Interim atmospheric model and reanalysis system uses cycle 31r2 of ECMWF's Integrated Forecast System (IFS), which was introduced operationally in September 2006, configured for the following spatial resolution: (1) 60 levels in the vertical, with the top level at 0.1 hPa; (2) T255 spherical-harmonic representation for the basic dynamical fields; (3) a reduced Gaussian grid with approximately uniform 79 km spacing for surface and other grid-point fields. In the study, all models have used boundary and initial conditions from ERA-Interim every 6 h with a spatial resolution of 0.7° × 0.7°.

With respect to the forcings along the ERA-Interim period, all models use a sea surface temperature interpolated from ERA-Interim data and updated every 6 h. Aerosols in PROMES follow the GADS climatology (d'Almeida et al. 1991, Koepke et al. 1997), while the rest of the models do not include aerosols in their forcing for regional climate simulations. For greenhouse gases, REMO uses trends in the amounts of specified radiatively active gases (CO_2, CH_4, N_2O, CFC-11, CFC-12) according to the ones used to generate the ERA reanalysis (Houghton et al. 1996). PROMES, WRF and MM5 use constant values of greenhouse gases.

2.3. Observational database

We used the Spain02 precipitation and temperature database (Herrera et al. 2012). Spain02 is a daily gridded dataset developed using surface station data from a set of 2756 quality-controlled stations over peninsular Spain and the Balearic islands. This data set covers the period 1950–2008, with a daily frequency and 0.2° × 0.2° resolution. A 2-step interpolation procedure was used: firstly, the monthly means were interpolated using thin-plate splines, and secondly, daily departures from the monthly means were interpolated using a kriging methodology. In the case of precipitation, occurrence and amounts are interpolated through indicator and ordinary kriging, respectively. The interpolation procedure for precipitation is similar to that of the E-OBS European database (Haylock et al. 2008), except for the absence of the elevation-dependent splines applied to the E-OBS database, which was not used in Spain02 since the topography is well represented by the large amount of stations. For temperature, the Spain02 database applied a more stringent filter to the station data in order to provide a product better suited for trend analyses (Herrera et al. 2012). Unlike the precipitation product, where the maximum number of available stations at a given day were used, the temperature data set was built from a reduced set of only those stations with a record longer than 40 yr and <1% of missing data, leaving 186 stations. These stations were gridded using the same methodology as E-OBS, the only difference being the better quality and higher number of stations over spain.

The large number of stations used in this product compared with the ECA database provides a better

representation of precipitation (Herrera et al. 2010) and temperature variability (Herrera et al. 2012) over Spain.

2.4. Validation methodology

The Spain02 regular latitude-longitude grid has been used as the reference for validation. Thus, all RCM data have been bilinearly interpolated onto the Spain02 grid. Since the resolution of the RCMs is similar to that of Spain02 and we are interested in the mean climate, the interpolation procedure is not expected to significantly alter any of our results.

All the statistical measures are calculated at individual grid points. Only land grid points over peninsular Spain and the Balearic islands are considered in the analysis. Since this work focuses on mean climate, we only worked with monthly averaged data. Thus, we will use the notation V^k_{iym} for a variable from model k at grid point i, in year $y = 1990, \ldots, 2007$ and month $m = 1, \ldots, 12$. If we use bracket notation for an average over a given index (e.g. $\langle\cdot\rangle_{ym}$ for an average over all years and months, i.e. a time average), we can express the bias (b) at a given grid point as:

$$b^k_i = \langle V^k_{iym} - O_{iym}\rangle_{ym} \quad (1)$$

where O_{iym} is the value observed. The model bias is the simplest measure of model performance.

The ensemble mean, $\langle V^k_{iym}\rangle_k$, is usually considered as an additional simulation that compensates the errors of the different ensemble members. Even though this is a very simplistic view of the ensemble (which should be considered from a probabilistic point of view), it can be useful to reinforce the common signal of the different models in our analysis of the mean climate. Notice, however, that the ensemble mean is not a physical realization of any of the models, but just a statistical average of different nonlinear trajectories (Knutti et al. 2010).

For seasonal analyses, the seasons were defined as winter (DJF), spring (MAM), summer (JJA) and autumn (SON). Seasonal biases can be defined by averaging over months in a specific season, e.g. $b^k_{i,\mathrm{DJF}} = \langle V^k_{iym} - O_{iym}\rangle_{y,m=\mathrm{DJF}}$.

The seasonal cycle was calculated as follows:

$$V^k_{im} = \langle V^k_{iym}\rangle_y \quad (2)$$

Then, the interannual variability was assessed on the series without considering their annual cycle ($\hat{V}^k_{iym}$). The averaged monthly annual cycle was removed from each corresponding monthly value:

$$\hat{V}^k_{iym} = V^k_{iym} - V^k_{im} \quad (3)$$

The ability to represent the interannual variability can be decomposed into (1) the ability to represent its size, which can be represented by the standard deviation (SD) of the deseasonalized series as:

$$\mathrm{SD}[V]^k_i = \sqrt{\langle\left(\hat{V}^k_{iym}\right)^2\rangle_{ym}} \quad (4)$$

and (2) can be compared to that of the observations $\mathrm{SD}[O]_i$, and (3) the ability to represent the year-to-year variations, which can be represented by the linear correlation coefficient with the observations:

$$\rho^k_i = \frac{\langle \hat{V}^k_{iym}\hat{O}_{iym}\rangle_{ym}}{\sqrt{\langle\left(\hat{V}^k_{iym}\right)^2\rangle_{ym}\langle\left(\hat{O}_{iym}\right)^2\rangle_{ym}}} \quad (5)$$

The latter ability can only be expected on RCM simulations nested into 'perfect' boundary conditions such as those considered in this study.

Finally, pattern agreement between simulated and observed seasonal climatologies was quantified by means of the spatial correlation and the ratio between simulated and observed SDs, $V^k_i = \langle V^k_{iym}\rangle_{ym}$, as:

$$r^k = \frac{\langle (V^k_i - \langle V^k_i\rangle_i)(O_i - \langle O_i\rangle_i)\rangle_i}{\sqrt{\langle (V^k_i - \langle V^k_i\rangle_i)^2\rangle_i \langle (O_i - \langle O_i\rangle_i)^2\rangle_i}} \quad (6)$$

$$s^k = \sqrt{\frac{\langle (V^k_i - \langle V^k_i\rangle_i)^2\rangle_i}{\langle (O_i - \langle O_i\rangle_i)^2\rangle_i}} \quad (7)$$

This information can be summarized in a Taylor (2001) diagram, which is a polar plot, with radial coordinate s^k and angular coordinate related to r^k.

3. RESULTS

3.1. Bias

Seasonal model biases in representing the climatology of maximum and minimum temperature and precipitation are represented in Figs. 2 & 3 and summarized in Table 2a by their spatial averages, along with the average annual bias. The ensemble average is also displayed as if it were an additional simulation.

Annual maximum temperatures are cold-biased (Table 2a) in most models (ranging from –2.43K in MM5 to –1.74K in PROMES). REMO is the only exception and shows a warm bias (+0.64K). The models keep in all seasons the same bias sign as the annual values: REMO shows warm biases in all seasons and the other models show cold biases. The models show similar bias spatial patterns (Fig. 2). For

Table 2. Spatial average of (a) the mean bias and (b) interannual variability for maximum (T_{max}) and minimum (T_{min}) temperature and precipitation in the models included in ESCENA and the ensemble of simulations. Data in (b): SD for temperature, CV for precipitation. Seasons: winter (DJF), spring (MAM), summer (JJA), autumn (SON)

