



# Evaluation of a new 12 km regional perturbed parameter ensemble over Europe

Simon O. Tucker<sup>1</sup> · Elizabeth J. Kendon<sup>1</sup> · Nicolas Bellouin<sup>2</sup> · Erasmo Buonomo<sup>1</sup> · Ben Johnson<sup>1</sup> · James M. Murphy<sup>1</sup>

Received: 5 February 2021 / Accepted: 19 August 2021 / Published online: 30 August 2021  
© Crown 2021, corrected publication 2021

## Abstract

We evaluate a 12-member perturbed parameter ensemble of regional climate simulations over Europe at 12 km resolution, carried out as part of the UK Climate Projections (UKCP) project. This ensemble is formed by varying uncertain parameters within the model physics, allowing uncertainty in future projections due to climate modelling uncertainty to be explored in a systematic way. We focus on present day performance both compared to observations, and consistency with the driving global ensemble. Daily and seasonal temperature and precipitation are evaluated as two variables commonly used in impacts assessments. For precipitation we find that downscaling, even whilst within the convection-parameterised regime, generally improves daily precipitation, but not everywhere. In summer, the underestimation of dry day frequency is worse in the regional ensemble than in the driving simulations. For temperature we find that the regional ensemble inherits a large wintertime cold bias from the global model, however downscaling reduces this bias. The largest bias reduction is in daily winter cold temperature extremes. In summer the regional ensemble is cooler and wetter than the driving global models, and we examine cloud and radiation diagnostics to understand the causes of the differences. We also use a low-resolution regional simulation to determine whether the differences are a consequence of resolution, or due to other configuration differences, with the predominant configuration difference being the treatment of aerosols. We find that use of the EasyAerosol scheme in the regional model, which aims to approximate the aerosol effects in the driving model, causes reduced temperatures by around 0.5 K over Eastern Europe in Summer, and warming of a similar magnitude over France and Germany in Winter, relative to the impact of interactive aerosol in the global runs. Precipitation is also increased in these regions. Overall, we find that the regional model is consistent with the global model, but with a typically better representation of daily extremes and consequently we have higher confidence in its projections of their future change.

**Keywords** Aerosol modelling · Perturbed parameter ensembles · Regional climate change · UKCP18

## 1 Introduction

A major update to the UK Climate Projections was released in 2018 (UKCP18). This provides updated probabilistic projections, as well as new global and regional climate model projections, allowing an assessment of how the climate not just of the UK but also for Europe and globally may change over the twenty-first century. The UKCP18 projections are intended to help inform climate change risk assessments

and adaptation plans. The regional component of UKCP18 includes a 12-member ensemble of regional climate simulations over Europe at 12 km resolution, that downscale twelve 60 km Hadley Centre global model simulations. The different ensemble members differ due to perturbations applied to uncertain parameters in the model physics, with perturbations in the global models mirrored in the downscaling regional simulations. The regional simulations use the same 0.11° grid as the EURO-CORDEX (Jacob et al. 2014, 2020) simulations for the time period 1980–2080 using the RCP8.5 scenario (Moss et al. 2010). This paper evaluates the 12 km simulations for the present day period (1980–2000) only.

The UKCP18 regional 12 km projections replace the previous UKCP09 25 km regional ensemble projections (Murphy et al. n.d.), which have been used extensively to inform impacts studies. These include assessments of drought

✉ Simon O. Tucker  
Simon.tucker@metoffice.gov.uk

<sup>1</sup> Met Office Hadley Centre, Fitzroy road, Exeter EX1 3PB, UK

<sup>2</sup> Department of Meteorology, University of Reading, Reading RG6 6EG, UK

(Burke et al. 2010), river flows (Prudhomme et al. 2012), water availability (Sanderson et al. 2012), flood frequency (Kay and Jones 2012), and effects on the electricity and rail networks (McColl et al. 2012; Palin et al. 2013). Compared to UKCP09, UKCP18 provides projections at increased resolution (global from 300 to 60 km and regional from 25 to 12 km) and includes a number of model developments. These have led to an improved representation of regional climate dynamics (Scaife et al. 2012, 2014), including a good simulation of mid latitude synoptic variability (Williams et al. 2018).

