Consistency of climate change projections from multiple global and regional model intercomparison projects

J. Fernández · M. D. Frías · W. D. Cabos · A. S. Cofiño · M. Domínguez · L. Fita · M. A. Gaertner · M. García-Díez · J. M. Gutiérrez · P. Jiménez-Guerrero · G. Liguori · J. P. Montávez · R. Romera · E. Sánchez

Abstract We present an unprecedented ensemble of 196 future climate projections arising from different global and regional model intercomparison projects (MIPs): CMIP3, CMIP5, ENSEMBLES, ESCENA, EURO- and Med-CORDEX. This multi-MIP ensemble includes all regional climate model (RCM) projections publicly available to date, along with their driving global climate models (GCMs). We illustrate consistent and conflicting messages using continental Spain and the Balearic Islands as target region. The study considers near future (2021-2050) changes and their dependence on several uncertainty sources sampled in the multi-MIP ensemble: GCM, future scenario, internal variability, RCM, and spatial resolution. This initial work focuses on mean seasonal precipitation and temperature changes. The results show that the potential GCM-RCM combinations have been

J. Fernández (✉) · M.D. Frías · A. S. Cofiño
Grupo de Meteorología y Computación. Dpto. Matemática Aplicada y Ciencias de la Computación. Univ. de Cantabria, Avda. de los Castros, s/n, 39005 Santander, Spain
Tel.: +34 942201723
E-mail: Jesus.Fernandez@unican.es

W. D. Cabos · G. Liguori
Dpto. Física y Matemática. Univ. de Alcalá, 28805 Madrid, Spain

M. García-Díez
Predictia Intelligent Data Solutions. 39005 Santander, Spain

M. Domínguez · M. A. Gaertner · R. Romera · E. Sánchez
Facultad Ciencias Ambientales y Bioquímica. Univ. Castilla-La Mancha, Avda. Carlos III s/n, 45071 Toledo, Spain

L. Fita
Centro de Investigaciones del Mar y la Atmósfera (CIMA), CONICET-UBA, Buenos Aires, Argentina

J. M. Gutiérrez
Grupo de Meteorología. Instituto de Física de Cantabria (IFCA). CSIC - Universidad de Cantabria, Avda. de los Castros, s/n, 39005 Santander, Spain

P. Jiménez-Guerrero · J. P. Montávez
Dpto. Física. Campus de Excelencia Internacional Mare Nostrum. Univ. de Murcia, 30100 Murcia, Spain

explored very unevenly, with favoured GCMs and large ensembles of a few RCMs that do not respond to any ensemble design. Therefore, the grand-ensemble is weighted towards a few models. The selection of a balanced, credible sub-ensemble is challenged in this study by illustrating several conflicting responses between the RCM and its driving GCM and among different RCMs. Sub-ensembles from different initiatives are dominated by different uncertainty sources, being the driving GCM the main contributor to uncertainty in the grand-ensemble. For this analysis of the near future changes, the emission scenario does not lead to a strong uncertainty. Despite the extra computational effort, for mean seasonal changes, the increase in resolution does not lead to important changes.

1 Introduction

Regional climate change information is increasingly being demanded by different vulnerability, impact and adaptation (VIA) research communities (Hewitson et al, 2013). This information is required to feed models which, eventually, will produce information for specific sectors (health, energy, food availability, risk management, water resources) and enter decision-making processes at different levels (Huynen and Martens, 2015; Koutroulis et al, 2015; Giannini et al, 2016). The distillation of information out of the huge amount of available data is a technical and ethical challenge (Hewitson et al, 2013). A key approach to produce future climate scenarios is the use of numerical models. A recent series of international projects and initiatives have produced, using both global (GCM) and regional (RCM) climate models, a huge ensemble of future climate projections, which samples most of the uncertainties affecting climate change. The model development cycle, along with the periodic Intergovernmental Panel on Climate Change (IPCC) reporting cycle, impose a rhythm in climate scenario production which can hardly be followed by the scientific community feeding from them. Specific technological infrastructures, such as the Earth System Grid Federation (ESGF; Williams et al, 2015), have been developed to face the data storage and discovery challenge posed by the 5th Coupled Model Intercomparison Project (CMIP5) feeding the last IPCC Assessment Report (AR5) (Stocker et al, 2013). Each new initiative increases model complexity and/or spatial resolution, resolving more and more processes relevant for the anthropogenic climate change assessment. Data users switch to the latest available products, without exhausting previously existing databases. In contrast, this work combines 6 ensembles of future climate projections from different model types and generations, showing common and conflicting messages that arise from very simple analyses.

VIA research communities are usually advised to consider an ensemble of model projections in order to propagate the uncertainty arising from different greenhouse gases (GHG) scenarios and climate simulation tools. In the past, data availability usually drove the selection of models, e.g. researchers used models from their own institution (Easterling et al, 2001; Knowlton et al, 2007), from a project consortium (Dessai, 2003; Sima et al, 2015), or data from the few models openly available (Kalkstein and Greene, 1997; Rötter et al, 2013). Since 2004, the advent of the

ESGF, climate change scenario model output data can be accessed homogeneously, through a single access infrastructure and in a common data format which eases the processing of any number of models and ensemble members. The use of the ESGF is still complex and slow for an average user, and a number of tools have been developed to bridge the gap with the users; for example, *climate4impact* (Plieger et al, 2015) or *ESGFToolsUI* (Terry, 2014). Despite these technical advances to access data, fundamental transdisciplinary issues might still hamper a proper usage of regional climate data (Rössler et al, 2017; Hewitson et al, 2017).

An important gap in the provision of regional climate change information is the scale mismatch between GCM model output and the spatial scale of most climate impact applications. Downscaling techniques (Wilby and Wigley, 1997) have been developed in the last decades to bridge this gap by means of two main approaches: dynamical (RCMs; Laprise, 2008; Rummukainen, 2010; Giorgi and Gutowski, 2015) and empirical-statistical (Maraun et al, 2010; Gutiérrez et al, 2012). The empirical-statistical downscaling (ESD) builds empirical relationships between large-scale and local observations, which are exploited to generate local climate information out of coarse GCM projections. Unfortunately, the availability of comprehensive ensembles of future climate projections derived from ESD is very scarce to date, although initiatives such as VALUE (Maraun et al, 2015) and the COordinated Regional climate Downscaling EXperiment (CORDEX; Giorgi and Gutowski, 2015) are currently devoting a strong effort to this approach. Dynamical downscaling through regional climate models, on the other hand, has been coordinated during the last two decades, particularly over certain regions. In Europe, a series of chained projects (Regionalization (1993-1994), RACCS (1995-1996), MERCURE (1997-2000), PRUDENCE (2001-2004) and ENSEMBLES (2004-2009)) coordinated a regional climate modelling community, nowadays under the umbrella of CORDEX forming the EURO-CORDEX (Jacob et al, 2014) and Med-CORDEX (Ruti et al, 2016) communities.

In parallel to international initiatives, some national efforts are also producing downscaled regional climate projections. Some examples are the DRIAS project for France (Lémond et al, 2011), KNMI'15 for the Netherlands (Tank et al, 2015), UKCP for the United Kingdom (Murphy et al, 2009) or PNACC-2012 for Spain (Gómez et al, 2016; San-Martín et al, 2017). We included dynamical downscaling (ESCENA project, see Section 2.2) results from the latter, given that it covers our target region. Other regional, national climate change projection studies rely on data from international initiatives and usually focus on a single data source: Soares et al (2014) for Portugal/ENSEMBLES, Ouzeau et al (2016) for France/EURO-CORDEX or Kis et al (2017) for Carpathian region/ENSEMBLES. Some studies combine a couple of initiatives, but this is not common practice. For example, Gobiet et al (2014) and Tolika et al (2012) consider ENSEMBLES and PRUDENCE, Jacob et al (2014) compare EURO-CORDEX and ENSEMBLES. Data from the global models used to drive regional projections are seldom considered (Turco et al, 2013).

The objective of this work is twofold:

1. Compare climate change signals arising from different recent ensembles of regional climate projections produced by numerical climate models. The comparison will also consider other factors sampled in the grand-ensemble, such as different GCMs, RCMs, RCM spatial resolutions, GHG scenarios, etc.

2. Provide basic climate change scenarios for continental Spain and the Balearic Islands. The body of the paper focuses on specific results illustrating the grand-ensemble heterogeneity, while the Electronic Supplementary Material provides results for other seasons, variables or factors.

The paper is structured as follows: Section 2 summarizes the data gathered for this study. The data processing and analysis methodology is described in Section 3. The results section (4) introduces spatially-averaged delta changes for precipitation and near surface temperature (Section 4.1), which are decomposed in different uncertainty sources in Section 4.2. Cautionary remarks on the use of delta changes (Section 4.3) and conflicting projections (Section 4.4) are provided next. The paper closes with a discussion (Section 5) and a brief summary of the main conclusions (Section 6).

2 Data

This study uses only dynamical (global and regional) model output from transient climate change simulations. Observations and evaluation simulations (using reanalysis data as boundaries) were not considered, but all models presented have been evaluated elsewhere (see below). A total of 196 future climate projections (Figure 1) for Spain have been collected from the global (CMIP3 and CMIP5) and the regional (ENSEMBLES, ESCENA, EURO-CORDEX and Med-CORDEX) initiatives described below. This is, to our knowledge, the largest ensemble of scenarios ever considered for a region.

2.1 ENSEMBLES

As part of the ENSEMBLES FP6 project (van der Linden and Mitchell, 2009), a multi-model ensemble of RCMs was produced. An evaluation of this ensemble forced by boundary conditions from reanalysis data for precipitation and maximum and minimum temperature over Europe can be found in Kjellström et al (2010). The quality of the climate change simulations for these variables has been reported in several project technical reports (`http://ensembles-eu.metoffice.com/tech_reports.html`) and elsewhere for particular RCMs or variables (Boberg et al, 2009; Kjellström et al, 2011).

13 RCMs from ENSEMBLES were used in this study, mostly at a horizontal resolution of about 25 Km, but also at about 50 Km. Boundary conditions for these RCMs were taken from 7 CMIP3 GCMs forced by the A1B emissions scenario. Publicly available data were obtained from the ENSEMBLES RCM data archive at the Danish Meteorological Institute (DMI).

2.2 ESCENA

Project ESCENA (2008-2012) was part of the Spanish Strategic action on energy and climate change (`http://meteo.unican.es/projects/escena`). Funded by the Spanish government, this call aimed at providing the scientific basis for the assessment of regional climate change impacts over Spain (including continental Spain,

Fig. 1 GCM-RCM matrix for the multi-project ensemble considered in this study and publicly available as of April 30th, 2017. Each square represents a future climate projection. For each GCM-RCM entry, a 3×3 matrix represents the availability of different forcing scenarios and RCM resolutions (see upper left legend). Marginal counts are shown for each GCM and RCM. The legend also shows, in parenthesis, the count for each project/initiative considered. Dashed rectangles show full factorial sub-ensembles within ESCENA and EURO-CORDEX (see Section 5.1).

the Balearic and Canary Islands). This project agglutinated the regional climate modelling efforts of this action and coordinated 4 different institutions using 4 different RCMs: Universidad de Castilla-La Mancha (PROMES model), Universidad de Murcia (MM5 model), Universidad de Alcalá de Henares (REMO model) and Universidad de Cantabria (WRF model). Two versions of the WRF-v3.1.1 model were considered, WRA and WRB, which differ in the planetary boundary layer physics scheme (using a local vs. non-local closure, respectively). The boundary forcing came from CMIP3, as in ENSEMBLES, and the horizontal resolution was also about 25 km. The historical and scenario simulations span the period 1951-2050. A brief description of the RCMs used in the project and an evaluation of their performance under present climate conditions in terms of precipitation and temperature can be found in Jiménez-Guerrero et al (2013) and Domínguez et al (2013). Gómez et al (2016) evaluated surface wind according to this data set and investigated future changes.

2.3 EURO-CORDEX

As part of the global CORDEX initiative, EURO-CORDEX provides climate scenarios for Europe mainly at two resolutions, $\sim$12 km (EUR-11) and $\sim$50 km (EUR-44), complementing previous data sets from ENSEMBLES. The regional simulations are based on CMIP5 global climate projections (Taylor et al, 2011), forced by different representative concentration pathways (RCPs; van Vuuren et al, 2011). In this case low (RCP 2.6), midrange (RCP 4.5) and high (RCP 8.5) emissions scenarios are considered depending on the model. Previous studies have evaluated the EURO-CORDEX results in present day climate (Vautard et al, 2013; Kotlarski et al, 2014) or compared the climate change projections to those from ENSEMBLES (Jacob et al, 2014) for two key variables for impact assessment studies: temperature and precipitation. Overall, the EURO-CORDEX evaluation highlights the general ability of the RCMs to represent the basic spatio-temporal patterns over Europe with some limitations for selected metrics, regions and seasons (Kotlarski et al, 2014). It was also found an agreement between EURO-CORDEX and ENSEMBLES results (Jacob et al, 2014).

10 RCMs from EUR-44 resolution and 8 from EUR-11 were considered in this analysis, (Figure 1). A few more simulations were considered since the beginning of this study, but during this time they were removed from the server due to different problems detected (modellers communication to EURO-CORDEX community). The volatility of these data is discussed later on in this work. This study finally considers only those RCM simulations available through the ESGF as of April 30th, 2017.