	T_{max} (K)				T_{min} (K)				Precipitation (mm d^{-1})				
	DJF	MAM	JJA	SON	Annual	DJF	MAM	JJA	SON	ANNUAL	DJF	MAM	JJA	SON	Annual
(a)															
PROMES	−2.08	−2.07	−0.62	−2.20	−1.70	0.51	0.06	0.27	0.53	0.34	0.30	0.65	0.52	0.38	0.44
WRF-A	−2.48	−2.33	−2.31	−2.49	−2.40	0.03	−0.42	−0.59	−0.45	−0.37	−0.48	−0.25	0.01	−0.80	−0.38
WRF-B	−1.80	−2.27	−2.44	−2.00	−2.13	1.25	0.67	0.60	0.65	0.79	−0.29	−0.17	−0.07	−0.71	−0.31
MM5	−1.72	−2.59	−2.99	−2.44	−2.43	0.79	−0.21	−0.15	0.09	0.13	−0.07	−0.13	0.00	−0.22	−0.10
REMO	0.30	0.26	1.01	0.98	0.64	2.34	2.10	3.21	2.87	2.63	−0.17	0.21	0.14	−0.24	−0.25
ENSEMBLE	−1.56	−1.80	−1.47	−1.63	−1.61	0.98	0.44	0.67	0.74	0.71	−0.14	0.06	0.12	−0.32	−0.08
(b)															
Spain02	1.64	1.86	1.86	1.80	1.81	1.83	1.26	1.20	1.59	1.50	0.92	0.74	1.07	0.74	0.87
PROMES	−0.07	0.33	0.32	0.22	0.21	−0.09	0.11	−0.04	0.07	0.00	0.04	0.16	0.28	0.04	0.29
WRF-A	0.07	−0.00	−0.02	0.04	0.02	−0.33	−0.04	0.06	−0.12	−0.13	−0.23	−0.18	−0.26	−0.28	−0.21
WRF-B	−0.11	−0.16	−0.12	−0.06	−0.12	−0.36	−0.18	0.01	−0.20	−0.20	−0.19	−0.14	−0.34	−0.23	−0.17
MM5	−0.27	−0.38	−0.20	−0.12	−0.24	−0.38	−0.16	−0.02	−0.01	−0.16	−0.06	−0.08	−0.15	−0.07	0.00
REMO	−0.08	−0.15	−0.17	−0.01	−0.10	−0.60	−0.21	0.08	−0.19	−0.25	−0.05	0.03	−0.07	−0.05	0.02
ENSEMBLE	−0.09	−0.07	−0.04	−0.01	−0.05	−0.35	−0.10	0.02	−0.09	−0.15	−0.09	−0.04	−0.10	−0.11	−0.01

instance, all models overestimate maximum temperature during summertime over the southern IP and some areas along the Mediterranean coast.

For minimum temperature (Table 2a), REMO is again strongly warm biased (+2.63K) compared with the other models, which range from −0.37 (WRF-A) to +0.79K (WRF-B). Two configurations of the same model provide a wide range of minimum temperatures. Regarding minimum temperature, the difference between different PBL parameterizations in a single model is larger than between completely different models in the ensemble (e.g. MM5 and PROMES), except in the case of REMO. Again, the spatial patterns (Fig. 2) of the different models are very similar. The topographical characteristics of minimum temperatures are well captured by the models, but they all show overestimations over mountainous areas and slight underestimations elsewhere. Note that these tend to partly compensate each other in the spatially averaged biases shown in Table 2a.

For precipitation, only those cells over land with observed precipitation >0.25 mm d^{-1} (monthly average) have been selected for comparisons. Annual biases (Table 2a) range from −0.38 (WRF-A) to +0.44 mm d^{-1} (PROMES). In this case, the spatial pattern of seasonal biases (Fig. 3) are presented as a relative difference (in %), since larger absolute biases tend to occur over areas with larger precipitation. The threshold chosen (0.25 mm d^{-1}) is especially noticeable for summertime in the south and southeastern Mediterranean coast. In autumn, most models show a large precipitation underestimation in southern and eastern peninsular Spain and the Balearic Islands, where torrential precipitation occurs. PROMES shows a different behaviour compared with the rest of the models, since it overestimates seasonal precipitation, especially during spring.

Table 2a and Figs. 2 & 3 include the biases of the ensemble mean, considered as an additional model simulation. As expected, the ensemble mean shows an intermediate behaviour, which in most cases leads to reduced biases. However, common biases, such as the cold-biased maximum temperature or the common pattern in minimum temperature, remain in the ensemble mean, although the most extreme biases are attenuated. For precipitation, the ensemble mean shows the smallest absolute bias compared to the individual simulations.

An especially remarkable result is the high spread of the simulated bias, despite the small amount of models used in this study. It is crucial to work with an

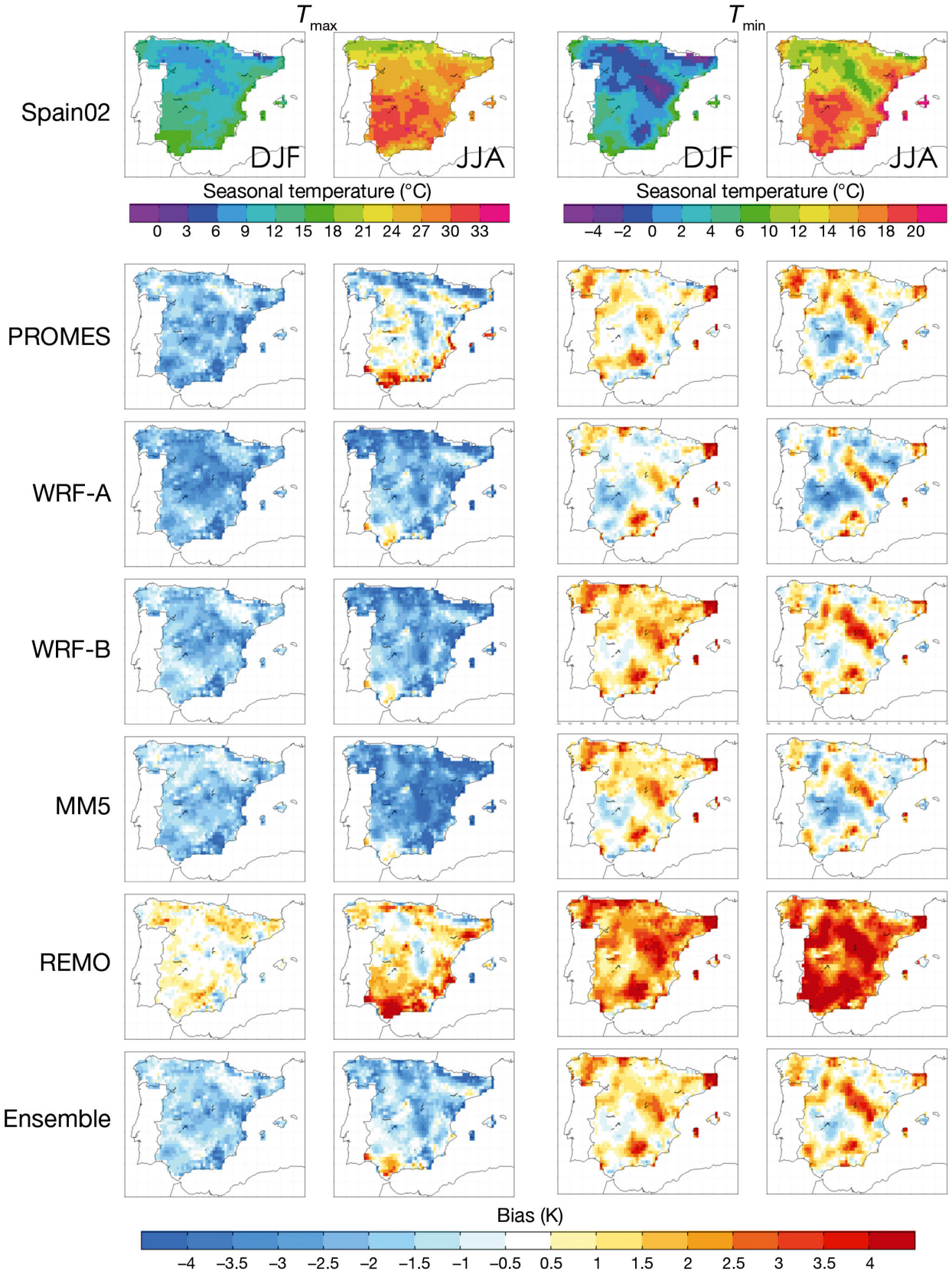

Fig. 2. Seasonal maximum (T_{max}) and minimum (T_{min}) temperature (°C; top row) for Spain02 and biases (K) for models. Columns: winter (DJF) and summer (JJA) values