The downscaled simulations are expected to provide more local detail due to the improved representation of surface features such as mountains and coastlines, and improved mesoscale dynamics, whilst being consistent with their driving global model over large time and spatial scales (e.g. Rummukainen 2010, 2016). Over Europe numerous studies compare pairs of 50 km and 12 km Euro-CORDEX simulations to each other. Kotlarski et al. (2014) found that for seasonal mean quantities averaged over large European subdomains, no clear benefit of an increased spatial resolution (12 km vs. 50 km) can be identified. Similarly, Vautard et al. (2013) report that heat waves are caused by large scale processes and hence little difference is found between the two resolutions. Meanwhile Prein et al. (2016) find that the 12 km simulations better reproduce mean and extreme precipitation for almost all regions and seasons, even on the scale of the coarser-gridded simulations (50 km), with the largest improvement seen over regions with substantial orographic features. Fantini et al. (2018) reinforce this finding, but also note the tendency for models to underestimate dry day frequency remains at 12 km. Similarly, Casanueva et al. (2016) calculate precipitation indices over both Spain and the Alps, and find that the spatial correlation with observations is higher in the 12 km simulations.

A perturbed parameter ensemble involves taking a single climate model and identifying plausible values for a set of uncertain model parameters. This parameter space can then be sampled in a systematic way. In contrast, initiatives such as EURO-CORDEX (Jacob et al. 2014, 2020) allow for a wider, but less systematic sampling of the differences between climate models (including structural differences due to different model architectures and parameterisation schemes) by downscaling a number of global climate models with a number of regional models, with each model having its own formulation. Impacts assessments can thus sample different types of climate model uncertainty by using data from both CORDEX and the Hadley Centre regional perturbed parameter ensemble presented here (hereafter RCM-PPE). Note that RCM-PPE includes the unperturbed RCM configuration (hereafter RCM-STD), and this configuration is also one of the RCMs used in CORDEX (where it has the standard name MOHC-HadREM3-GA7-05).

Another difference with the CORDEX multi model approach is that each regional model simulation in RCM-PPE has the same atmosphere and land surface configuration as its driving global model, including the same set of parameter perturbations. This is done to ensure that the RCM scenarios are as consistent as possible with the global simulations at large regional scales, and can help minimise mismatches between the lateral boundary conditions supplied by the global model, and the regional model (Davies 2014). There are some exceptions, where configuration differences between regional and global model were necessary, these are outlined in Sect. 2.2.2. Part of the analysis in this paper tries to quantify the impact that these configuration differences have.

This paper describes the RCM-PPE ensemble, examining its performance compared to observations, and its consistency with the driving model. The ensemble design and model description are described in Sect. 2. The results are presented in Sect. 3 and include an analysis of seasonal mean temperature and precipitation differences (Sect. 3.1)—some potential causes of which are discussed in Sect. 3.2—and daily extremes of precipitation and temperature (Sect. 3.3). We focus on temperature and precipitation as they are commonly used in impacts assessments and extremes are generally better represented in higher resolution regional models (e.g. Prein et al. 2016; Fantini et al. 2018; Torma et al. 2015). Section 4 discusses the implications of the results in terms of suitability for use in different applications and our relative confidence (compared to that of the driving GCMs) for making climate change projections.

## 2 Methods

### 2.1 Experimental design

Twelve global perturbed parameter simulations have been downscaled using the HadREM3-GA7-05 model over the EURO-CORDEX domain (Jacob et al. 2014) for the time period 1980–2080 using the RCP85 scenario. Each regional model simulation has the same atmosphere and land surface configuration as its driving global model (apart from aerosol modelling, see below), including the same set of parameter perturbations.

The selection of the 12 RCM-PPE members was carried out as follows:

1. Initially, a 25-member global PPE of coupled ocean–atmosphere model variants was produced. The 25 perturbed configurations were themselves chosen from a larger preliminary PPE using performance and diversity criteria, in order to increase the range of global and

regional changes sampled in the projections (Sexton et al. 2021; Murphy et al. 2018).

2. This set of 25 coupled variants was then filtered down to 20, with the 5 members excluded because their simulated climate was unrealistically cool by 1970, or because they suffered from numerical instabilities (Murphy et al. 2018; Yamazaki et al. 2021).
3. Due to cost constraints, 16 of these 20 were selected for downscaling by the RCM. This was done by selecting the standard (unperturbed) member, plus the four members with outlying aerosol forcing and climate feedback strength, plus the remaining 11 by maximising sampling of spread within PPE parameter space (Murphy et al. 2018).
4. Subsequent to this selection of 16, the global set was reduced by 5 following further assessment of biases in European climatology, Atlantic meridional overturning circulation (AMOC) strength and historical trends in northern hemisphere surface temperature. Of this set of 5 excluded members, four were in the set of 16 above that had been downscaled. Thus these four downscaled simulations were also excluded, resulting in an RCM-PPE of 12 members.