2.4 Med-CORDEX

Med-CORDEX is a coordinated contribution to the CORDEX initiative focused on the Mediterranean basin. The particularity of Med-CORDEX is that it includes regional atmospheric, land surface, river and oceanic climate models and coupled RCMs (Ruti et al, 2016). Evaluation simulations from Med-CORDEX are analysed by Dell'Aquila et al (2016) and compared to the previous generation simulations from the ENSEMBLES project. Ayar et al (2016) evaluate the precipitation of 3 RCMs from Med-CORDEX, 2 from EURO-CORDEX and 6 statistical downscaling methods. There are a few other evaluation studies for particular Med-CORDEX models and/or regions (Tramblay et al, 2013; Flaounas et al, 2013; Stéfanon et al, 2015; Liguori et al, 2017).

Currently, most of the simulations available at the Med-CORDEX database (`http://www.medcordex.eu`) are from atmospheric models. We considered 3 RCMs and one fully-coupled (atmosphere-land-ocean) model for 50 km resolution (MED-44) and 2 RCMs for 12 km resolution (MED-11).

2.5 Global model data

Finally, we considered also the output from the GCMs used as driving boundary conditions for the RCM simulations in the previous projects/initiatives. They were downloaded from different repositories. CMIP3 data (Meehl et al, 2007) was mainly
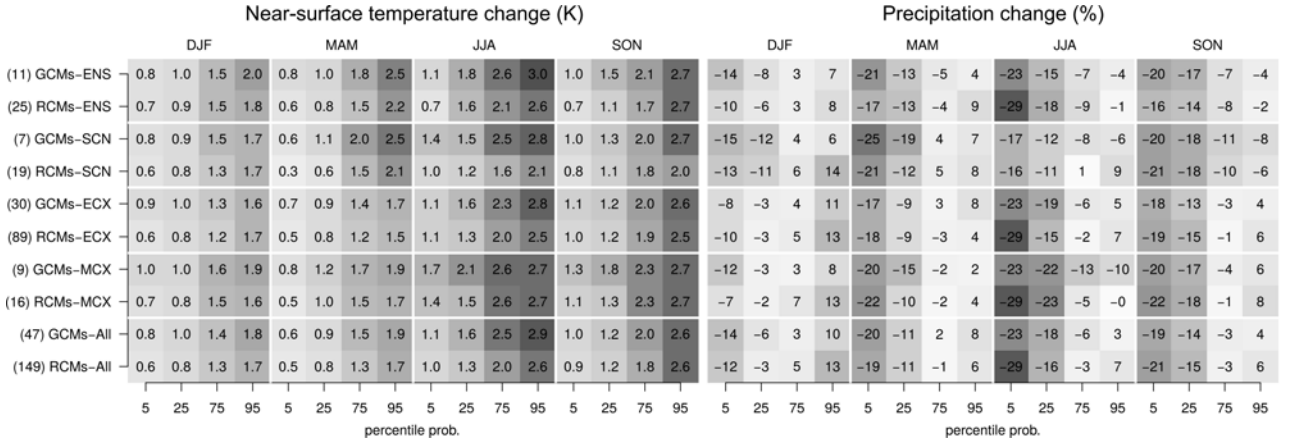
Near-surface temperature change (K)

	DJF 5	25	75	95	MAM 5	25	75	95	JJA 5	25	75	95	SON 5	25	75	95
(11) GCMs-ENS	0.8	1.0	1.5	2.0	0.8	1.0	1.8	2.5	1.1	1.8	2.6	3.0	1.0	1.5	2.1	2.7
(25) RCMs-ENS	0.7	0.9	1.5	1.8	0.6	0.8	1.5	2.2	0.7	1.6	2.1	2.6	0.7	1.1	1.7	2.7
(7) GCMs-SCN	0.8	0.9	1.5	1.7	0.6	1.1	2.0	2.5	1.4	1.5	2.5	2.8	1.0	1.3	2.0	2.7
(19) RCMs-SCN	0.6	0.8	1.3	1.7	0.3	0.6	1.5	2.1	1.0	1.2	1.6	2.1	0.8	1.1	1.8	2.0
(30) GCMs-ECX	0.9	1.0	1.3	1.6	0.7	0.9	1.4	1.7	1.1	1.6	2.3	2.8	1.1	1.2	2.0	2.6
(89) RCMs-ECX	0.6	0.8	1.2	1.7	0.5	0.8	1.2	1.5	1.1	1.3	2.0	2.5	1.0	1.2	1.9	2.5
(9) GCMs-MCX	1.0	1.0	1.6	1.9	0.8	1.2	1.7	1.9	1.7	2.1	2.6	2.7	1.3	1.8	2.3	2.7
(16) RCMs-MCX	0.7	0.8	1.5	1.6	0.5	1.0	1.5	1.7	1.4	1.5	2.6	2.7	1.1	1.3	2.3	2.7
(47) GCMs-All	0.8	1.0	1.4	1.8	0.6	0.9	1.5	1.9	1.1	1.6	2.5	2.9	1.0	1.2	2.0	2.6
(149) RCMs-All	0.6	0.8	1.3	1.7	0.5	0.8	1.3	1.7	1.0	1.3	2.0	2.6	0.9	1.2	1.8	2.6

Precipitation change (%)

	DJF 5	25	75	95	MAM 5	25	75	95	JJA 5	25	75	95	SON 5	25	75	95
(11) GCMs-ENS	-14	-8	3	7	-21	-13	-5	4	-23	-15	-7	-4	-20	-17	-7	-4
(25) RCMs-ENS	-10	-6	3	8	-17	-13	-4	9	-29	-18	-9	-1	-16	-14	-8	-2
(7) GCMs-SCN	-15	-12	4	6	-25	-19	4	7	-17	-12	-8	-6	-20	-18	-11	-8
(19) RCMs-SCN	-13	-11	6	14	-21	-12	5	8	-16	-11	1	9	-21	-18	-10	-6
(30) GCMs-ECX	-8	-3	4	11	-17	-9	3	8	-23	-19	-6	5	-18	-13	-3	4
(89) RCMs-ECX	-10	-3	5	13	-18	-9	-3	4	-29	-15	-2	7	-19	-15	-1	6
(9) GCMs-MCX	-12	-3	3	8	-20	-15	-2	2	-23	-22	-13	-10	-20	-17	-4	6
(16) RCMs-MCX	-7	-2	7	13	-22	-10	-2	4	-29	-23	-5	-0	-22	-18	-1	8
(47) GCMs-All	-14	-6	3	10	-20	-11	2	8	-23	-18	-6	3	-19	-14	-3	4
(149) RCMs-All	-12	-3	5	13	-19	-11	-1	6	-29	-16	-3	7	-21	-15	-3	6

percentile prob.

Fig. 2 Central 50% (25th to 75th percentile) and 90% (5th to 95th percentile) ranges for the delta changes spatially averaged over Continental Spain and the Balearic Islands. Changes are shown for near-surface temperature (left, in K) and precipitation (right, as percent change). The changes projected by different sub-ensembles are shown in rows. Each project (ENSembles, eSCeNa, EuroCordeX, MedCordeX) is divided into the GCMs used to drive the regional projections and the RCM results. The last rows (All) show grand-ensemble results, separately for GCMs and RCMs.

obtained from the Climate and Environmental Retrieval and Archive (CERA) database (Lautenschlager et al, 2015). This is the driving data for the ENSEMBLES and ESCENA projects. CMIP5 data were obtained through the ESGF. Model data from this experiment drove EURO- and Med-CORDEX regional simulations. Some driving models were not publicly available and, therefore, are missing in this study (See empty positions in the first column of Figure 1).

3 Methodology

There is a high number of dimensions involved in regional climate change projections. Giorgi et al (2008) envisaged them as components of a regional climate change "hyper-matrix", which is now at the core of the CORDEX framework. One can (should) consider different: Concentration scenarios, GCMs, RCMs, RCM resolutions, GCM (RCM) realizations, GCM (RCM) versions, etc.; each spanning a different uncertainty, which depends on the projection horizon (Hawkins and Sutton, 2009; Déqué et al, 2012). Uncertainties can be properly evaluated using ensembles of simulations (Palmer, 2000), if these are sufficiently large and well designed (Daron and Stainforth, 2013). However, even when projecting the available ensemble on just two dimensions, such as GCM and RCM, the phase space spanned is relatively empty (Figure 1). Moreover, the analysis can focus on different variables, seasons, future time-slices, statistics (mean, variability, extremes, ...).

With this perspective, we focus on a simple approach. We examine only precipitation and mean surface temperature, usually considered in the applications and climate change projection summaries, although we discuss (Section 5.2) the limited view these two variables provide. We consider only the mean climate, in particular, seasonal climatologies of 30-year periods.

We concentrate on a near future (2021-2050), which maximises the ensemble members available (limited by ESCENA and some ENSEMBLES RCMs). We will see that this also has the advantage to consider only weak GHG-concentration

differences. The response to these weak differences is smaller than the internal model variability. Pooling these projections is, therefore, closer to a random sample from the same population (of realizations from the same model climate), than in far future periods, when the different scenarios clearly represent different populations and cannot be easily merged (what is the weight of each GHG scenario?).

We selected the reference historical period 1971-2000 and consider the standard delta-change approach (Räisänen, 2007) to explore future climate change projections. We highlight some of the caveats behind this approach in Section 4.3.

Even though we pool the ensembles from different sources in a grand-ensemble, in recognition of potentially different populations being merged, we avoid grand-ensemble summaries (means, PDFs) and illustrate individual members factorized by the many dimensions considered (GCM, RCM, project/initiative, resolution, ...). We warn against grand-ensemble summaries by showing the split of bimodal PDFs under some factorizations. The lack of independence (complex and unknown dependence among projections) and the incompleteness of the interactions between dimensions prevents the use of many standard statistical tools. Instead of summary/factorial statistics, we favour graphical displays of the full grand-ensemble.

For brevity, we show mainly summer, June through August (JJA), results. The Electronic Supplementary Material provides equivalent results for other seasons, which are commented where appropriate and, in any case, left as reference for downscaled scenario users interested in this region. Figure 2 shows that in JJA occur the largest differences among projects.

All data were interpolated by a nearest neighbour algorithm from their different projected grids to a regular 0.2° latitude-longitude grid spanning continental Spain and the Balearic Islands. Spatial averages (e.g. those in Section 4.1) consider only these land areas.

4 Results

4.1 Projected delta changes for precipitation and temperature

Projected delta changes for precipitation and temperature are summarized for the different model intercomparison projects (MIP) in Figure 2. To emphasize the inherent uncertainty in these projections, no central estimate is provided. The ranges encompassing 50 and 90% of the projections for each MIP are provided instead, split into forcing GCMs and dynamically downscaled (RCMs) results. Overall, the different MIPs agree, showing the most prominent change during summer (JJA) for both variables. Some differences can also be identified across MIPs, e.g. the extended upper end of the temperature change range during spring (MAM) for the CMIP3-based projections (ENSEMBLES and ESCENA), or the less dry summer projection in ESCENA. However, one of the most striking features of Figure 2 is the systematic reduction of the temperature change by the RCM ensembles with respect to their driving GCMs. No MIP projected a single range limit higher than their driving GCM ensemble. For precipitation, this is not systematic, and RCMs even show opposite effects on the deltas in different seasons: drier summers and wetter winters than in the driving GCMs. We look into individual ensemble members to trace the origin of these summary statistics.
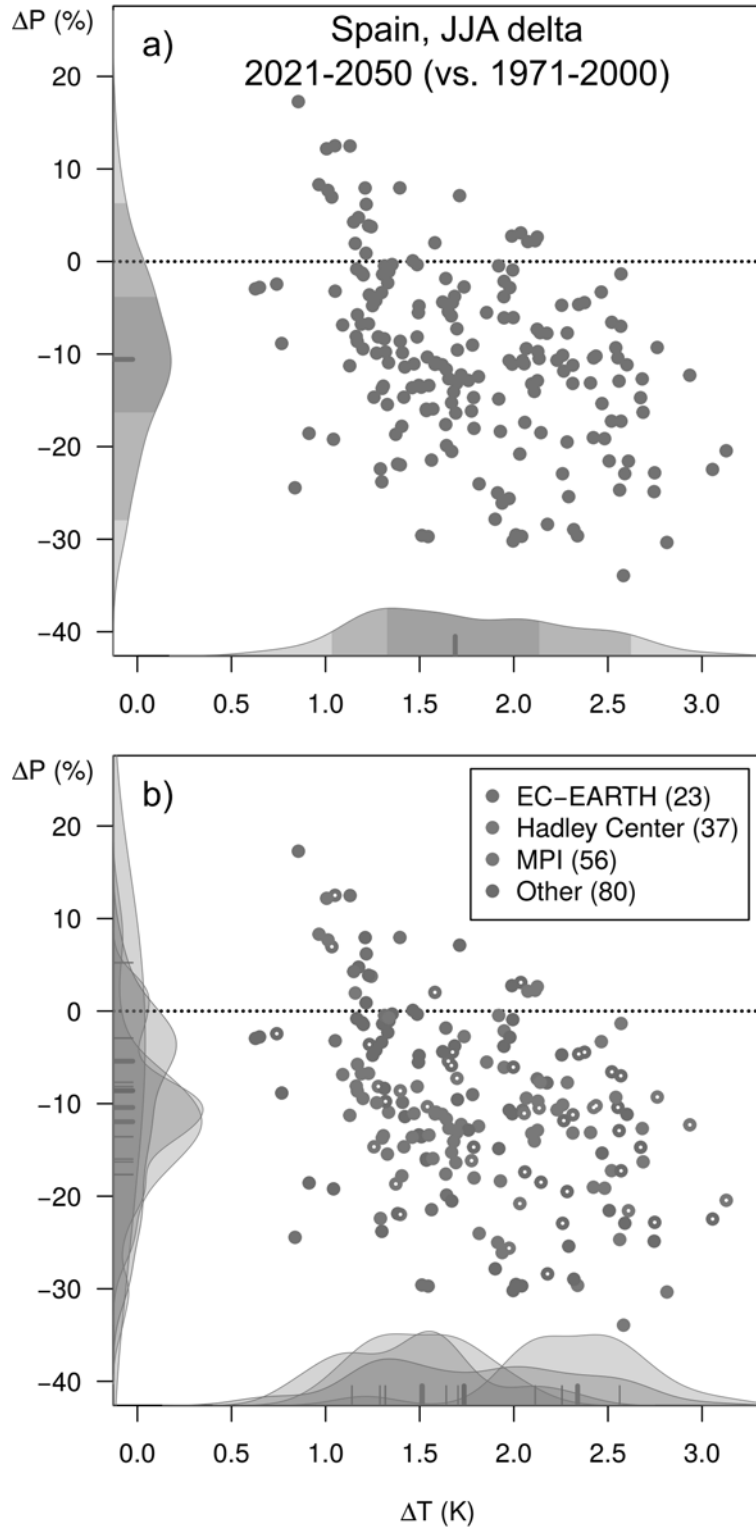
Fig. 3 Precipitation vs mean surface temperature delta changes spatially averaged over Continental Spain and the Balearic Islands for each ensemble member. In (a), marginal probability density functions are shown for each variable pooling the whole 196-member ensemble, using a Gaussian kernel density estimator. Different shading shows the ranges from the 5th to 95th percentile (90% of the sample) and from the 1st to 3rd quartile (50%). The median is also shown as an inner tickmark. In (b), the scatter plot and PDFs are factorized by GCM family (see text). In this case, only the quartiles and median are shown (as tickmarks). Raw GCM output deltas are shown as empty circles.