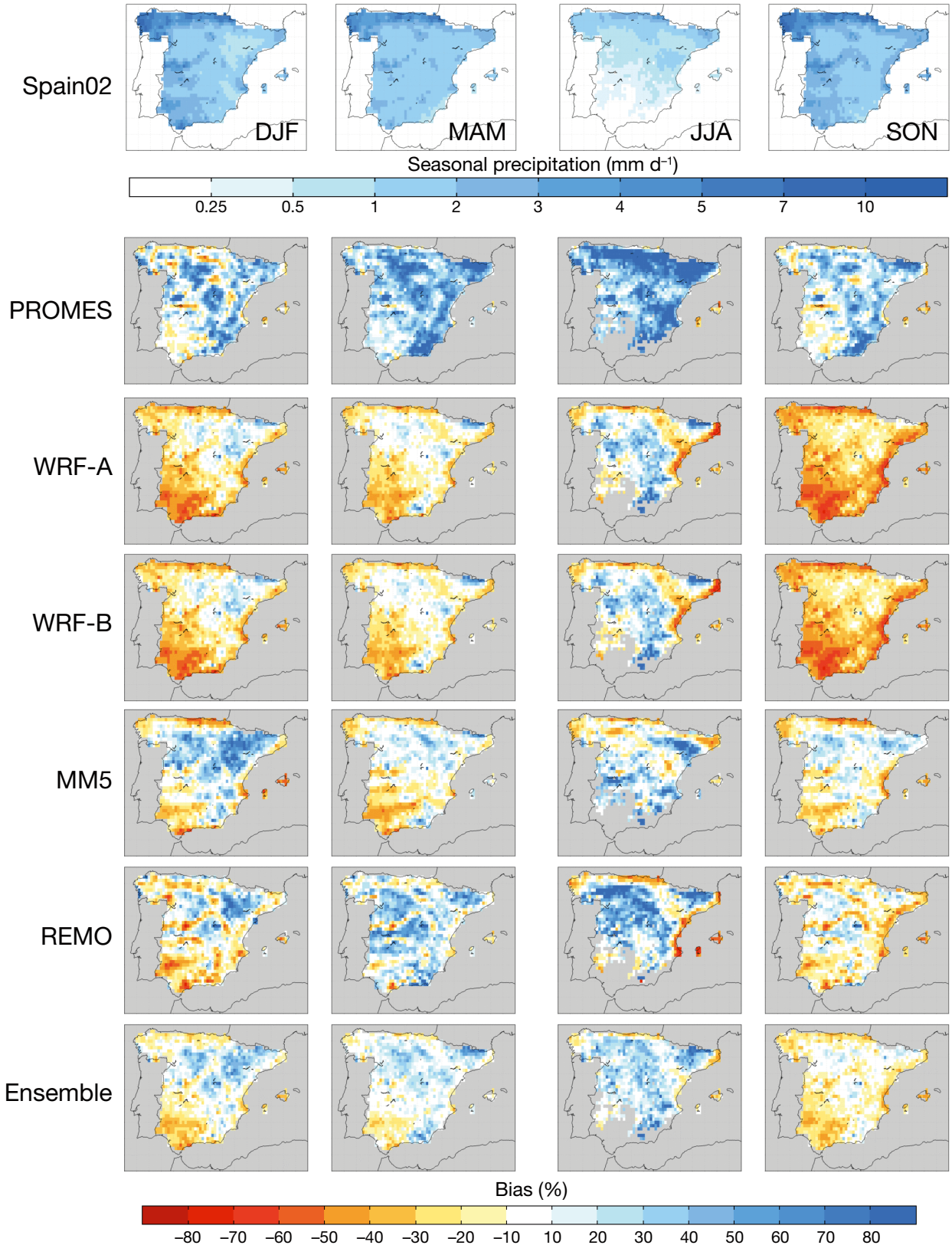

Fig. 3. Seasonal precipitation for Spain02 (mm d^{-1}, top row) and biases (%) for models. Columns (left to right): winter (DJF), spring (MAM), summer (JJA) and autumn (SON) values. Grey shading: areas where biases are not shown (ocean and precipitation < 0.25 mm d^{-1})

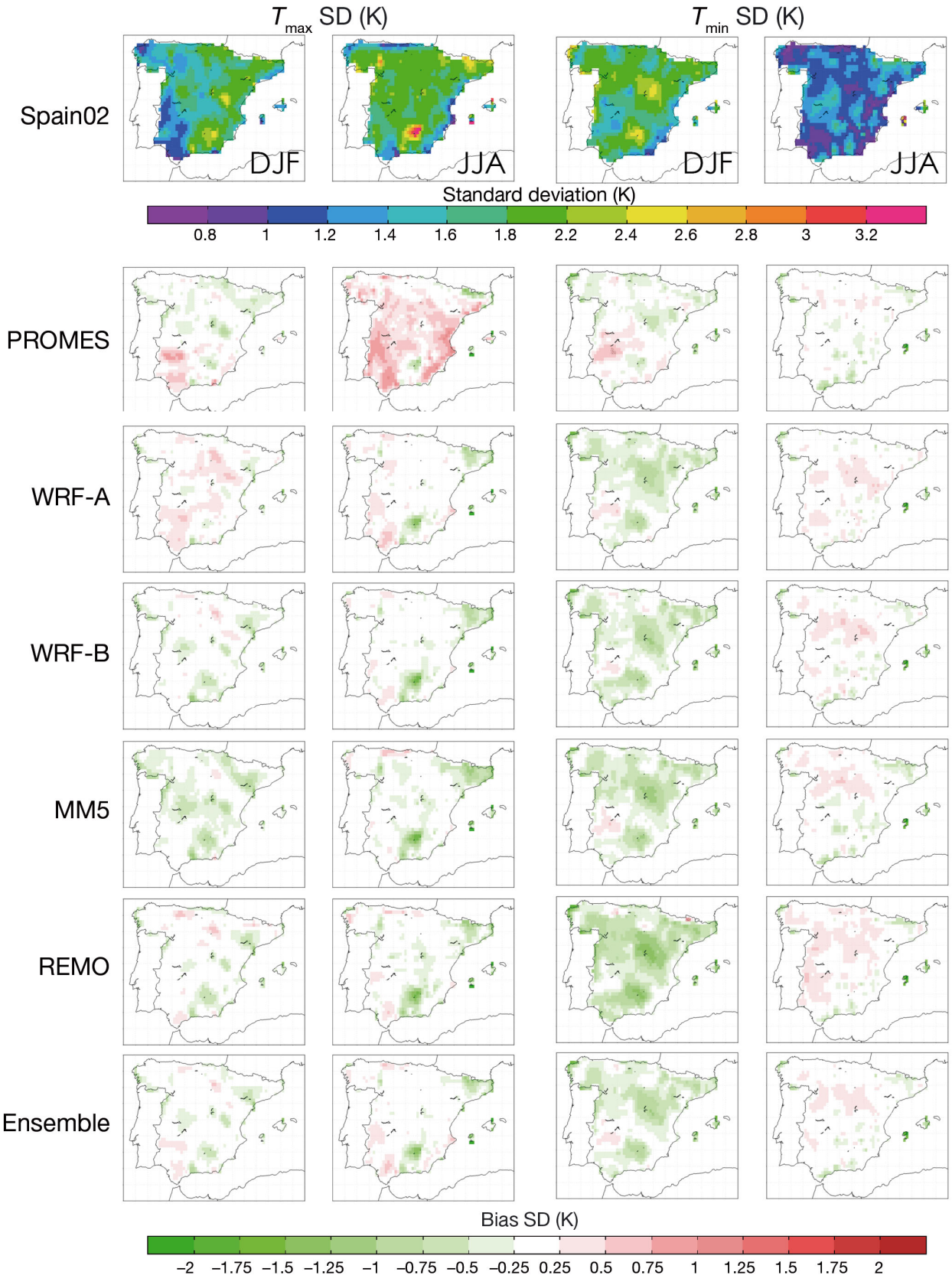

Fig. 4. Standard deviation (SD) of seasonal maximum (T_{max}) and minimum (T_{min}) temperature for Spain02 (K; top row) and SD biases (K) for models. Columns: winter (DJF) and summer (JJA) values