Note when we refer to GCM-PPE in this paper we mean specifically the 12 global models driving RCM-PPE (rather than the larger ensemble that members were selected to downscale from).

In order to facilitate understanding of model biases two additional simulations have been run. The first, hereafter ERAI-RCM-STD, was run from 1981 to 2002, and used the RCM-STD model configuration, but driven by quasi observed lateral boundary forcing taken from ERA-Interim reanalyses (Dee et al. 2011). Sea surface temperatures and sea-ice extents were also prescribed from analyses of observations (Reynolds et al. 2002). The second simulation, hereafter RCM-LOWRES was run for 10 years (1980–1989) with the same configuration as RCM-STD, but with a resolution of  $0.55^\circ$ , approximately the same resolution as GCM-STD ( $\sim 60$  km over Europe, see Sect. 2.2.2). Due to the larger grid spacing of RCM-LOWRES, the rim region is larger than that of RCM-STD, meaning that the internal domain is unintentionally smaller by around  $3^\circ$  on each side. The purpose of ERAI-RCM-STD is to assess the performance of the RCM when it is not inheriting errors from the driving GCM, whilst the purpose of RCM-LOWRES is to understand which differences between RCM and GCM are due to resolution. Additionally, an error has since been found, known as the daylight hours or DLH error, and we use a rerun of RCM-LOWRES with the error corrected (hereafter RCM-LOWRES-DLHFIX) to assess the impact. More details on the DLH error are provided later.

## 2.2 Model configuration and forcing

### 2.2.1 Model forcing

The simulations used CMIP5 observed historical forcings until 2005 and RCP8.5 thereafter (Moss et al. 2010). These included changes in well-mixed greenhouse gases, ozone, solar radiation, major volcanic eruptions, land use changes and for GCM-PPE natural and anthropogenic aerosol precursors.

Time-dependent changes in fractional coverage of land vegetation types were prescribed, according to the harmonised land-use reconstructions used in CMIP5 (Hurt et al. 2011). Following a similar approach to that taken in HadGEM2-AO (Baek et al. 2013), land use change is represented by applying a time varying anomaly to the GA7 standard (non-time-varying) land cover ancillary (IGBP (Loveland 2000) mapped to the nine tiles used in the JULES land surface scheme). Anomalies with respect to the year 1992 that IGBP represents, in total crop and pasture (Hurt et al. 2011) are mapped to changes in the combined coverage of C3 and C4 grass plant functional types (PFTs). Changes in fractional grass cover were compensated for by opposite changes to the PFTs that are assumed to represent natural undisturbed land: namely broadleaf trees, needleleaf trees and shrubs. Coverage of urban, soil, inland water and land ice classifications were kept unchanged.

Volcanic forcing up to year 2000 was prescribed from an observed estimate (Sato et al. 1993, updated). Subsequently, it followed a profile similar to that of Jones et al. (2011), ramping down to a low level by 2020, and then recovering to climatological values by 2040. Total solar irradiance was specified using the Lean et al. (2009) data to 2008 followed by a fixed 12-year cycle to 2100, obtained by continuously repeating the observations for 1996–2008.

Each RCM member has the same prescribed  $\text{CO}_2$  concentrations as its driving GCM member, however from 2005 concentrations vary between members. This was done to reflect the global effects of carbon cycle uncertainties on projected changes, further details are in Murphy et al. (2018). GCM-PPE members also sampled uncertainty in the observed emissions of sulphur dioxide, using a scaling factor ranging from 0.5 to 1.5 that constituted one of the perturbed parameters in GCM-PPE. The same scaling factors were applied to future emissions.

### 2.2.2 Model configurations

The GCM-PPE is based on the GC3.05 coupled configuration of the Met Office Hadley Centre ocean–atmosphere model. That is the GC3.0 configuration with a number of the changes that were included in GC3.1 (the Met Office model submitted to CMIP6) in order to reduce the strong negative

forcing due to anthropogenic aerosol emissions found in GC3.0 (Williams et al. 2018). Appendix D of the UKCP report (Murphy et al. 2018) gives full details of the changes included. The configuration consists of the following model components: atmosphere: GA7.0 (Walters et al. 2017), land: GL7.0 (Walters et al. 2017), ocean: GO6.0 (Storkey et al. 2018) and sea ice: GSI8.0 (Ridley et al. 2017). The atmosphere and land components of GCM-PPE are configured on a regular latitude–longitude grid at N216 resolution, which gives a horizontal grid spacing of approximately 60 km at mid-latitudes. There are 85 vertical levels, 30 of which are in or above the stratosphere. Surface heat and water flux adjustments have been applied to each member of GCM-PPE to ensure realistic sea surface temperature and sea ice patterns (Murphy et al. 2018).