The analysis of the full 196-member ensemble (Figure 3a) shows that 50% (90%) of the simulated summer near surface temperatures project an increase of 1.3 to 2.1 K (1.0 to 2.6 K) by 2021-2050, with respect to the average of the 1971-2000 period, on average for continental Spain and the Balearic Islands. The median increase is about 1.7 K, but this value is misleading, since the sample of projections is slightly bimodal, and there are relatively few projections close to this value.

Marginal results for precipitation show that more than 75% of the models project a decrease in summer rainfall, with a median decrease of 11%. Even though this decrease could be not statistically significant, what is clear is the inverse relationship between precipitation and temperature, typical of moisture-limited land-atmosphere feedback regimes (Seneviratne et al, 2010; Jerez et al, 2013; Stegehuis et al, 2012), which respond to increasing temperatures by drying out the soils and limiting moisture availability for summer precipitation. Causality is always difficult to establish (a decrease in precipitation also leads to increasing temperature by reducing soil moisture, thus increasing the Bowen ratio), but the relationship in Figure 3a is clear. The relationship also emerges during spring (MAM) and slightly in autumn (SON) but, as expected, is absent during winter (DJF) (Figure ESM2). In winter, an energy-limited regime prevails and frontal precipitation feeds from moisture advected from the Atlantic ocean and less from local evaporation (Rios-Entenza et al, 2014).

The bimodality of the projected near surface temperature can be partly explained by the driving GCM. Moreover, GCM "families" tend to cluster and provide high or low delta values (Figure 3b). In particular, due to their different climate sensitivities, models developed by the EC-EARTH international consortium and the Max Plank Institute (MPI) for Meteorology in Germany give rise to deltas in the lower range (median about 1.5 K), while Hadley Center MetOffice (UK) models project the largest deltas (median close to 2.3 K). The "Other" family labelled in Figure 3b includes the rest of the models, which contribute fewer members to the grand-ensemble, and consist of driving GCMs developed in research centres across the world (France, Norway, Canada, USA, Japan, Australia). The clustering in temperature change induces a clustering in precipitation according to the relationship previously mentioned. Most Hadley Center-derived projections show a precipitation decrease, while the EC-EARTH family shows mixed sign projections. Interestingly, the MPI model family shows lower end temperature deltas, but mostly precipitation decreases. The leading role of GCM family in partitioning the delta range is also clear in other seasons (not shown).

Other factors, such as source project (Figure ESM1b) or model resolution (Figure ESM1c) do not lead to specific clusters. The GHG scenario leads to some clusters (Figure ESM1d), but not consistent with their radiative forcing. A high-emission scenario, such as SRES A2, lies at the lower end of the temperature delta range. Very low concentration scenarios (RCP2.6) do not rank at the lower end and intermediate- to high-concentration scenarios (A1B, RCP4.5, RCP8.5) span the whole temperature delta range. Incidentally, the coldest and the warmest member of the ensemble is driven by the same scenario (A1B). Therefore, scenario forcing does not seem to be playing an important role in the near future projections.

Global vs. regional climate model projections show differences in near future temperature which are likely statistically significant (Figure ESM1f). They arise from an upper end of the temperature delta range mostly populated by GCM
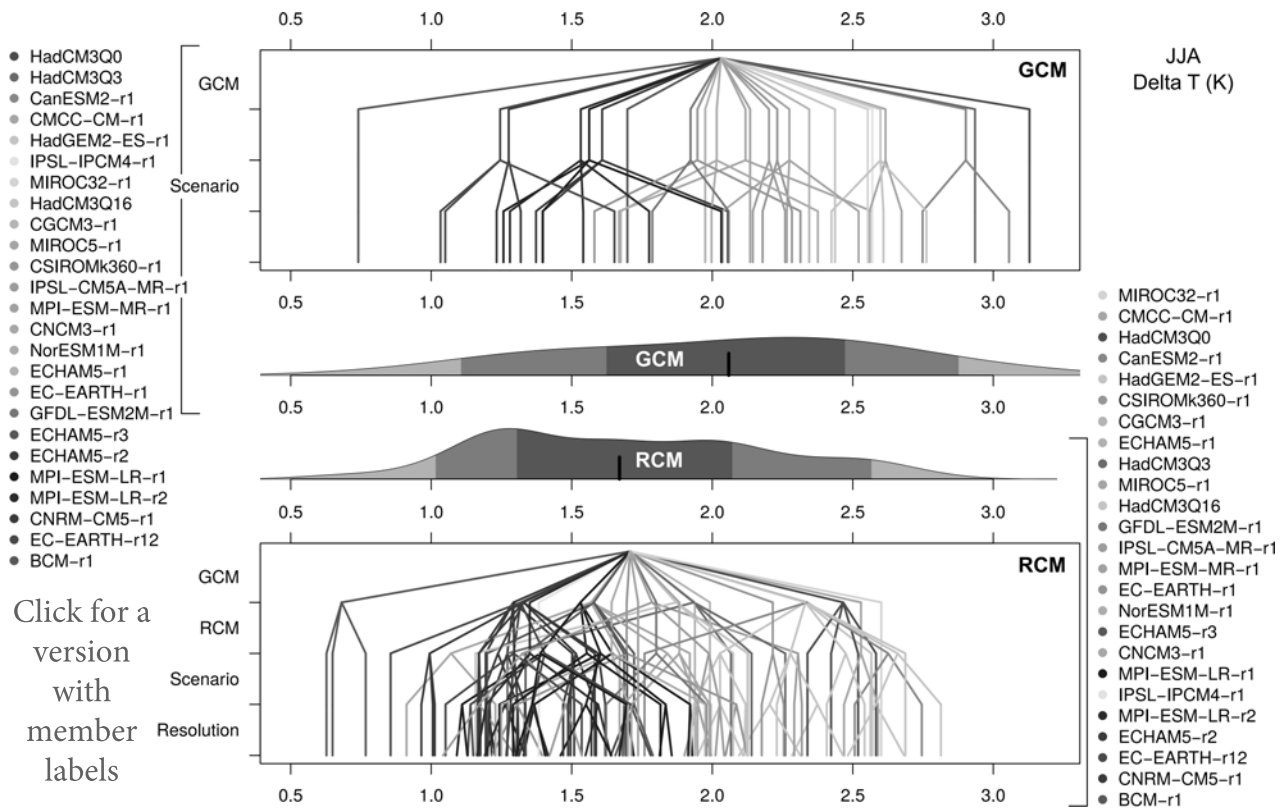
Fig. 4 Uncertainty cascade representing spatially averaged JJA near surface temperature deltas over Spain averaged over all GCMs (upper tip where all lines converge; top) and RCMs (bottom). From top to bottom, for each line plot, members forced by the same GCM are averaged (and coloured by GCM). Next, each line splits into the mean for each RCM nested into that GCM. Next, each line splits into the mean for each scenario. Finally, those lines corresponding to RCMs nested with two different horizontal resolutions in the same scenario and GCM are split. Therefore, at the bottom of each plot, all individual member projections can be found. The right legends show each individual GCM ranked by their projected JJA temperature delta. Additionally, a kernel density estimation of the PDF for each ensemble is shown, shading 90 and 50% sample ranges and marking the ensemble median. Cascades follow the visualization proposed by Ed Hawkins on his Climate Lab Book `http://www.climate-lab-book.ac.uk/2014/cascade-of-uncertainty`

projections and a lower end populated by RCMs. In precipitation, the driest future projections are provided by RCMs. Statistically sound tests could be applied to assign a likelihood to the occurrence of such differences in two random samples if their population means were equal. However, we will see in the next section that the deltas are not a random sample, but the result of very specific choices in the driving GCMs used in the different projects.

4.2 Uncertainty cascade

The distribution of temperature delta projections is clearly different in the direct GCM model output (median above 2K) and the dynamically downscaled projections (median 1.6K; see empirical PDF estimates in Figure 4). Given the non-systematic GCM-RCM combinations in our ensemble (Figure 1), the behaviour of individual nestings is of interest.

In order to decompose the uncertainty cascade down to individual projections, we start decomposing by GCM, given that this is the main source of uncertainty

identified so far. Figure 4 (top) shows the JJA near future temperature uncertainty cascade considering only GCM projections. The mean JJA delta derived from the ensemble of GCM projections (47 members) is 2K. Ensembles of GHG scenario forcing for each GCM are averaged next and coloured according to their mean JJA temperature delta. The GCM delta ranking shows the 4 Hadley Centre models in the top 8, while BCM projects the smallest delta change (0.75K, the only one below 1K). Individual members are split when reaching the scenario step. Although, in general, higher radiative forcing scenarios from a single GCM lead to warmer projections, this is not systematic. This reinforces the view, for this time horizon, of GHG scenario as a smaller source of uncertainty as compared to internal model variability. For instance, EC-EARTH r1 and r12 members forced by RCP4.5 show larger differences than EC-EARTH-r12 forced by RCP2.6 and RCP4.5.

Dynamically-downscaled projections usually lead to smaller JJA temperature deltas than their driving GCM projections in this area (Figure 4, bottom). Specific GCM choices tend to enhance this effect. For instance, the GCM showing less warming (BCM) was downscaled by 3 different RCMs, increasing its weight in the ensemble. The GCM delta ranking is only slightly altered by downscaling. RCMs nested into the same GCM tend to produce similar deltas (small spread at the RCM step), although the are exceptions. WRF- and MM5-based projections tend to give rise to smaller delta changes than other models nested in the same GCM. In the next section, we will see that this is linked to absolute temperature biases.

Scenario and RCM resolution (0.44, 0.22 or 0.11) make little difference (small spread at those steps) in the country-level-averaged projected delta changes. The only exceptions, where the scenario step shows some spread, are inherited from the GCM (e.g. from large differences between EC-EARTH-r12 RCP2.6/4.5 and RCP8.5).

4.3 Delta method caveats

The so-called delta method (Räisänen, 2007) has been used as a standard method to present future projections from our ensemble. In the simple terms applied here, we assume that the difference in future minus reference mean climate will cancel out likely model errors. This is related to bias correction methods. In particular, delta changes are insensitive to local shift bias correction methods, the simplest correction applied to temperatures to match the modelled and observed climatology. Relative delta changes are insensitive to local scaling, usually applied to precipitation due to its statistical distribution and to preserve its lower bound. More sophisticated bias correction methods could be applied, such as quantile mapping. However, for those methods, bias corrected and delta change projections differ (Ho et al, 2012; Räisänen and Räty, 2013), leading to a new source of uncertainty.