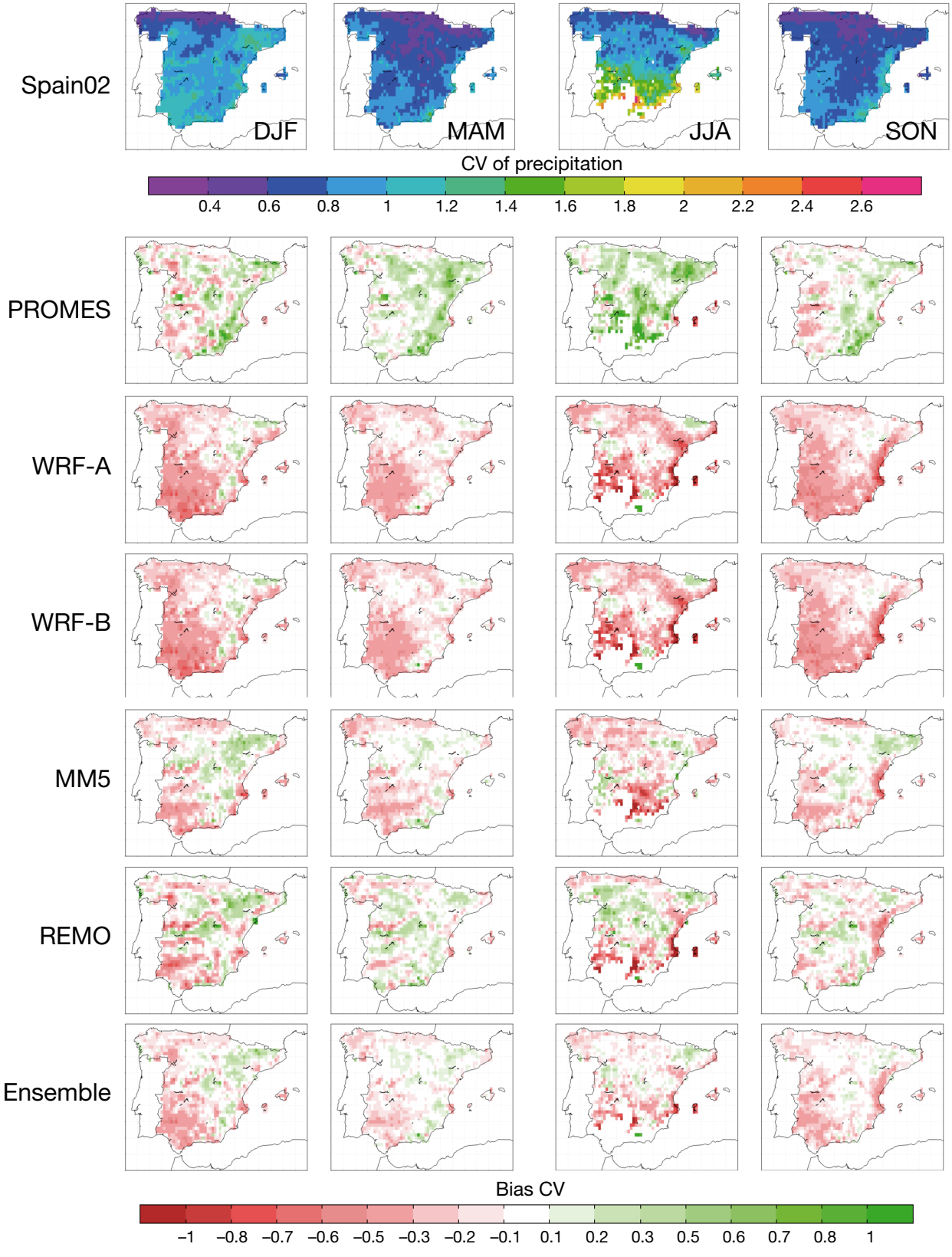

Fig. 5. Coefficient of variation (CV) of seasonal precipitation for Spain02 and CV biases for models. Columns: same order as in Fig. 3

ensemble of regional climate models instead of one individual model to get more robust future climate projections, for a better assessment of uncertainties due to the dynamical downscaling process. These climate projections have already been carried out in the ESCENA project. An important issue, which we do not address here but will be considered in future stages of the ESCENA project, is the relationship between past and future performance of models (as evaluated by Whetton et al. 2007 and Abe et al. 2009, among others). As stated by Annan & Hargreaves (2010), climate models have already been tuned to some extent to the recent climate data and therefore accurately reproduce present-day climate conditions. It would be interesting to further test the reliability of the ensemble in other ways, for example considering simulations of other epochs or other climatic observations that are less widely used during model construction and tuning.

3.2. Interannual variability

For precipitation, SD is normalized by the period average, therefore by obtaining a coefficient of variation (CV). This is done because the SD of precipitation is typically related to the mean (Giorgi et al. 2004), so that the CV is a more independent measure of interannual variability.

Table 2b shows the spatially averaged seasonal and annual interannual variability biases for maximum and minimum temperature and the spatially averaged CV biases for precipitation. The seasonal spatial distribution of the mentioned statistics are represented in Figs. 4 & 5, for maximum and minimum temperature and precipitation, respectively.

All models except for PROMES and WRF-A tend to slightly underestimate the interannual SD of maximum temperature (Fig. 4). Similar values and responses from models are observed for different seasons. In winter, the biases of maximum temperature SD vary between −0.27K in MM5 to +0.07K in WRF-A (Table 2b). The biases range from −0.38 (MM5) to +0.33K (PROMES) in spring, and −0.20 (MM5) to +0.32K (PROMES) during summer; meanwhile for autumn MM5 once more provides the largest underestimation of the SD (−0.12K) and PROMES the highest overestimation (+0.22K). The rest of the models show an intermediate behavior for representing the interannual SD, with the best skills during the autumn season.

For minimum temperature (Fig. 4), the results are similar to maximum temperature, but in this case it is REMO which tends to provide the largest biases in the minimum temperature variability (Table 2b). The SD is generally underestimated by all models, especially during winter over the highest mountain chains of the IP. In this season, biases vary between −0.60 (REMO) and −0.09K (PROMES). All models predict spring variability better than for wintertime (−0.21K in REMO and 0.11K in PROMES during springtime). However, in summer all the models show values close to Spain02.

Regarding precipitation, the models (except PROMES) pervasively underestimate the interannual variability over the domain, especially in southwestern Spain in winter, and the Mediterranean coast and the Balearic Islands during summer and autumn (Fig. 5). The biases of the CV range from +0.04 (winter and autumn) to +0.28 (summer) in the case of PROMES, to underestimations in WRF-A (−0.28 in autumn to −0.18 in spring). MM5 and REMO present the smallest average biases, ranging from −0.15 to 0.03.

The multi-model ensemble mean outperforms most of the individual models in all seasons, showing the added value of using an ensemble of RCMs to improve the amount of interannual variability. However, common model deficiencies still remain (e.g. the low precipitation variability in autumn on the Mediterranean coast).

The temporal correlation (ρ) between simulated and observed deseasonalized series is shown in Fig. 6 for maximum and minimum temperature and precipitation. Maximum temperatures show the highest correlations (>0.85) over most of the domain, especially over the river valleys of northern Spain. The fact that the lowest correlations (~0.4) are found in northeastern Spain for all models may point to (1) problems with the quality of raw observed data used to create the Spain02 database or (2) to local climatic features that the models are unable to capture at the selected resolution. For minimum temperature, ρ is generally slightly lower than for maxima. The spatial structure of the correlation presents common minima in all models, and this points to a relation with topography, since the lowest correlation coefficients ($\rho < 0.5$) are located in the highest mountain chains in Spain, such as the Pyrenees, Sistema Ibérico or Sierra Nevada, from north to south of the domain. On the other hand, the highest correlations for minimum temperatures are observed over the Central Plateaus and the river valleys. The precipitation correlations show a strong east-west gradient reaching the lowest values (~0.30 to 0.50) towards the Ebro Valley and the southeastern IP (Fig. 6).

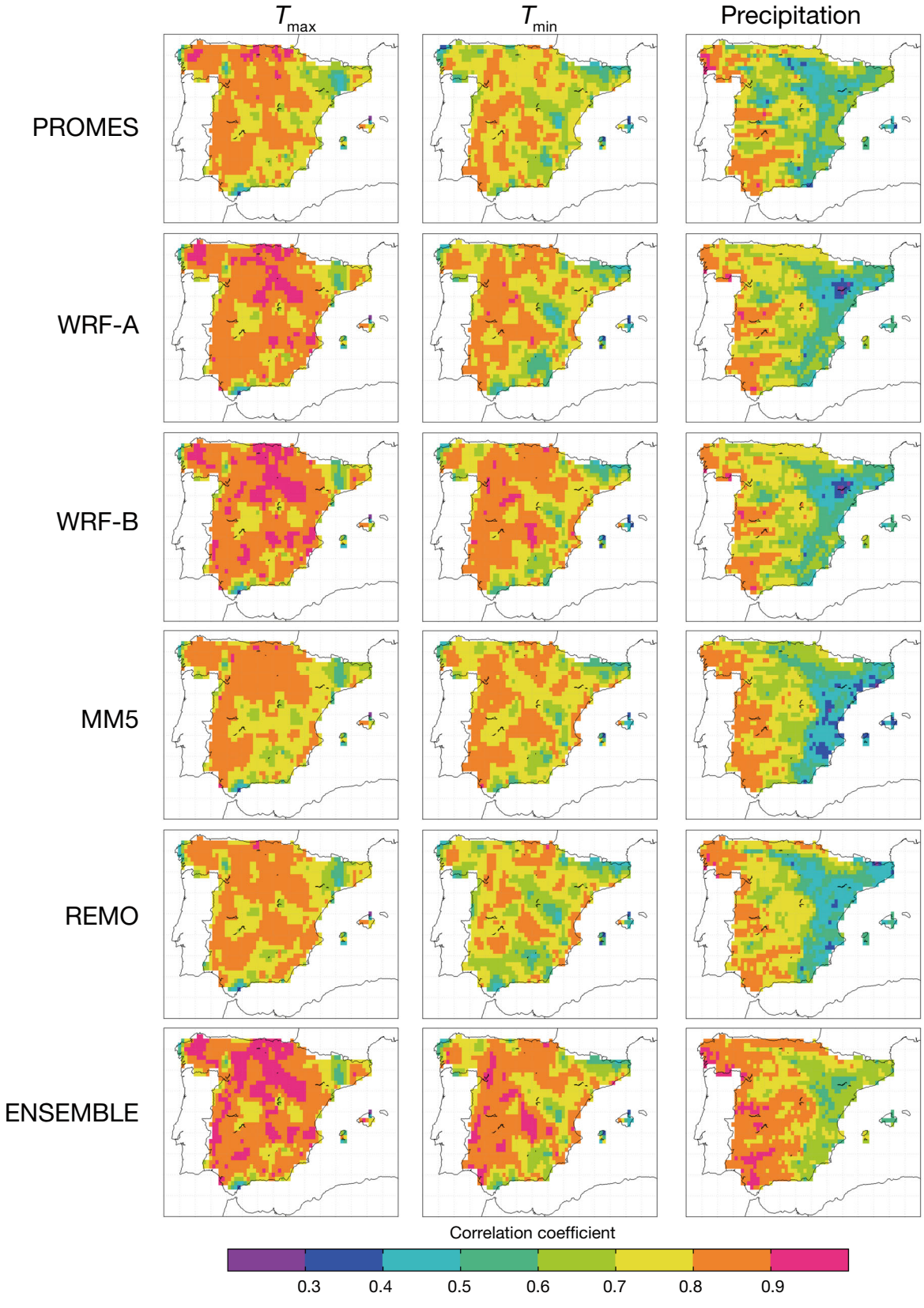

Fig. 6. Mean temporal correlation coefficients for simulated maximum (T_{max}) and minimum (T_{min}) temperature and precipitation vs. Spain02 observations. Correlations were calculated using individual months in each season yr^{-1} and for each point

In order to discriminate the seasonal behavior of this temporal correlation, Table 3 represents spatially averaged correlations for maximum and minimum temperature and precipitation. Maximum temperature shows the highest correlations and they present a low seasonal variability. Spring has slightly higher correlations than the rest of the seasons. In general, there is also a low model-to-model variability. Only WRF shows somewhat higher correlations (mostly >0.8). REMO, unlike the rest of the models, presents a value <0.70 in summer, indicating a lower capability to capture maximum temperature variability in summer. For minimum temperatures, correlations are, in general, lower than for maximum temperature and they exhibit a clearer seasonal variability, with higher correlations in winter and autumn. In this case, spring shows the lowest correlations, and several models show correlations <0.7 in different seasons. Only WRF-B is above this threshold throughout the year and performs best on average. For this variable, WRF-A is more similar to other models, such as MM5, than to WRF-B. Lastly, for precipitation the correlations are generally much lower than for temperature, and the seasonal dependence is very strong. As expected, higher correlation coefficients are attained during winter (0.77 to 0.80), when synoptic systems, well captured in the reanalysis data, drive precipitation. During summer, the synoptic activity is lower and precipitation is mostly driven by convective systems, less constrained by the boundary conditions; therefore, the lowest correlations are attained (0.42 to 0.55). An intermediate behaviour is shown in the transition seasons.

The most noticeable aspect regarding correlation (Fig. 6), is the large improvement achieved when considering the ensemble mean. Spatially (Fig. 6), there are areas where correlations are above any of the component members; most noticeably in precipitation over the eastern coast, where a common RCM defect is improved by the ensemble, softening the strong gradient present in each individual model. Also, after spatial averaging (Table 3), the ensemble mean performs at least as well as the best member, and in most cases shows better correlation than any of the members. This is true especially for precipitation.

Table 3. Time-correlation for modeled (a) maximum and (b) minimum temperature, and (c) precipitation vs. Spain02 observations for all seasons and diverse models included in ESCENA. Seasons: winter (DJF), spring (MAM), summer (JJA), autumn (SON). Light to dark shading indicates low to high correlation coefficients, respectively

	DJF	MAM	JJA	SON	Annual
(a)					
PROMES	0.77	0.79	0.77	0.80	0.78
WRF-A	0.78	0.85	0.83	0.84	0.83
WRF-B	0.80	0.86	0.83	0.84	0.84
MM5	0.77	0.83	0.74	0.80	0.78
REMO	0.78	0.83	0.69	0.83	0.79
Ensemble	0.81	0.87	0.83	0.86	0.84
(b)					
PROMES	0.67	0.74	0.77	0.76	0.73
WRF-A	0.76	0.69	0.74	0.80	0.75
WRF-B	0.82	0.74	0.74	0.82	0.78
MM5	0.81	0.70	0.73	0.79	0.76
REMO	0.78	0.66	0.68	0.78	0.72
Ensemble	0.82	0.74	0.77	0.83	0.79
(c)					
PROMES	0.79	0.68	0.55	0.60	0.67
WRF-A	0.78	0.68	0.52	0.62	0.68
WRF-B	0.80	0.69	0.53	0.63	0.69
MM5	0.77	0.64	0.44	0.62	0.66
REMO	0.77	0.67	0.42	0.64	0.66
Ensemble	0.85	0.77	0.63	0.73	0.77

3.3. Spatial variability

Here, we focus on the ability of the models to represent the observed spatial variability by using Taylor diagrams (Taylor 2001), which enable an easy comparison between the spatial and temporal patterns of 2 fields. In our diagrams (Fig. 7), the statistics displayed are the relative spatial SD (radial distance from the origin) and the correlation (cosine of the angular coordinate). Better models, in terms of centered root mean squared error (RMSE), are located closer to the black circle shown in Fig. 7, which corresponds to Spain02. Centered RMSE increases with distance from this point.

With respect to the mean field of maximum temperature (Fig. 7a), all models perform well, especially for winter, with very high spatial correlations and a normalized SD close to observations. However, PROMES and REMO represent excessive spatial variability during summertime. The spatial variability of the interannual SD (Fig. 7b), as well as its spatial mean (Table 2) is strongly overestimated by PROMES in spring and summer, providing higher values of spatial patterns of variability. There is a strong variation in the skill of the models as a function of the season of the year, with better correlation values during spring and autumn.

In the simulation of the mean minimum temperature (Fig. 7c), the models perform very similarly with each other, showing a high spatial correlation with the observations, but in most cases with a small underestimation of the spatial variability, except for

the cold season. The models do not capture the spatial structure of the variability during summer and wintertime (Fig. 7d). This variability is pervasively underpredicted for winter and summer, when spatial correlation does not exceed 0.45. Both for maximum and minimum temperature, the diverse models (and seasons) present a high spread in the representation of the spatial structure of the SD.

The mean precipitation pattern (Fig. 7e) is captured worse than those of temperature. This is expected, since the temperature patterns are strongly related to the orography. PROMES tends to provide a higher spatial variability than observations, especially for summertime (see red diamond outside the Taylor diagram in 7e), but with a high correlation (>0.88). In this case, REMO tends to better capture the field variability. For the SD pattern (Fig. 7f), a similar analysis can be performed, with PROMES tending to overestimate the SD of the spatial variability.

In this case, the added value of considering the ensemble mean is not clear for the maximum temperature (the ensemble mean outperforms most but not all models), albeit it generally improves the skill of most models. For minimum temperature and, overall, precipitation, the ensemble mean outperforms the individual models in all seasons, although it somewhat underestimates the variance of the SD patterns.

4. DISCUSSION AND CONCLUSIONS

In this work we have analyzed the RCM evaluation simulations of the ESCENA project over peninsular Spain and the Balearic Islands. This work evaluates, for the first time over a European region, the MM5 and WRF open-source, US-developed models along with European RCMs (PROMES and REMO). This study is also one of the earliest works using ERA-Interim as boundary conditions over Europe. As an initial evaluation of the ESCENA simulations, we used simple, widely used metrics to enable the comparison of the model performance with earlier studies. Also, we focused on the mean climate and interannual variability, therefore relying on monthly data.