The contribution of sophisticated bias correction methods to climate change signals is still a field of active study (see e.g. Gobiet et al, 2015; Casanueva et al, 2017; Turco et al, 2017, and references therein). However, there are clear indications that the simple delta method has a limited accuracy. An illustration is shown in Figure 5, where JJA near surface temperature deltas are plotted as a function of mean near surface temperature (relative to the coldest estimate) during the reference period. One striking fact is that the 30-year temperature average, spatially
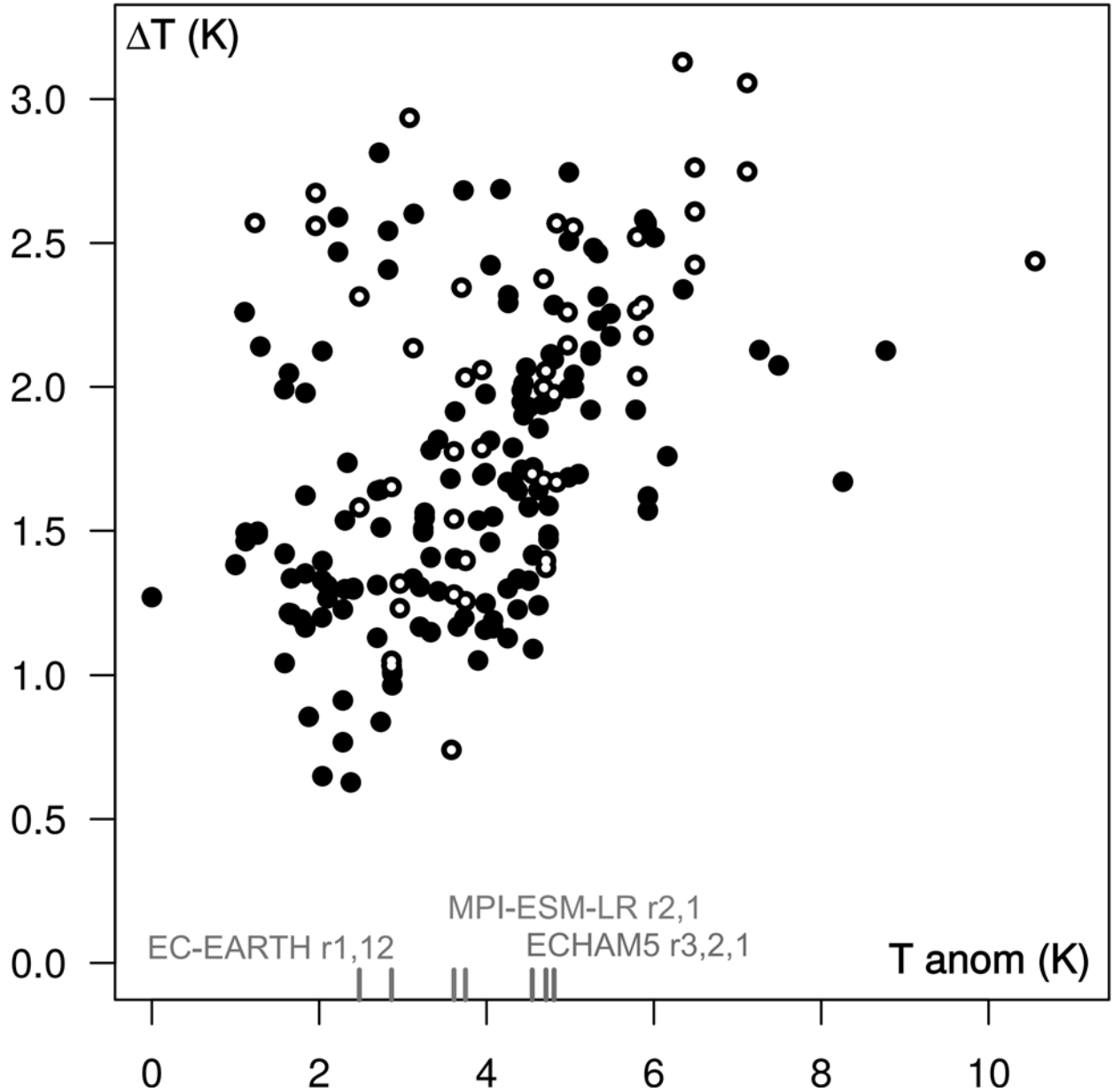
Fig. 5 Scatter plot showing projected JJA near-surface temperature delta changes vs. the temperature anomaly during the reference period (1971-2000) with respect to the coldest member of the ensemble in that period. Open circles represent GCM direct model output. The abscissas of the 3 GCMs with multiple realizations in our ensemble are highlighted in colour (see text).

averaged over a relatively large region, shows a range of 10K across the grand ensemble. This range is not the result of some outliers, but it is densely filled at least up to 7K. The range is smaller for GCMs, which start with a minimum anomaly of 1.5K, probably due to the lower orography. Decadal variability cannot explain this spread, as illustrated by the GCMs with multiple realizations present in our ensemble (Figure 5). The spread of these sub-ensembles, started from different initial conditions in 1850, is well below 0.5K after more than a century of simulation. Differences among models dominate the spread in simulated present climate.

Another evident feature is the non-random distribution of delta changes. Warm models during the reference period tend to produce stronger delta changes. This could be explained by soil moisture-temperature feedbacks. Warm (cool) models during summer in this area are likely to be due to dry (wet) soils. The extra radiative forcing from future scenarios goes directly into sensible heat (i.e. temperature

increase) for those models with dry soils, while it can be partly used to evaporate soil water (latent heat) in the models that preserve wet soils in the future. The relationship shown in Figure 5 is, therefore, typical of regions/seasons with transitional soil moisture-temperature feedback regimes between energy- and moisture-limited. Models have shown problems reproducing this transitional regimes (Stegehuis et al, 2012) and slight changes can have drastic consequences on temperature. For instance, the cold summers produced by WRF are due to excessive rainfall and soil moisture (Knist et al, 2017).

The delta change dependence on absolute temperature has been known for some time and there are several proposals to correct (constrain) the projections based on it (Boberg and Christensen, 2012) or some other physically relevant processes (Hall and Qu, 2006; Bellprat et al, 2013; Stegehuis et al, 2013). These corrections are subject to their own uncertainties, since they are based on additional models for the relationship and involve observations, with their own uncertainties.

4.4 Conflicting messages

What to expect from downscaling? GCMs provide coarse resolution climate change projections. Individual grid points should not be considered representative of their exact location due to the so-called skilful scale (Grotch and MacCracken, 1991; von Storch et al, 1993). Moreover, many important features affecting the climate of a region could be entirely absent in a GCM. For example, the Pyrenees mountain chain or the Balearic Islands are simply not there in many of the GCMs considered in this study. The point of downscaling is combining the climate information at skilful scales of the GCM with local, non-resolved features (orography, land-sea contrasts, etc.). As such, downscaling is not expected to change GCM projections overall for extended regions, but smaller scale changes could appear, as a result of the interaction of the GCM dynamics with the regional characteristics. The ability, or even possibility, of RCMs in changing (improving?) the GCM large-scale circulation is debated (Diaconescu and Laprise, 2013; Xue et al, 2014; Hall, 2014). The principle of downscaling is to estimate local climate subject to the large-scale climate generated by the GCM. In this sense, strong changes of the overall climate change pattern should be analysed with care. Also, conflicting messages could arise between the GCM and RCM or between several RCMs nested into the same GCM (Turco et al, 2013). These should also be analysed to discern genuine modelling uncertainty from unrealistic response. The line between these two might be difficult to draw, though.

In this work, no attempt is made to assess the added value of downscaling. Added value could be behind some of the RCM projected changes shown next that differ from those projected by their driving GCM. Some of these differences could be interpreted as added value, but this would require a more in-depth, case-by-case analysis and a clear definition of added value (see e.g. Rummukainen, 2016, for a recent review).

Figure 6 shows some spatial responses to climate change forcing scenarios by several GCMs and RCMs nested into them. They were selected to illustrate the plausibility of the downscaled response. Each row shows the forcing GCM JJA precipitation relative delta change (first column) and the dynamical downscaling by different RCMs.
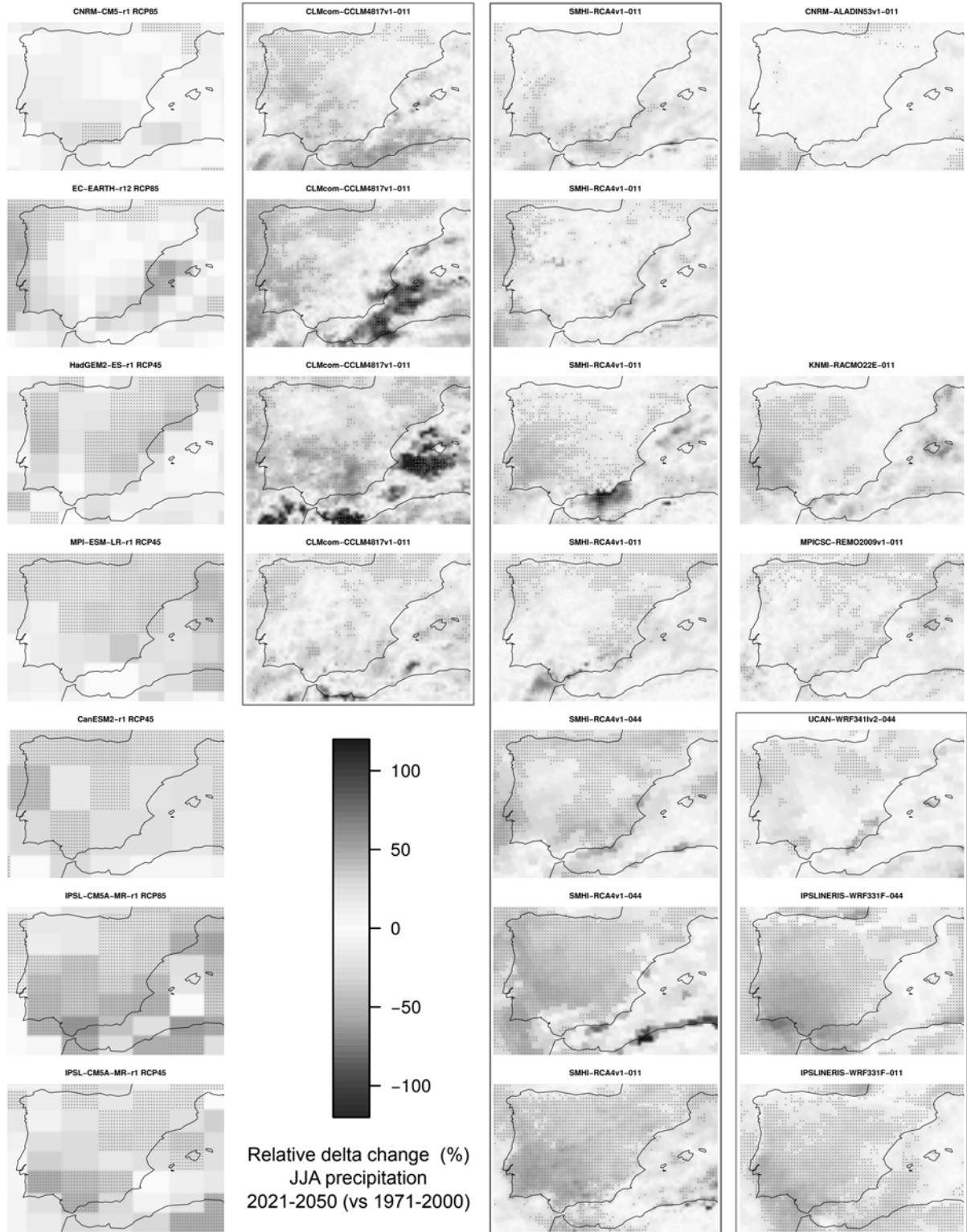
Fig. 6 Summer (JJA) precipitation relative delta changes (in %) for several members of the EURO-CORDEX ensemble, along with the direct model output (in the corresponding first column) in which the RCMs were nested. Dotted grid cells indicate significant changes with 90% confidence after a t-test on the absolute mean difference.

For example, the top row shows the climate change response by the CNRM-CM5 model (realization r1) to the RCP8.5 scenario. There is a slight southwest-to-northeast dipole, from a summer precipitation decrease to an increase. An overall similar pattern can be discerned in the CCLM-4.8.17 RCM nested into it. However, small scale features appear. CCLM projects a stronger precipitation decrease over southern parts of the domain (especially over the Alboran Sea, just east of the
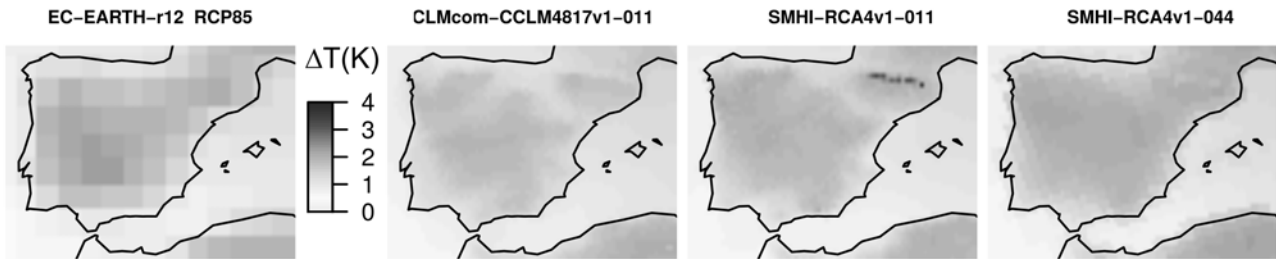
Fig. 7 Summer (JJA) near surface temperature delta (K) patterns for the direct global model output of EC-EARTH (r12) forced by the RCP8.5 scenario and downscaled by CCLM4.8.17-v1 (0.11) and RCA4-v1 (0.11° and 0.44° resolutions).

Gibraltar Strait). In EC-EARTH-r12 (second row), the dipole is northwest-to-southeast, and this pattern is, again, put into regional context by CCLM, which projects a precipitation increase west of the Balearic Islands, consistent with a similar but coarser pattern in the GCM. This Mediterranean precipitation increase also appears in CCLM when nested into HadGEM2-ES-r1 (third row). However, this global model does not show a trace of this precipitation increase. This increase in summer precipitation over the western Mediterranean Sea could be interpreted as added value of the dynamical downscaling, responding more properly to the large scale forcing, or as a specific bias of this RCM.