As an indication of the quality of the simulations, we can compare the spatial correlation values for precipitation deduced from Fig. 7 with those obtained for RCMs in ENSEMBLES included in Herrera et al. (2010). That work classified the RCMs into 2 groups, according to a gap in the spatial correlation of precipitation. According to this classification, all individual RCMs in ESCENA would be in the group of the 'best' models. Also, the ensemble mean in ESCENA is similar to the best ensemble in the aforementioned work. Therefore, the ability to reproduce precipitation within ESCENA is high, at least in this aspect, when compared to Herrera et al. (2010) and references therein. The temperature results can be qualitatively compared to the bias maps shown by Christensen et al. (2010), which show typical biases within ENSEMBLES of the order of 2K in this region, similar to those found in ESCENA. Temperature-outlier models in ENSEMBLES show biases much larger than those in ESCENA. Therefore, overall, the 5 regional climate simulations included in ESCENA show good quality in reproducing the climatology of the IP, even though these models are based on completely different approaches and physical parameterizations.

WRF was used in 2 different settings, which differ only in the representation of the PBL. WRF-A used a local closure scheme, which has been reported to develop shallower, colder, and moister PBLs than WRF-B, which used a non-local closure scheme (García-Díez et al. 2013). The weaker low-level mixing in WRF-A leads to colder nighttime (i.e. minimum) temperatures. Therefore, the largest differences between the WRF ensemble members appear in the minimum temperature. Remarkably, MM5 is the only other model in the ensemble using a non-local closure PBL scheme, but the results are closer to WRF-A. It seems that biases arising from other components (model dynamics, numerics or other parameterizations) are compensating the expected summer warm bias.

Regarding the inherent uncertainties of model validation methodologies, climate models here are assessed on their capacity to reproduce present climate conditions, which in turn are established by comparing the output of climate simulations with observational datasets including gridded products (here, Spain02 dataset). However, due to the nature of the procedures to obtain observations and the statistical techniques employed to extrapolate this information onto reference gridded databases, they contain important uncertainties which may compromise the evaluation process. A recent work by Gómez-Navarro et al. (2012) indicates that, even in areas covered by dense monitoring networks such as Spain, uncertainties in the observations are comparable to the uncertainties within state-of-the-art RCMs driven by reanalysis, such as those used in ESCENA. Therefore, some of the common deficiencies found, can probably be traced to problems in the observational database. For instance, the warm biased minimum temperatures over northeastern Spain are a

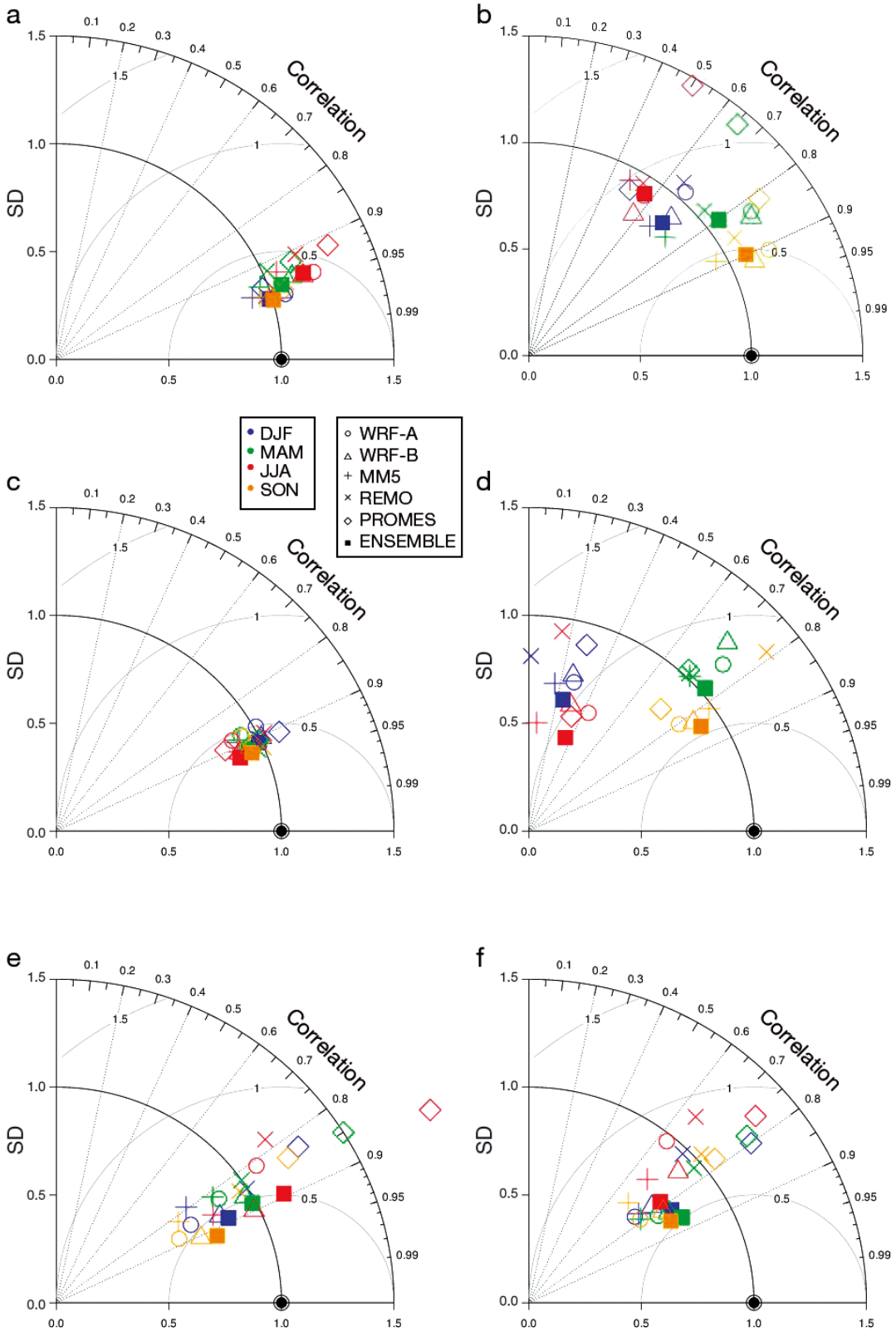

Fig. 7. Taylor diagrams for seasonal (a,b) maximum temperature, (c,d) minimum temperature, and (e,f) precipitation for all models included in the analysis. (a,c,e) Mean fields; (b,d,f) SD. Seasons: winter (DJF), spring (MAM), summer (JJA), autumn (SON)

candidate to a misrepresentation in the observational database, since (1) the problem is common to all models, (2) there are no complex topographic features over that area, and (3) the problem does not arise in any other variable.

No single model outperforms the rest of the models for all the variables analyzed. Depending on the variable and season, different models stand out. For example, REMO stands out for temperature (especially minimum temperature), PROMES shows a wet bias when the rest tend to be drier than observed, and WRF is also especially dry during autumn. Regarding the inclusion of US-developed models in the ensemble, we found no 'home court advantage' (Takle et al. 2007) for the European models. Unlike shown in other recent studies over north America (Mearns et al. 2012), the use of RCMs out of their 'home' domain did not lead to poorer performance.

The different performance of the RCMs in different seasons and variables encourages the use of the whole ensemble of simulations, taking the range of biases as an indicator of model uncertainty. The simplest way of considering the ensemble of RCMs is through the use of the ensemble mean as an additional simulation. The results show that the ensemble mean is usually less biased than the individual members or is close to the best member. Remarkably, the ensemble of simulations is able to correct the problems associated with the interannual variability for precipitation, showing substantially higher temporal correlation than the best individual model. These results are in accordance with previous works (Gleckler et al. 2008, Coppola et al. 2010, Kjellström et al. 2011). As stated by Annan & Hargreaves (2011), one hypothesis for the improvement of the ensemble mean when compared to the performance of the individual models is the paradigm of models being considered as independent samples from some distribution that is centered on the truth, as in this case the ensemble mean could be expected to converge to the truth as more models are added to the ensemble (Tebaldi & Knutti 2007). However, this hypothesis has been refuted by Knutti et al. (2010). Annan & Hargreaves (2010, 2011) defend the point that the statistically indistinguishable paradigm provides a reasonable basis for explanation of the properties of the CMIP3 ensemble. However, their results are only directly applicable to their specific comparisons and a convincing explanation for the outperformance of the ensemble mean is still an open question in climate science.