Other RCM (RCA4, in the third column) nested into the same GCMs shows overall similar patterns, but regional details differ instead of reinforce. For instance, a GCM-conflicting response appears north of the Alboran Sea when nested into HadGEM2-ES-r1, with a centre of strong precipitation increase where CCLM shows a decrease. Other simulations nested into HadGEM2-ES-r1 also show inconsistent responses so it is likely that a problem exists in the coupling with this GCM. All KNMI runs nesting the RACMO RCM into this GCM have been withdrawn from the ESGF and re-run during the development of this study. The new version (v2 in ESGF, shown in Figure 6) shows a moderate summer rainfall increase over the Mediterranean Sea, unlike the initial run, which showed larger and more extended increases (not shown).

Several other examples of conflicting messages are shown in Figure 6 and need to be analysed in detail to find out the causes behind the differences (this is out of the scope of this work). Some of them could be regional added value of the downscaling, while others might simply reflect downscaling deficiencies, which artificially enlarge uncertainties. See, e.g. the increased precipitation developed over the north African coast by CCLM nested into MPI-ESM-LR-r1, over the Cantabrian Sea (north of the Iberian Peninsula) by WRFv3.3.1-F nested into the IPSL-CM5A or the conflicting pattern between WRFv3.4.1-I and RCA4, nested into CanESM2.

An example of physically consistent, but unplausible, delta changes is shown in Figure 7. RCA4 at 0.11 resolution provides a very local pattern of change (reaching 4K) over the Pyrenees. This elevation-dependent warming (Pepin et al, 2015) pattern appears in every RCA4 0.11 projection, and it is systematically absent in the lower resolution runs (0.44) or in any other RCM projection at 0.11 resolution (CCLM-4.8.17 shown as an example). This pattern is typical, in most models and resolutions, during winter (not shown) and is related to reduced snow cover due to warming and a positive snow-albedo feedback. We checked the snow depth (not shown) and confirmed that RCA4 produces significant snow cover

in the Pyrenees during summer when the resolution resolves high terrain. The resolution increase makes the modelled surface temperature cross the threshold to support snow cover, but slightly enough to lie back behind the threshold in the near future projections. Since extended snow cover in the Pyrenees during summer is unrealistic under present conditions, so it is the small-scale climate change pattern depicted by RCA4 0.11 over this area. Of course, during winter, when similar small-scale features are produced by many RCMs, this is a clear added value of downscaling. Given that the smooth orography of GCMs cannot resolve snow accumulating at mountain tops, they miss important local delta temperature changes arising from changes in the snowline.

This is another example of a threshold-dependent, non-linear process breaking the delta method basic assumption that model errors remain constant. Processes depending on absolute thresholds give rise to delta changes which cannot be improved by assuming that model error evolves as some smooth function.

5 Discussion

5.1 Lack of ensemble design

The data shown in Figure 1 are essentially all future projections publicly available for our target region in the last decade. GCM output is restricted to those models used as boundary conditions in the downscaled projections. Therefore, the first column could be greatly enlarged (only from CMIP5, there are more than 200 other projections available through the ESGF), but the interior of the matrix would remain empty.

The GCM-RCM matrix is not only mostly empty, but heterogeneously filled, with preferred GCMs and RCMs. A few GCMs were selected by most regional climate modelling groups. For instance: MPI-ESM-LR was downscaled 23 times, ECHAM5 (20), CNRM-CM5 (19), EC-EARTH (18). Few RCMs contribute a large fraction of the grand-ensemble, notably RCA4 with 36 scenario simulations.

The decision of the GCM to downscale usually falls within 2 categories:

1. GCM from the same institution as the downscaling group. Some examples from ENSEMBLES: HadRM nested only into HadCM3 projections, MRCC nested only into CGCM. This practice remains in CORDEX: ALADIN only nested into CNRM-CM5 runs, REMO only nested into MPI-ESM runs, etc.
2. GCM was easily available, or available earlier. Even though GCM data is available from different sources, the amount of model levels and the time frequency required to drive an RCM, reduces drastically the available options. Thus, GCMs archiving full model levels at high frequency (usually only a given run for each scenario) are candidates to feed dynamical downscaling.

These can also be combined: e.g. in category 1, apart from the obvious institutional interest, we can assume motivations regarding the early and ease of access to full resolution GCM data. RCM simulation consortia including institutions running global simulations (e.g. ENSEMBLES) have traditionally eased the access to GCM boundary conditions. GCM selection criteria in ESCENA also followed ease of GCM data availability by several participating groups. Incidentally (not by design), the GCMs in ESCENA span a wide range of future projections, including the ECHAM and Hadley Center families.

As a result of the above, none of the ensembles considered in this study fully explores future climate change uncertainties. Thanks to GCM boundary sharing within consortia, some uncertainties can be partially explored:

In ENSEMBLES, RCM uncertainty can be explored in the ECHAM5-r3 sub-ensemble (5 RCMs) or in the HadCM3Q0 sub-ensemble, with a different set of 4 RCMs. These are the most populated sub-ensembles to explore RCM uncertainty and their intersection is empty: none of the RCMs downscaled both GCMs.

ESCENA was designed to have a GCM (ECHAM5-r2) downscaled by all RCMs and 2 RCMs (PROMES and MM5) downscaling all GCMs. Unfortunately, a different ECHAM5 realization (r2) from that in ENSEMBLES (r3) was selected, preventing the exploration of domain position (in REMO) or the inclusion of an RCM (PROMES) in both, the ECHAM5-r3 and HadCM3Q0 sub-ensembles.

In CORDEX, the situation is not better. Despite the large number of scenario simulations (105), the largest, complete GCM-RCM matrix is 4×2, formed by (CNRM-CM5-r1, EC-EARTH-r12, MPI-ESM-LR-r1, HadGEM2-ES-r1)× (CCLMv4.8.17 RCA4), exploring 2 scenarios (RCP 4.5 and 8.5) with a single resolution (0.11). A critical example is provided by the HadGEM2-ES-r1≫RegCM4.3v1 coupling, which offers publicly the 0.44 RCP8.5 run on the EURO-CORDEX domain and the 0.11 RCP8.5 and 0.44 RCP4.5 on the Med-CORDEX domain, thus preventing any resolution, domain position or scenario comparison.

A 3-way analysis of variance using full factorial GCM-RCM-Scenario sub-ensembles within ESCENA ($4 \times 2 \times 2$) and EURO-CORDEX ($4 \times 2 \times 2$) shows conflicting results due to the very partial representativity of the full matrix. No significant interaction is found for either two- or three-factor interaction terms. However, in ESCENA, the only significant (confidence above 99.99%) main factor is the RCM, while in EURO-CORDEX the GCM shows the main significant effect on summer temperature deltas (above 99.99% confidence), along with the emissions scenario (above 95% confidence).

The problem to fill a GCM-RCM matrix is not the computational cost. A single EURO-CORDEX 0.11-resolution simulation requires, at least, 8 times as much computer resources as a 0.22 ENSEMBLES simulation. The full GCM-RCM matrix from ENSEMBLES (or an updated version using CMIP5 input and the latest generation RCMs) consists of 143 (11×13) simulations, accounting for an effort of less than 18 EURO-CORDEX 0.11 simulations (of which, at least, 42 have been produced with little chance for systematic comparison). At a much lower cost, a full 0.44-resolution matrix could have been filled.

There are some examples of ensemble design out of Europe: for instance, the NARCliM project in Australia (Evans et al, 2014) or the NARCCAP program in North America (Mearns et al, 2013). These two examples use different design criteria to maximize the resulting information according to practical limitations defined in each case. The selection of models in NARCliM is based on different evaluation metrics and model independence, although some subjective choices were also required. In case of NARCCAP, the selection of 24 nesting combinations was based on a balanced RCM-GCM design.

Up to now, there is no multi-RCM/multi-GCM experiment where all cross-combinations have been simulated (full matrix). In this situation, the effect of having relatively empty GCM-RCM matrices to explore uncertainties cannot be quantified. There are several attempts in the literature to use non-designed en-

sembles to explore uncertainties (Déqué et al, 2012). Strong assumptions on the independence of the contribution of different uncertainty sources need to be taken.

5.2 Observational constraints

Observational constraints are increasingly used to reduce the uncertainty in future projections. In essence, the idea is to use a plot such as Figure 5, along with the observed anomaly (it would be a vertical line in this Figure, showing the observed temperature anomaly with respect to the coldest simulation) to produce a distribution of future delta changes conditional on the observations. This is a sort of model weighting, favouring the models which better reproduce the reference climate for the particular variable used as abscissa.

Of course, this introduces plenty of new uncertainties. The abscissa (predictor) should be able to provide a relationship as clear as possible with the delta change being corrected. Figure 5 is probably not enough and introduces more errors than improvements in the delta. The predictor should also be a densely observed variable (to be able to produce an observational product comparable to the grid-cell output of a model), and ideally with a low observational uncertainty (otherwise, the vertical line becomes a wide distribution, and the conditional delta distribution tends back to that of the unweighted ensemble). Up to now, observational constraints have relied on predictors such as: temperature trends (Boberg and Christensen, 2012), the seasonal cycle of snow albedo (Hall and Qu, 2006), precipitation from previous seasons (Quesada et al, 2012), sensible heat flux (Stegehuis et al, 2012), or soil moisture (Bellprat et al, 2013). There is no overall solution, and these procedures work only where and when the relationship found is strong enough.

In any case, these works illustrate that the assumption that delta change cancels out model error is ill posed. Moreover, observational constraints are likely to favour models which reproduce the right predictor for the wrong reason (Räisänen, 2007). For instance, radiation biases can easily be compensated by soil moisture to produce reasonable surface temperatures (García-Díez et al, 2015). Therefore, model evaluation and future projection analyses should extend beyond temperature and precipitation to include variables controlling key surface, radiative and dynamical processes that drive the changes. Data availability, both from models and observations, is a challenge for this kind of analyses, which are a natural continuation of this basic study.

5.3 Earth System Grid Federation

The ESGF infrastructure (Williams et al, 2015) is a breakthrough in climate data availability and it makes possible unprecedented analyses of multi-model ensembles of climate projections, like the one presented in this work. Based on the experience retrieving these data, we found that the ESGF data discovery services are powerful, but the access and retrieval is still cumbersome because of the file granularity of the data. Accessing and retrieving datasets effectively requires some advanced skills in data analysis and management, including metadata standards and file formats like NetCDF, and handling a huge batch of datasets.

Despite the strong quality controls on metadata prior to data publishing, there is a lack of quality control regarding the match of data and metadata (e.g. units not matching the data order of magnitude) or the completeness of the dataset, either in time (missing years) or variables (despite mandatory variable lists). Many of such mismatches have been reported back to the modelling teams during the development of this study.

These metadata mismatches take time to realize and correct for every user, and mine the core of the ESGF concept of providing homogeneous access to data. Times or variables missing prevent models to enter studies, thus artificially weighting the ensemble. As an example, the direct output for EC-EARTH-r1 RCP4.5 was missing for years 2027 and 2029 and seasonal means used in this work did not consider these years. Currently, one of the priorities for CMIP6 data distribution is the deployment of a new errata service in ESGF to provide a central repository for reporting and accessing documentation related to problems with the data.

ESGF data un-publishing also poses some challenges on reproducibility. During the development of this study, data from several RCMs were un-published. These simulations were excluded from this study, but it is likely that similar problems arise in the future and, then, the work will no longer be reproducible. Also, a mean to link ESGF entries with scientific literature analysing them would ease quality assessment by data users and scientific development. Related to this, Hewitson et al (2017) have recently shown some key issues in climate information websites, including ESGF, highlighting the need to include user guidance to overcome the confusion arising from the varied messages of different data products. This relevant issue is part of the current priorities for future developments of ESGF metadata services. They include persistent identifiers to each dataset to trace its version history, an early citation service or some provenance capture to ease the traceability and reproducibility of results. See Williams et al (2017) for more details on the current status and developments of ESGF.

5.4 Sub-ensemble selection

Users of future climate change projections are commonly advised to use as many projections as possible, in order to sample the strong uncertainties present. This advise is hardly followed, and recent analyses show that the sub-ensembles selected by users of climate information can be near-random and not fit for their purpose (Hewitson et al, 2017).

Even though ESGF homogenizes the access to data, data availability can still be an issue. There are now dozens of RCM projections over Europe and more are still to come. Studies usually rely on the first models publicly available, leading to a race to publish data, ultimately driven by IPCC assessment report cycles, which drive CMIPs, which drive CORDEX and regional climate change studies.

Despite the diversity of publicly available projections, there are plenty of recent works still relying on a few models (Bavay et al, 2013; Ruelland et al, 2015) or even just one (Tramblay et al, 2013; Lemesios et al, 2016). Burke et al (2015) show the underestimation of the range of projected climate impacts in 7 well-cited articles that failed to account for multi-model uncertainty. Commonly, there is no mention of the motivation behind the selection of a particular sub-ensemble (e.g. Räisänen and Räty, 2013; Filahi et al, 2017). The motivation for particular

sub-ensembles usually includes (1) good performance in current climate against observations (e.g. Herrera et al, 2010) or (2) span a large range of possible future scenarios (e.g. Bavay et al, 2013). Both criteria can hardly be met simultaneously, given that the largest spread usually comes from outliers (Figure 5), and opposite outliers cannot simultaneously fit the observations.

The assumption that good performing models under current climate will project better climate changes can easily be challenged due to differing relevance of parameterized processes in the present and in the future (Jerez et al, 2013) and a potential good performance due to the wrong reason (Räisänen, 2007; García-Díez et al, 2015). In a recent work, Monerie et al (2016) show that present-day biases in precipitation and temperature are not good metrics for the credibility of future projections and propose other methods based on clustering.

Sub-ensemble selection is an open and complicated problem. Unfortunately, the alternative –use all projections available (see e.g. Burke et al, 2015)– is no better: as we show in this study, the full ensemble is dominated by somewhat arbitrary decisions on which GCMs were downscaled or which RCMs produced more projections, thus weighting the full ensemble towards particular GCM/RCM deficiencies. Our recommendation would be to cherry-pick GCM/RCMs based on their ability to represent the process of interest in current climate. Not just the variable of interest (e.g. precipitation or temperature) but the mechanisms behind their variability (circulation, water/energy budgets, ...). The credibility of the dynamical downscaling response (Section 4.4) can also be used as guidance.

6 Summary and conclusions

We presented an ensemble of future climate change projections of unprecedented size (196 members) combining several dynamical downscaling projects and initiatives. All data used are publicly available and include not only the high resolution RCM projections, but also their driving GCMs. For simplicity, this initial study has been limited to precipitation and temperature changes in the near future period 2021-2050. The target region selected is continental Spain.

The GCM-RCM combination matrix is sparse and all initiatives, except ES-CENA, lacked a systematic design to explore the uncertainty sources (GCM, RCM, scenario, resolution, ...) and their relative contribution to the total uncertainty range. In this scenario, we avoided summary statistics and favoured graphical displays of the full ensemble.

Near-future projections of precipitation and temperature essentially agree across the different initiatives, based on different model generations. The summer season shows the largest differences across initiatives and we focused on it. There is a common tendency in the different initiatives to project smaller temperature changes by the RCMs than their driving GCMs. This is partly explained by an uneven downscaling, favouring low climate sensitivity GCMs, and some particularly cold-biased RCMs (WRF, MM5).

The main contributor to the uncertainty range in this area is the GCM, which dominates the resulting delta changes. This agrees with previous studies (Déqué et al, 2012; Rajczak et al, 2013), although they establish that this result depends on the region (southwestern Europe), and also on the season (Christensen and

Christensen, 2007; Mearns et al, 2013). We also found particular full factorial sub-ensembles (e.g. in ESCENA [CNRM-CM,ECHAM5]×[PROMES,MM5]×[A1B,B1]) where the RCM dominates the uncertainty range. Emissions scenario and RCM resolution have a much lower contribution to uncertainty; the former should take prevalence on longer-term projections (Hawkins and Sutton, 2009).

This large ensemble shows a clear relation between projected temperature changes and present-day temperature biases, indicating an overestimation of the uncertainty range by the standard delta method. Several methods have been proposed in the literature to alleviate this problem (Hall and Qu, 2006; Quesada et al, 2012; Boberg and Christensen, 2012; Stegehuis et al, 2012; Bellprat et al, 2013) and this ensemble is an ideal testbed for their comparison. However, this is out of the scope of the current study.

Focusing on specific projections, we found several conflicting results between RCM projections and their driving GCM, and between RCMs nested to the same GCM. These conflicts need to be further analysed to identify the causes behind them, and discern if they consist of added value of downscaling and a genuine contribution to uncertainty. We checked one of these conflicts, linking it to unrealistic summer snow cover on the Pyrenees, but a full identification of the causes behind conflicting messages is beyond this study.

Finally, we discuss on the process to obtain and homogenize the large ensemble of projections used in the study, which took a large amount of time, unveiling the potential of the ESGF, but also drawbacks which might affect the reproducibility of results. This initial study just covers an overview of the potential of this multi-initiative ensemble, which can feed many future studies, e.g. on the impact of sub-ensemble selection strategies, observational constrain of the projected uncertainty, explanation of conflicting messages, etc., which are currently planned and ongoing.

References

Ayar PV, Vrac M, Bastin S, Carreau J, Déqué M, Gallardo C (2016) Intercomparison of statistical and dynamical downscaling models under the EURO- and MED-CORDEX initiative framework: present climate evaluations. Clim Dyn 46(3-4):1301–1329, DOI 10.1007/s00382-015-2647-5

Bavay M, Grünewald T, Lehning M (2013) Response of snow cover and runoff to climate change in high Alpine catchments of Eastern Switzerland. Advances in Water Resources 55:4–16, DOI 10.1016/j.advwatres.2012.12.009

Bellprat O, Kotlarski S, Lüthi D, Schär C (2013) Physical constraints for temperature biases in climate models. Geophysical Research Letters 40(15):4042–4047, DOI 10.1002/grl.50737, URL http://dx.doi.org/10.1002/grl.50737

Boberg F, Christensen JH (2012) Overestimation of Mediterranean summer temperature projections due to model deficiencies. Nature Clim Change 2(6):433–436, DOI 10.1038/nclimate1454

Boberg F, Berg P, Thejll P, Gutowski WJ, Christensen JH (2009) Improved confidence in climate change projections of precipitation further evaluated using daily statistics from ENSEMBLES models. Clim Dyn 35(7-8):1509–1520, DOI 10.1007/s00382-009-0683-8

Burke M, Dykema J, Lobell DB, Miguel E, Satyanath S (2015) Incorporating climate uncertainty into estimates of climate change impacts. The Review of Economics and Statistics 97(2):461–471, DOI 10.1162/REST_a_00478

Casanueva A, Bedia J, Herrera S, Fernández J, Gutiérrez J (2017) Bias correction of multi-variate climate indices: A new insight into the effect on the climate change signal. Climatic Change (Submitted)

Christensen JH, Christensen OB (2007) A summary of the PRUDENCE model projections of changes in European climate by the end of this century. Climatic Change 81(1):7–30, DOI 10.1007/s10584-006-9210-7

Daron JD, Stainforth DA (2013) On predicting climate under climate change. Environmental Research Letters 8(3):034,021, DOI 10.1088/1748-9326/8/3/034021

Dell'Aquila A, Mariotti A, Bastin S, Calmanti S, Cavicchia L, Deque M, Djurdjevic V, Dominguez M, Gaertner M, Gualdi S (2016) Evaluation of simulated decadal variations over the Euro-Mediterranean region from ENSEMBLES to Med-CORDEX. Climate Dynamics pp 1–20, DOI 10.1007/s00382-016-3143-2

Déqué M, Somot S, Sánchez-Gómez E, Goodess CM, Jacob D, Lenderink G, Christensen OB (2012) The spread amongst ENSEMBLES regional scenarios: regional climate models, driving general circulation models and interannual variability. Clim Dyn 38(5-6):951–964, DOI 10.1007/s00382-011-1053-x

Dessai S (2003) Heat stress and mortality in Lisbon Part II. An assessment of the potential impacts of climate change. Int J Biometeorol 48(1):37–44, DOI 10.1007/s00484-003-0180-4

Diaconescu EP, Laprise R (2013) Can added value be expected in RCM-simulated large scales? Climate Dynamics 41(7):1769–1800, DOI 10.1007/s00382-012-1649-9

Domínguez M, Romera R, Sánchez E, Fita L, Fernández J, Jiménez-Guerrero P, Montávez JP, Cabos WD, Liguori G, Gaertner MA (2013) Present-climate precipitation and temperature extremes over Spain from a set of high resolution RCMs. Clim Res 58(2):149–164, DOI 10.3354/cr01186

Easterling WE, Mearns LO, Hays CJ, Marx D (2001) Comparison of agricultural impacts of climate change calculated from high and low resolution climate change scenarios: Part II. Accounting for adaptation and CO2 direct effects. Climatic Change 51(2):173–197, DOI 10.1023/A:1012267900745

Evans JP, Ji F, Lee C, Smith P, Argüeso D, Fita L (2014) Design of a regional climate modelling projection ensemble experiment – NARCliM. Geosci Model Dev 7(2):621–629, DOI 10.5194/gmd-7-621-2014

Filahi S, Tramblay Y, Mouhir L, Diaconescu EP (2017) Projected changes in temperature and precipitation indices in Morocco from high-resolution regional climate models. International Journal of Climatology (In print), DOI 10.1002/joc.5127, URL http://dx.doi.org/10.1002/joc.5127

Flaounas E, Drobinski P, Vrac M, Bastin S, Lebeaupin-Brossier C, Stéfanon M, Borga M, Calvet JC (2013) Precipitation and temperature space–time variability and extremes in the Mediterranean region: evaluation of dynamical and statistical downscaling methods. Clim Dyn 40(11-12):2687–2705, DOI 10.1007/s00382-012-1558-y

García-Díez M, Fernández J, Vautard R (2015) An RCM multi-physics ensemble over Europe: multi-variable evaluation to avoid error compensation. Clim Dyn 45(11-12):3141–3156, DOI 10.1007/s00382-015-2529-x

Giannini V, Bellucci A, Torresan S (2016) Sharing skills and needs between providers and users of climate information to create climate services: lessons from the Northern Adriatic case study. Earth Perspectives 3(1), DOI 10.1186/s40322-016-0033-z

Giorgi F, Gutowski WJ (2015) Regional dynamical downscaling and the CORDEX initiative. Annual Review of Environment and Resources 40(1):467–490, DOI 10.1146/annurev-environ-102014-021217

Giorgi F, Diffenbaugh NS, Gao XJ, Coppola E, Dash SK, Frumento O, Rauscher SA, Remedio A, Sanda IS, Steiner A, Sylla B, Zakey AS (2008) The regional climate change hyper-matrix framework. Eos, Transactions American Geophysical Union 89(45):445–446, DOI 10.1029/2008EO450001

Gobiet A, Kotlarski S, Beniston M, Heinrich G, Rajczak J, Stoffel M (2014) 21st century climate change in the European Alps—A review. Science of The Total Environment 493:1138–1151, DOI 10.1016/j.scitotenv.2013.07.050

Gobiet A, Suklitsch M, Heinrich G (2015) The effect of empirical-statistical correction of intensity-dependent model errors on the temperature climate change signal. Hydrol Earth Syst Sci 19(10):4055–4066, DOI 10.5194/hess-19-4055-2015

Gómez G, Cabos WD, Liguori G, Sein D, Lozano-Galeana S, Fita L, Fernández J, Magariño ME, Jiménez-Guerrero P, Montávez JP, Domínguez M, Romera R, Gaertner MA (2016) Characterization of the wind speed variability and future change in the Iberian Peninsula and the Balearic Islands. Wind Energy 19(7):1223–1237, DOI 10.1002/we.1893, wE-14-0271.R2

Grotch SL, MacCracken MC (1991) The use of general circulation models to predict regional climatic change. J Climate 4(3):286–303, DOI 10.1175/1520-0442(1991)004⟨0286:TUOGCM⟩2.0.CO;2

Gutiérrez JM, San-Martín D, Brands S, Manzanas R, Herrera S (2012) Reassessing statistical downscaling techniques for their robust application under climate change conditions. J Climate 26(1):171–188, DOI 10.1175/JCLI-D-11-00687.1

Hall A (2014) Projecting regional change. Science 346(6216):1461–1462, DOI 10.1126/science.aaa0629

Hall A, Qu X (2006) Using the current seasonal cycle to constrain snow albedo feedback in future climate change. Geophys Res Lett 33(3):L03,502, DOI 10.1029/2005GL025127

Hawkins E, Sutton R (2009) The potential to narrow uncertainty in regional climate predictions. Bull Amer Meteor Soc 90(8):1095–1107, DOI 10.1175/2009BAMS2607.1

Herrera S, Fita L, Fernández J, Gutiérrez JM (2010) Evaluation of the mean and extreme precipitation regimes from the ensembles regional climate multi-model simulations over spain. Journal of Geophysical Research: Atmospheres 115(D21):n/a–n/a, DOI 10.1029/2010JD013936, URL http://dx.doi.org/10.1029/2010JD013936

Hewitson B, Waagsaether K, Wohland J, Kloppers K, Kara T (2017) Climate information websites: an evolving landscape. Wiley Interdisciplinary Reviews: Climate Change pp e470–n/a, DOI 10.1002/wcc.470, URL http://dx.doi.org/10.1002/wcc.470

Hewitson BC, Daron J, Crane RG, Zermoglio MF, Jack C (2013) Interrogating empirical-statistical downscaling. Climatic Change 122:539–554, DOI 10.1007/s10584-013-1021-z

Ho CK, Stephenson DB, Collins M, Ferro CAT, Brown SJ (2012) Calibration strategies: A source of additional uncertainty in climate change projections. Bull Amer Meteor Soc 93(1):21–26, DOI 10.1175/2011BAMS3110.1

Huynen MMTE, Martens P (2015) Climate change effects on heat- and cold-related mortality in The Netherlands: A scenario-based integrated environmental health impact assessment. Int J Environ Res Public Health 12(10):13,295–13,320, DOI 10.3390/ijerph121013295

Jacob D, Petersen J, Eggert B, Alias A, Christensen OB, Bouwer LM, Braun A, Colette A, Déqué M, Georgievski G, Georgopoulou E, Gobiet A, Menut L, Nikulin G, Haensler A, Hempelmann N, Jones C, Keuler K, Kovats S, Kröner N, Kotlarski S, Kriegsmann A, Martin E, van Meijgaard E, Moseley C, Pfeifer S, Preuschmann S, Radermacher C, Radtke K, Rechid D, Rounsevell M, Samuelsson P, Somot S, Soussana JF, Teichmann C, Valentini R, Vautard R, Weber B, Yiou P (2014) EURO-CORDEX: new high-resolution climate change projections for European impact research. Reg Environ Change 14(2):563–578, DOI 10.1007/s10113-013-0499-2

Jerez S, Montávez JP, Jiménez-Guerrero P, Gómez-Navarro JJ, Lorente-Plazas R, Zorita E (2013) A multi-physics ensemble of present-day climate regional simulations over the Iberian Peninsula. Clim Dyn 40(11-12):3023–3046, DOI 10.1007/s00382-012-1539-1