We can, therefore, conclude that the use of ensemble simulations in this kind of study substantially improves the representativity of the climatologies, and we also expect that studies of future climates, by using ensemble methodologies, will provide more robust climate projections and a valuable estimation of the associated uncertainty. As shown, these data can be complementary to other European projects such as ENSEMBLES over the Iberian Peninsula and surrounding areas. The data from this study are publicly available on the ESCENA project server (http://proyectoescena.uclm.es). The analyses shown are just an initial evaluation that needs to be extended in forthcoming studies by the climate and impact communities in the Iberian Peninsula, who are encouraged to use the whole RCM ensemble, thereby propagating the uncertainties found.

Acknowledgements. The Spanish R+D Programme of the Ministry of the Environment (Ministerio de Medio Ambiente y Rural y Marino) is acknowledged for the funding provided for the ESCENA project (Ref: 200800050084265). P.J.G. also thanks the Ramón y Cajal Programme of the Spanish Ministry of Science and Technology. This ministry is also acknowledged for the financial support through grants CGL2008-05112-C02-02, CGL2010-18013 and CGL2010-22158-C02 (this latter also funded by the European Fund for Regional Development [FEDER]. The authors also acknowledge the 5 anonymous reviewers for their valuable comments.

LITERATURE CITED

Abe M, Shiogama H, Hargreaves J, Annan J, Nozawa T, Emori S (2009) Correlation between inter-model similarities in spatial pattern for present and projected future mean climate. SOLA 5:133–136

Annan JD, Hargreaves JC (2010) Reliability of the CMIP3 ensemble. Geophys Res Lett 37:L02703, doi:10.1029/2009GL041994

ä Annan JD, Hargreaves JC (2011) Understanding the CMIP3 multimodel ensemble. J Clim 24:4529–4538

ä Argüeso D, Hidalgo-Muñoz JJ, Gamiz-Fortis SR, Esteban-Parra MJ, Castro-Diez Y (2012) Evaluation of WRF mean and extreme precipitation over Spain: present climate (1970–1999). J Clim 25:4883–4897

Arribas A, Gallardo C, Gaertner MA, Castro M (2003) Sensitivity of Iberian Peninsula climate to land degradation. Clim Dyn 20:477–489

ä Boo KO, Kwon WT, Baek HJ (2006) Change of extreme events of temperature and precipitation over Korea using regional projection of future climate change. Geophys Res Lett 33:L01791, doi:10.1029/2005GL023378

ä Boulanger JP, Brasseur G, Carril AF, de Castro M and others (2010) A Europe-South America network for climate change assessment and impact studies. Clim Change 98: 307–329

Brinkop S, Roeckner E (1995) Sensitivity of a general circulation model to parameterizations of cloud-turbulence interactions in the atmospheric boundary layer. Tellus A 47(2):197–220

Castro M, Fernández C, Gaertner MA (1993) Description of a mesoscale atmospheric numerical model. In: Díaz JI,

Lions JL (eds) Mathematics, climate and environment. Rech Math Appl Ser 27, Masson, Paris, p 230–253

Chaboureau JP, Bechtold P (2002) A simple cloud parameterization derived from cloud resolving model data: diagnostic and prognostic applications. J Atmos Sci 59:2362–2372

Chaboureau JP, Bechtold P (2005) Statistical representation of clouds in a regional model and the impact on the diurnal cycle of convection during tropical convection, cirrus and nitrogen oxides (TROCCINOX). J Geophys Res 110: D17103, doi:10.1029/2004JD005645

Chen F, Dudhia J (2001a) Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. I. Model implementation and sensitivity. Mon Weather Rev 129:569–585

Chen F, Dudhia J (2001b) Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. II. Preliminary model validation. Mon Weather Rev 129:587–604

Christensen JH, Christensen OB (2007) A summary of the PRUDENCE model projections of changes in European climate by the end of this century. Clim Change 81:7–30

Christensen JH, Kjellström E, Giorgi F, Lenderink G, Rummukainen M (2010) Weight assignment in regional climate models. Clim Res 44:179–194

Collins WD, Rasch PJ, Boville DA, Hack JJ, McCaa JR, Williamson DK, Briegleb BP (2006) The formulation and atmospheric simulation of the Community Atmosphere Model. J Clim 19:2144–2161

Coppola E, Giorgi F, Rauscher SA, Piani C (2010) Model weighting based on mesoscale structures in precipitation and temperature in an ensemble of regional climate models. Clim Res 44:121–134

Cuxart J, Bougeault P, Redelsperger JL (2000) A turbulence scheme allowing for mesoscale and large-eddy simulations. Q J R Meteorol Soc 126:1–30

d'Almeida GA, Koepke P, Shettle EP (1991) Atmospheric aerosols: global climatology and radiative characteristics. A Deepak Publ, Hampton, VA

Davies HC (1976) A lateral boundary formulation for multilevel prediction models. Q J R Meteorol Soc 102:405–418

Dee DP, Uppsala SM, Simmons AJ, Berrisford P and others (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137:553–597

Domínguez M, Gaertner MA, de Rosnay P, Losada T (2010) A regional climate model simulation over West Africa: parameterization tests and analysis of land-surface fields. Clim Dyn 35:249–265

Domínguez M, Romera R, Sánchez E, Fita L and others (2013) Present climate precipitation and temperature extremes over Spain from a set of high resolution RCM. Clim Res (in press) doi:10.3354/cr01186

Dudhia J (1989) Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. J Atmos Sci 46: 3077–3107

Dudhia J (1993) A nonhydrostatic version of the Penn State NCAR mesoscale model: validation tests and simulation of an Atlantic cyclone and cold front. Mon Weather Rev 121:1493–1513

ECMWF (2004) IFS Documentation CY28r1, Chapter 4: physical processes, Section 2: radiation. Available at www.ecmwf.int/research/ifsdocs/CY28r1/index.html

Fernández J, Montávez JP, Saenz J, Gonzalez-Rouco JF, Zorita E (2007) Sensitivity of the MM5 mesoscale model to physical parameterizations for regional climate studies: annual cycle. J Geophys Res 112(D4):D04101, doi:10.1029/2005JD006649

Fernández J, Primo C, Cofiño AS, Gutiérrez JM, Rodríguez MA (2009) MVL spatiotemporal analysis for model intercomparison in EPS: application to the DEMETER multimodel ensemble. Clim Dyn 33:233–243

Fernández-Quiruelas V, Fita L, Fernández J, Cofiño A (2010) WRF workflow on the Grid with WRF4G. Proc 11th WRF Users' Workshop, Boulder, CO

Fita L, Fernández J, García-Díez M (2010) CLWRF: WRF modifications for regional climate simulation under future scenarios. Proc 11th WRF Users' Workshop, Boulder, CO

Font-Tullot I (2000) Climatología de España y Portugal, 2nd edn. Universidad de Salamanca, Salamanca

Gaertner MA, Castro M (1996) A new method for vertical interpolation of the mass field. Mon Weather Rev 124: 1596–1603

Gaertner MA, Christensen OB, Prego JA, Polcher J, Gallardo C, Castro M (2001) The impact of deforestation on the hydrological cycle in the western Mediterranean: an ensemble study with two regional climate models. Clim Dyn 17:857–873

Gaertner MA, Jacob D, Gil V, Domínguez M, Padorno E, Sánchez E, Castro M (2007) Tropical cyclones over the Mediterranean Sea in climate change simulations. Geophys Res Lett 34:L14711, doi:10.1029/2007GL029977

Gaertner MA, Domínguez M, Garvert MA (2010) Modelling case-study of soil moisture-atmosphere coupling. Q J R Meteorol Soc 136:483–495

Gallardo C, Arribas A, Prego JA, Gaertner MA, de Castro M (2001) Multi-year simulations using a regional-climate model over the Iberian Peninsula: current climate and doubled CO_2 scenario. Q J R Meteorol Soc 127:1659–1682

Gao X, Pal JS, Giorgi F (2006) Projected changes in mean and extreme precipitation over the Mediterranean region from a high resolution double nested RCM simulation. Geophys Res Lett 33:L03706, doi:10.1029/2005GL024954

García-Díez M, Fernández J, Fita L, Yagüe C (2013) Seasonal dependence of WRF model biases and sensitivity to PBL schemes over Europe. Q J R Meteorol Soc 139:501–514