Jiménez-Guerrero P, Montávez JP, Domínguez M, Romera R, Fita L, Fernández J, Cabos WD, Liguori G, Gaertner MA (2013) Mean fields and interannual variability in RCM simulations over Spain: the ESCENA project. Clim Res 57(3):201–220, DOI 10.3354/cr01165

Kalkstein LS, Greene JS (1997) An evaluation of climate/mortality relationships in large U.S. cities and the possible impacts of a climate change. Environmental Health Perspectives 105(1):84–93, DOI 10.2307/3433067

Kis A, Pongrácz R, Bartholy J (2017) Multi-model analysis of regional dry and wet conditions for the Carpathian region. International Journal of Climatology pp n/a–n/a, DOI 10.1002/joc.5104, URL http://dx.doi.org/10.1002/joc.5104

Kjellström E, Boberg F, Castro M, Christensen J, Nikulin G, Sánchez E (2010) Daily and monthly temperature and precipitation statistics as performance indicators for regional climate models. Climate Research 44(2-3):135–150, DOI 10.3354/cr00932

Kjellström E, Nikulin G, Hansson U, Strandberg G, Ullerstig A (2011) 21st century changes in the European climate: uncertainties derived from an ensemble of regional climate model simulations. Tellus A 63(1):24–40, DOI 10.1111/j.1600-0870.2010.00475.x

Knist S, Goergen K, Buonomo E, Christensen OB, Colette A, Cardoso RM, Fealy R, Fernández J, García-Díez M, Jacob D, Kartsios S, Katragkou E, Keuler K, Mayer S, van Meijgaard E, Nikulin G, Soares PMM, Sobolowski S, Szepszo G, Teichmann C, Vautard R, Warrach-Sagi K, Wulfmeyer V, Simmer C (2017) Land-atmosphere coupling in euro-cordex evaluation experiments. Journal of Geophysical Research: Atmospheres 122(1):79–103, DOI 10.1002/2016JD025476, URL http://dx.doi.org/10.1002/2016JD025476, 2016JD025476

Knowlton K, Lynn B, Goldberg RA, Rosenzweig C, Hogrefe C, Rosenthal JK, Kinney PL (2007) Projecting heat-related mortality impacts under a changing climate in the New York city region. Am J Public Health 97(11):2028–2034, DOI 10.2105/AJPH.2006.102947

Kotlarski S, Keuler K, Christensen OB, Colette A, Déqué M, Gobiet A, Goergen K, Jacob D, Lüthi D, van Meijgaard E, Nikulin G, Schär C, Teichmann C, Vautard R, Warrach-Sagi K, Wulfmeyer V (2014) Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. Geoscientific Model Development Discussions 7(1):217–293, DOI 10.5194/gmdd-7-217-2014

Koutroulis AG, Grillakis MG, Tsanis IK, Jacob D (2015) Exploring the ability of current climate information to facilitate local climate services for the water sector. Earth Perspectives 2(1), DOI 10.1186/s40322-015-0032-5

Laprise R (2008) Regional climate modelling. Journal of Computational Physics 227(7):3641–3666, DOI 10.1016/j.jcp.2006.10.024

Lautenschlager M, Adamidis P, Kuhn M (2015) Big data research at DKRZ – climate model data production workflow. In: Big Data and High Performance Computing, IOS Press, pp 133 – 155

Lemesios G, Giannakopoulos C, Papadaskalopoulou C, Karali A, Varotsos KV, Moustakas K, Malamis D, Zachariou-Dodou M, Petrakis M, Loizidou M (2016) Future heat-related climate change impacts on tourism industry in Cyprus. Reg Environ Change pp 1–13, DOI 10.1007/s10113-016-0997-0

Lémond J, Dandin P, Planton S, Vautard R, Pagé C, Déqué M, Franchistéguy L, Geindre S, Kerdoncuff M, Li L, Moisselin JM, Noël T, Tourre YM (2011) DRIAS: a step toward Climate Services in France. Adv Sci Res 6(1):179–186, DOI 10.5194/asr-6-179-2011

Liguori G, Di Lorenzo E, Cabos W (2017) A multi-model ensemble view of winter heat flux dynamics and the dipole mode in the mediterranean sea. Climate Dynamics 48(3):1089–1108, DOI 10.1007/s00382-016-3129-0

Maraun D, Wetterhall F, Ireson AM, Chandler RE, Kendon EJ, Widmann M, Brienen S, Rust HW, Sauter T, Themeßl M, Venema VKC, Chun KP, Goodess CM, Jones RG, Onof C, Vrac M, Thiele-Eich I (2010) Precipitation downscaling under climate change: Recent developments to bridge the gap be-

tween dynamical models and the end user. Rev Geophys 48(3):RG3003, DOI 10.1029/2009RG000314

Maraun D, Widmann M, Gutiérrez JM, Kotlarski S, Chandler RE, Hertig E, Wibig J, Huth R, Wilcke RA (2015) VALUE: A framework to validate downscaling approaches for climate change studies. Earth's Future 3:2014EF000,259, DOI 10.1002/2014EF000259

Mearns LO, Sain S, Leung LR, Bukovsky MS, McGinnis S, Biner S, Caya D, Arritt RW, Gutowski W, Takle E, Snyder M, Jones RG, Nunes AMB, Tucker S, Herzmann D, McDaniel L, Sloan L (2013) Climate change projections of the North American Regional Climate Change Assessment Program (NARCCAP). Climatic Change 120(4):965–975, DOI 10.1007/s10584-013-0831-3

Meehl GA, Covey C, Taylor KE, Delworth T, Stouffer RJ, Latif M, McAvaney B, Mitchell JFB (2007) The WCRP CMIP3 multimodel dataset: A new era in climate change research. Bull Amer Meteor Soc 88(9):1383–1394, DOI 10.1175/BAMS-88-9-1383

Monerie PA, Sanchez-Gomez E, Boé J (2016) On the range of future Sahel precipitation projections and the selection of a sub-sample of CMIP5 models for impact studies. Clim Dyn pp 1–20, DOI 10.1007/s00382-016-3236-y

Murphy JM, Sexton DMH, Jenkins GJ, Booth BBB, Brown CC, Clark RT, Collins M, Harris GR, Kendon EJ, Betts RA, Brown SJ, Humphrey KA, McCarthy MP, McDonald RE, Stephens A, Wallace C, Warren R, Wilby R, Wood RA (2009) UK climate projections science report: climate change projections. Meteorological Office Hadley Centre, Exeter, UK

Ouzeau G, Soubeyroux JM, Schneider M, Vautard R, Planton S (2016) Heat waves analysis over France in present and future climate: Application of a new method on the EURO-CORDEX ensemble. Climate Services 4:1 – 12, DOI https://doi.org/10.1016/j.cliser.2016.09.002

Palmer TN (2000) Predicting uncertainty in forecasts of weather and climate. Rep Prog Phys 63(2):71, DOI 10.1088/0034-4885/63/2/201

Pepin N, Bradley RS, Diaz HF, Baraer M, Caceres EB, Forsythe N, Fowler H, Greenwood G, Hashmi MZ, Liu XD, Miller JR, Ning L, Ohmura A, Palazzi E, Rangwala I, Schöner W, Severskiy I, Shahgedanova M, Wang MB, Williamson SN, Yang DQ (2015) Elevation-dependent warming in mountain regions of the world. Nature Climate Change 5(5):424–430

Plieger M, Som de Cerff W, Pagé C, Tatarinova N, Cofiño A, Vega Saldarriaga M, Hutjes R, de Jong F, Bärring L, Sjökvist E (2015) The climate4impact portal: bridging the CMIP5 and CORDEX data infrastructure to impact users. In: EGU General Assembly Conference Abstracts, EGU General Assembly Conference Abstracts, vol 17, p 9721

Quesada B, Vautard R, Yiou P, Hirschi M, Seneviratne SI (2012) Asymmetric European summer heat predictability from wet and dry southern winters and springs. Nature Clim Change 2(10):736–741, DOI 10.1038/nclimate1536

Räisänen J (2007) How reliable are climate models? Tellus A 59(1):2–29, DOI 10.1111/j.1600-0870.2006.00211.x

Räisänen J, Räty O (2013) Projections of daily mean temperature variability in the future: cross-validation tests with ENSEMBLES regional climate simulations. Clim Dyn 41(5-6):1553–1568, DOI 10.1007/s00382-012-1515-9

Rajczak J, Pall P, Schär C (2013) Projections of extreme precipitation events in regional climate simulations for Europe and the Alpine Region. Journal of

Geophysical Research: Atmospheres 118(9):3610–3626, DOI 10.1002/jgrd.50297, URL http://dx.doi.org/10.1002/jgrd.50297

Rios-Entenza A, Soares PMM, Trigo RM, Cardoso RM, Miguez-Macho G (2014) Moisture recycling in the Iberian Peninsula from a regional climate simulation: Spatiotemporal analysis and impact on the precipitation regime. J Geophys Res Atmos 119(10):2013JD021,274, DOI 10.1002/2013JD021274

Rötter RP, Höhn J, Trnka M, Fronzek S, Carter TR, Kahiluoto H (2013) Modelling shifts in agroclimate and crop cultivar response under climate change. Ecol Evol 3(12):4197–4214, DOI 10.1002/ece3.782

Ruelland D, Hublart P, Tramblay Y (2015) Assessing uncertainties in climate change impacts on runoff in Western Mediterranean basins. In: Vaze J, Chiew F, Hughes D, Andreassian V (eds) Hydrologic Non-Stationarity and Extrapolating Models to Predict the Future, vol 371, Copernicus Gesellschaft Mbh, Gottingen, pp 75–81, wOS:000365021100014

Rummukainen M (2010) State-of-the-art with regional climate models. WIREs Clim Change 1(1):82–96, DOI 10.1002/wcc.8

Rummukainen M (2016) Added value in regional climate modeling. WIREs Clim Change 7(1):145–159, DOI 10.1002/wcc.378

Ruti PM, Somot S, Giorgi F, Dubois C, Flaounas E, Obermann A, Dell'Aquila A, Pisacane G, Harzallah A, Lombardi E, Ahrens B, Akhtar N, Alias A, Arsouze T, Aznar R, Bastin S, Bartholy J, Béranger K, Beuvier J, Bouffies-Cloché S, Brauch J, Cabos W, Calmanti S, Calvet JC, Carillo A, Conte D, Coppola E, Djurdjevic V, Drobinski P, Elizalde-Arellano A, Gaertner M, Galàn P, Gallardo C, Gualdi S, Goncalves M, Jorba O, Jordà G, L'Heveder B, Lebeaupin-Brossier C, Li L, Liguori G, Lionello P, Maciàs D, Nabat P, Önol B, Raikovic B, Ramage K, Sevault F, Sannino G, Struglia MV, Sanna A, Torma C, Vervatis V (2016) Med-CORDEX initiative for Mediterranean climate studies. Bull Amer Meteor Soc 97(7):1187–1208, DOI 10.1175/BAMS-D-14-00176.1

Rössler O, Fischer AM, Huebener H, Maraun D, Benestad RE, Christodoulides P, Soares PMM, Cardoso RM, Pagé C, Kanamaru H, Kreienkamp F, Vlachogiannis D (2017) Challenges to link climate change data provision and user needs – perspective from the cost-action value. International Journal of Climatology DOI 10.1002/joc.5060

San-Martín D, Manzanas R, Brands S, Herrera S, Gutiérrez JM (2017) Reassessing model uncertainty for regional projections of precipitation with an ensemble of statistical downscaling methods. Journal of Climate 30(1):203–223, DOI 10.1175/JCLI-D-16-0366.1

Seneviratne SI, Corti T, Davin EL, Hirschi M, Jaeger EB, Lehner I, Orlowsky B, Teuling AJ (2010) Investigating soil moisture–climate interactions in a changing climate: A review. Earth-Science Reviews 99(3–4):125–161, DOI 10.1016/j.earscirev.2010.02.004

Sima M, Popovici EA, Bălteanu D, Micu DM, Kucsicsa G, Dragotă C, Grigorescu I (2015) A farmer-based analysis of climate change adaptation options of agriculture in the Bărăgan Plain, Romania. Earth Perspectives 2(1), DOI 10.1186/s40322-015-0031-6

Soares PMM, Cardoso RM, Ferreira JJ, Miranda PMA (2014) Climate change and the Portuguese precipitation: ENSEMBLES regional climate models results. Clim Dyn pp 1–17, DOI 10.1007/s00382-014-2432-x

Stéfanon M, Martin-StPaul NK, Leadley P, Bastin S, Dell'Aquila A, Drobinski P, Gallardo C (2015) Testing climate models using an impact model: what are the advantages? Climatic Change 131(4):649–661, DOI 10.1007/s10584-015-1412-4

Stegehuis AI, Vautard R, Ciais P, Teuling AJ, Jung M, Yiou P (2012) Summer temperatures in Europe and land heat fluxes in observation-based data and regional climate model simulations. Clim Dyn 41(2):455–477, DOI 10.1007/s00382-012-1559-x

Stegehuis AI, Teuling AJ, Ciais P, Vautard R, Jung M (2013) Future European temperature change uncertainties reduced by using land heat flux observations. Geophysical Research Letters 40(10):2242–2245, DOI 10.1002/grl.50404, URL http://dx.doi.org/10.1002/grl.50404

Stocker T, Qin D, Plattner G, Tignor M, Allen S, Boschung J, Nauels A, Xia Y, Bex V, Midgley P (eds) (2013) Climate change 2013 - The physical science basis: Working Group I contribution to the fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge

Tank A, Beersma J, Bessembinder J, van den Hurk B, Lenderink G (2015) KNMI, 2015: KNMI'14 Climate scenarios the Netherlands; A guide for professionales in climate adaptation. Tech. rep., KNMI, De Bilt, The Netherlands

Taylor KE, Stouffer RJ, Meehl GA (2011) An overview of CMIP5 and the experiment design. Bull Amer Meteor Soc 93(4):485–498, DOI 10.1175/BAMS-D-11-00094.1

Terry K (2014) Analysis and prototyping of a climate data manager. Master's thesis, Science School. Universidad de Cantabria, URL https://repositorio.unican.es/xmlui/handle/10902/4485

Tolika CK, Zanis P, Anagnostopoulou C (2012) Regional climate change scenarios for Greece: future temperature and precipitation projections from ensembles of RCMs. Global NEST Journal 14(4):407–421

Tramblay Y, Ruelland D, Somot S, Bouaicha R, Servat E (2013) High-resolution Med-CORDEX regional climate model simulations for hydrological impact studies: a first evaluation of the ALADIN-Climate model in Morocco. Hydrol Earth Syst Sci 17(10):3721–3739, DOI 10.5194/hess-17-3721-2013

Turco M, Sanna A, Herrera S, Llasat MC, Gutiérrez JM (2013) Large biases and inconsistent climate change signals in ensembles regional projections. Climatic Change 120(4):859–869, DOI 10.1007/s10584-013-0844-y

Turco M, Llasat MC, Herrera S, Gutiérrez JM (2017) Bias correction and downscaling of future RCM precipitation projections using a MOS-Analog technique. Journal of Geophysical Research: Atmospheres 122(5):2631–2648, DOI 10.1002/2016JD025724

van der Linden P, Mitchell J (2009) ENSEMBLES: Climate change and its impacts: Summary of research and results from the ENSEMBLES project. Tech. rep., Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK.

Vautard R, Gobiet A, Jacob D, Belda M, Colette A, Déqué M, Fernández J, García-Díez M, Goergen K, Güttler I, Halenka T, Karacostas T, Katragkou E, Keuler K, Kotlarski S, Mayer S, van Meijgaard E, Nikulin G, Patarčić M, Scinocca J, Sobolowski S, Suklitsch M, Teichmann C, Warrach-Sagi K, Wulfmeyer V, Yiou P (2013) The simulation of European heat waves from an ensemble of regional climate models within the EURO-CORDEX project. Climate Dynamics 41(9):2555–2575, DOI 10.1007/s00382-013-1714-z

von Storch H, Zorita E, Cubasch U (1993) Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime. J Climate 6(6):1161–1171, DOI 10.1175/1520-0442(1993)006⟨1161:DOGCCE⟩2.0.CO;2

van Vuuren DP, Edmonds J, Kainuma M, Riahi K, Thomson A, Hibbard K, Hurtt GC, Kram T, Krey V, Lamarque JF, Masui T, Meinshausen M, Nakicenovic N, Smith SJ, Rose SK (2011) The representative concentration pathways: an overview. Climatic Change 109(1):5, DOI 10.1007/s10584-011-0148-z

Wilby RL, Wigley TML (1997) Downscaling general circulation model output: a review of methods and limitations. Progress in Physical Geography 21(4):530–548, DOI 10.1177/030913339702100403

Williams DN, Balaji V, Cinquini L, Denvil S, Duffy D, Evans B, Ferraro R, Hansen R, Lautenschlager M, Trenham C (2015) A global repository for planet-sized experiments and observations. Bull Amer Meteor Soc 97(5):803–816, DOI 10.1175/BAMS-D-15-00132.1

Williams DN, Lautenschlager M, Denvil S, Cinquini L, Ferraro R, Duffy D, Evans B, Trenham C (2017) 6th annual Earth System Grid Federation face-to-face conference report. DOE/SC-0188, U.S. Department of Energy Office of Science, DOI 10.2172/1369382, URL https://esgf.llnl.gov/media/pdf/2017-ESGF_F2F_Conference_Report.pdf

Xue Y, Janjic Z, Dudhia J, Vasic R, Sales FD (2014) A review on regional dynamical downscaling in intraseasonal to seasonal simulation/prediction and major factors that affect downscaling ability. Atmospheric Research 147–148:68 – 85, DOI https://doi.org/10.1016/j.atmosres.2014.05.001

Supplementary information for:

Consistency of climate change projections from multiple global and regional model intercomparison projects

J. Fernández, M. D. Frías, W. D. Cabos, A. S. Cofiño, M. Domínguez, L. Fita, M. A. Gaertner, M. García-Díez, J. M. Gutiérrez, P. Jiménez-Guerrero, G. Liguori, J. P. Montávez, R. Romera, E. Sánchez

Corresponding author: J. Fernández Grupo de Meteorología. Dpto. Matemática Aplicada y Ciencias de la Computación. Univ. de Cantabria, Avda. de los Castros, s/n, 39005 Santander, Spain
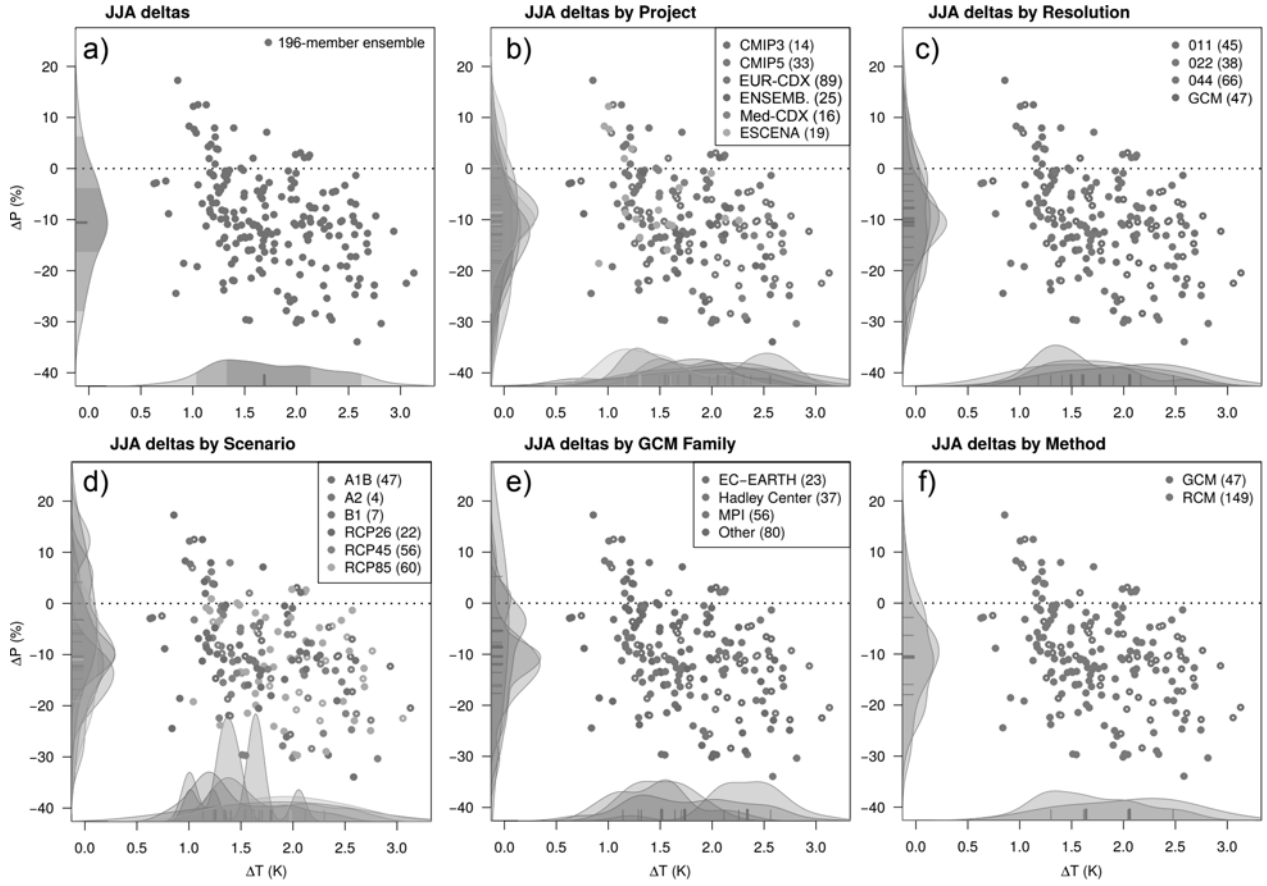jesus.fernandez@unican.es

Submitted to *Climate Dynamics*

Fig. ESM1 Precipitation vs mean surface temperature JJA delta changes spatially averaged over Continental Spain and the Balearic Islands for each ensemble member. (a) Deltas for the 196-member ensemble. (b) Deltas by Project. (c) Deltas by Resolution. (d) Deltas by Scenario. (e) Deltas by GCM Family. (f) Deltas by Method. Marginal probability density functions are shown for each variable pooling the whole 196-member ensemble for the different classification, using a Gaussian kernel density estimator. Tickmarks show the ranges from the 5th to 95th percentile (90% of the sample) and from the 1st to 3rd quartile (50%). The median is also shown as an inner tickmark (thicker). Raw GCM output deltas are shown as empty circles.
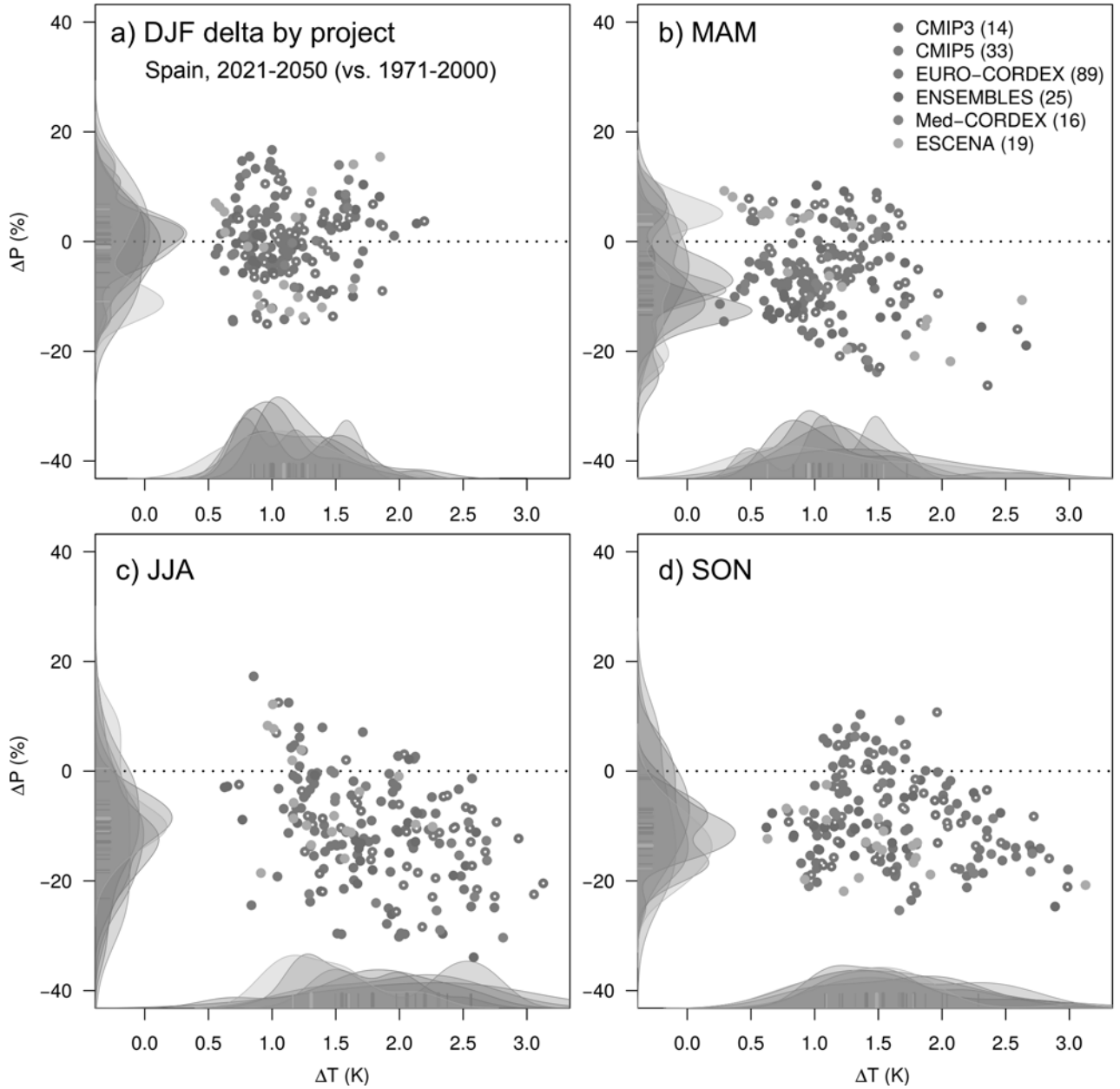
Fig. ESM2 Precipitation vs mean surface temperature delta changes spatially averaged over Continental Spain and the Balearic Islands for each ensemble member. Values classified by project. (a) DJF. (b) MAM. (c) JJA. (d) SON. Marginal probability density functions are shown for each variable and project pooling the whole 196-member ensemble, using a Gaussian kernel density estimator. Tickmarks show the ranges from the 5th to 95th percentile (90% of the sample) and from the 1st to 3rd quartile (50%). The median is also shown as a thicker inner tickmark.