Giorgi F, Bi X, Pal JS (2004) Mean, interannual variability and trends in a regional climate change experiment over Europe. I. Present-day climate (1961–1990). Clim Dyn 22:733–756

Gleckler P, Taylor K, Doutriax C (2008) Performance metrics for climate models. J Geophys Res 113:D06104, doi:10.1029/2007JD008972

Gómez-Navarro JJ, Montávez JP, Jiménez-Guerrero P, Jerez S, García-Valero JA, González-Rouco JF (2010) Warming patterns in regional climate projections over the Iberian Peninsula. Met Zeit (Hambg) 19:275–285

Gómez-Navarro JJ, Montávez JP, Jerez S, Jiménez-Guerrero P, Lorente-Plazas R, González-Rouco JF, Zorita E (2011) A regional climate simulation over the Iberian Peninsula for the last millennium. Clim Past 7:451–472

Gómez-Navarro JJ, Montávez JP, Jerez S, Jiménez-Guerrero P, Zorita E (2012) What is the role of the observational dataset in the evaluation and scoring of climate models? Geophys Res Lett 39:L24701, doi:10.1029/2012GL054206

Grell GA (1993) Prognostic evaluation of assumptions used by cumulus parameterizations. Mon Weather Rev 121: 764–787

ä Grell GA, Dévényi D (2002) A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. Geophys Res Lett 29(14), doi:10.1029/2002GL015311

Grell GA, Dudhia J, Stauffer DR (1994) A description of the fifth-generation PennState/NCAR mesoscale model (MM5). NCAR Tech Note NCAR/TN-398+STR, NCAR, Boulder, CO. Available at www.mmm.ucar.edu/mm5

ä Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M (2008) A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. J Geophys Res 113:D20119, doi:10.1029/2008JD010201

ä Herrera S, Fita L, Fernández J, Gutiérrez J (2010) Evaluation of the mean and extreme precipitation regimes from the ENSEMBLES regional climate multimodel simulations over Spain. J Geophys Res 115:D21117, doi:10.1029/2010JD013936

ä Herrera S, Gutiérrez JM, Ancell R, Pons MR, Frías MD, Fernández J (2012) Development and analysis of a 50-year high-resolution daily gridded precipitation dataset over Spain (Spain02). Int J Climatol 32:74–85

ä Hong SY, Lin JOJ (2006) The WRF Single-Moment 6-Class Microphysics Scheme (WSM6). J Korean Meteorol Soc 42:129–151

Hong SY, Pan HL (1996) Comparison of NCEP/NCAR Reanalysis with 1987 FIFE data. Mon Weather Rev 124:1480–1498

ä Hong SY, Dudhia J, Chen S (2004) A revised approach to ice microphysical processes for the bulk parameterization of cloud and precipitation. Mon Weather Rev 132:103–120

Houghton JT, Meira Filho LG, Callander BA, Harris N, Kattenberg A, Maskell K (eds) (1996) Climate change 1995. The science of climate change. Cambridge University Press, Cambridge

ä Jacob D, Van den Hurk BJJM, André U, Elgered G and others (2001) A comprehensive model inter-comparison study investigating the water budget during the BALTEX-PIDCAP period. Meteorol Atmos Phys 77:19–43

ä Jacob D, Bärring L, Christensen OB, Christensen JH and others (2007) An intercomparison of regional climate models for Europe: model performance in present-day climate. Clim Change 81:31–52

ä Jerez S, Montávez JP, Gomez-Navarro JJ, Jiménez-Guerrero P, Jiménez J, González-Rouco JF (2010) Temperature sensitivity to the land-surface model in MM5 climate simulations over the Iberian Peninsula. Met Zeit 19:363–374

ä Jerez S, Montávez JP, Jiménez-Guerrero P, Gómez-Navarro JJ, Lorente-Plazas R, Zorita E (2012a) A multiphysics ensemble of present-day climate regional simulations over the Iberian Peninsula. Clim Dyn 40:3023–3046

Jerez S, Montávez JP, Gómez-Navarro JJ, Jiménez PA, Jiménez-Guerrero P, Lorente R, González-Rouco JF (2012b) The role of the land-surface model for climate change projections over the Iberian Peninsula. J Geophys Res 117:D01109, doi:10.1029/2011JD016576

Jerez S, Montávez JP, Gómez-Navarro JJ, Lorente-Plazas R, García-Valero JA, Jiménez-Guerrero P (2013) A multiphysics ensemble of regional climate change projections over the Iberian Peninsula. Clim Dyn 41:1749–1768

Kain JS, Fritsch JM (1993) Convective parameterization for mesoscale models: the Kain-Fritsch scheme. The representation of cumulus convection in numerical models. Meteor Monogr 24:165–170

Kjellström E, Nikulim F, Hanson U, Strandberg G, Ullerstig A (2011) 21st century changes in the European climate: uncertainties derived from an ensemble of regional climate model simulations. Tellus 63A:24–40

ä Klemp JB, Skamarock WC, Dudhia J (2007) Conservative split-explicit time integration methods for the compressible non-hydrostatic equations. Mon Weather Rev 135:2897–2913

ä Knutti R, Furrer R, Tebaldi C, Cermak J, Meehl GA (2010) Challenges in combining projections from multiple climate models. J Clim 23:2739–2758

Koepke P, Hess M, Schult I, Shettle EP (1997) Global aerosol data set. Rep 243, Max Planck Inst Meteorol, Hamburg. Available at http://opac.userweb.mwn.de/radaer/gads.html

ä Krinner G, Viovy N, Noblet-Ducoudre N, Ogee J and others (2005) A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. Global Biogeochem Cycles 19:BG1015, doi:10.1029/2003GB002199

ä Mearns LO, Arritt R, Biner S, Bukovsky MS and others (2012) The North American Regional Climate Change Assessment Program: overview of phase I results. Bull Am Meteorol Soc 93:1337–1362

ä Mlawer EJ, Taubman SJ, Brown PD, Iacono MJ, Clough SA (1997) Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. J Geophys Res 102:16663–16682

Morcrette J, Smith L, Fourquart Y (1986) Pressure and temperature dependence of the absorption in longwave radiation parameterizations. Beitr Phys Atmos 59:455–469

Nordeng T (1994) Extended versions of the convective parametrization scheme at ECMWF and their impact on the mean and transient activity of the model in the tropics. Res Dep, Tech Memo 206, ECMWF, Reading

ä Nunez MN, Solman SA, Cabre MF (2009) Regional climate change experiments over southern South America. II. Climate change scenarios in the late twenty-first century. Clim Dyn 32:1081–1095

Roeckner E, Arpe K, Bengtsson L, Christoph M and others (1996) The atmospheric general circulation model ECHAM4: model description and simulation of present-day climate. Rep 218, Max Planck Inst Meteorol, Hamburg

Sánchez E, Romera R, Gaertner MA, Gallardo C, Castro M (2009) A weighting proposal for an ensemble of regional climate models over Europe driven by 1961–2000 ERA40 based on monthly precipitation probability density functions. Atmos Sci Lett 10:241–248

ä Simmons AJ, Burridge DM (1981) An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. Mon Weather Rev 109:758–766

Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Duda MG (2008) A description of the Advanced Research WRF version 3. NCAR Tech Note NCAR/TN20201c475+STR, available at www.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf

ä Takle ES, Roads J, Rockel WJ, Gutowski WJ and others (2007) Transferability intercomparison. Bull Am Meteorol Soc 88:375–384

ä Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. J Geophys Res 106:7183–7192

ä Tebaldi C, Knutti R (2007) The use of the multi-model ensemble in probabilistic climate projections. Philos Trans R Soc Lond A 365:2053–2075

Tiedtke M (1989) A comprehensive mass flux scheme for cumulus parameterization in large–scale models. Mon Weather Rev 117(8):1779–1800

ä Uppala SM, Kallberg PW, Simmons AJ, Andrae U and others (2005) The ERA-40 re-analysis. Q J R Meteorol Soc 131: 2961–3012

Van der Linden P, Mitchell JFB (2009) ENSEMBLES: climate change and its impacts: summary of research and results from the ENSEMBLES project. Met Office Hadley Centre Tech Rep, Exeter

ä Whetton P, Macadam I, Bathols J, O'Grady J (2007) Assessment of the use of current climate patterns to evaluate regional enhanced greenhouse response patterns of climate models. Geophys Res Lett 34:L14701, doi:10.1029/2007GL030025

Editorial responsibility: Filippo Giorgi, Trieste, Italy

Submitted: July 29, 2011; Accepted: April 24, 2013
Proofs received from author(s): September 26, 2013