The RCM consists of a limited area version of the atmosphere and land surface configuration used in GC3.05. It uses the EURO-CORDEX latitude–longitude grid with 0.11° resolution and a rotated pole set at 39.25° N, 198° E. This gives a quasi-uniform grid spacing of 12 km over the European domain. The RCM has 63 vertical levels with the first 50 levels being the same as those used in the global model, but with a lower model top (~40 km). It is driven in a one-way nesting approach using daily sea surface temperature (SST) and sea ice extent, and 3 hourly time series of atmospheric prognostic variables at its lateral boundaries provided by its corresponding GCM-PPE member. In its internal domain, the RCM simulation evolves freely. Some inland water bodies are resolved on the RCM grid, but not in the GCM. For these grid boxes a decision was taken on what will provide more credible characteristics: either interpolation of SST and sea-ice fields from the coarser global model fields, or editing the land sea mask to represent the lake as land points and modelling them as inland water in the JULES land surface scheme. High elevation Swedish lakes were set as land points on the assumption that most of them would normally be frozen in winter, whereas nearby GCM-PPE sea points may not be. However, the Finnish lakes are included as sea points, as well as the large Russian lakes (Onega and Ladoga) and the Bosphorus. The latter are important for simulation of surface heat and moisture fluxes in Eastern Europe, especially during summer.

Other than the difference in model resolution and time-step, the representations of atmospheric dynamics and the parameterisations of land and atmospheric processes are largely identical in the regional and global simulations, with parameter perturbations applied in each driving GC3.05-PPE member mirrored in its RCM counterpart. The exceptions are the treatment of aerosols (see next paragraph) and two schemes that are only available in global models namely a stochastic physics package (Sanchez et al. 2016), and the TRIP (Total Runoff Integrating Pathways model: Oki and Sud 1998) river routing scheme.

GCM-PPE uses GLOMAP-mode (Global Model of Aerosol Processes) aerosol scheme described in Mann et al. (2010) to provide a physically based treatment of aerosol microphysics and chemistry. This scheme is computationally expensive: inclusion increases model runtime by 50% in atmosphere and land only simulations (Walters et al. 2017). This meant that we were unable to include GLOMAP-mode in RCM-PPE. It was therefore decided that the RCM would instead approximately replicate aerosol radiation and cloud effects simulated by the driving GCM-PPE member using the simpler EasyAerosol Scheme (Stevens et al. 2017). In the RCM, aerosol optical properties (absorption, extinction and asymmetry) for the 6 shortwave and 9 longwave wavebands of the SOCRATIES radiative transfer code (Edwards and Slingo 1996; Manners et al. 2015), and also cloud droplet number concentrations (CDNC) were prescribed from the driving GCM using time-varying monthly mean, full 3D spatial fields. In a global model, Stevens et al. (2017) showed that this approach replicates quite well the aerosol forcing found in interactive simulations, and testing in GC3.05 supported this conclusion. Further assessment (Bellouin and Thornhill 2018) using GA7.1 at N96 resolution has since found that EasyAerosol successfully replicates clear sky aerosol forcing, but the prescription of monthly-mean CDNC led to radiative imbalance in all sky conditions as non-linearities in aerosol cloud interactions meant that clouds were on average optically thicker compared to simulations with interactively varying CDNC. An error affecting the seasonality of aerosol-radiation forcing in the RCM-PPE (the DLH-error) has since been found. This led to shortwave extinction and absorption properties being over prescribed in summer by ~20% and under prescribed by ~40% in winter. A full description and assessment on the impact of this error is provided in supplementary material Sect. 2, and also briefly in Sect. 3.3.

For the ERAI-RCM-STD simulation, it was not possible to provide EasyAerosol forcing fields from the driving GCM. In this simulation, changes in monthly anthropogenic aerosol optical properties from MACv2-SP (Stevens et al. 2017) have been added to a climatology from a pre-industrial simulation that used the GLOMAP-mode aerosol. These easy aerosol inputs also include an estimate of volcanic forcing that replaces the (Sato et al. 1993, updated) estimates used in RCM-PPE. A comparison of the two volcanic forcings is provided in supplementary Fig. 2.2.1

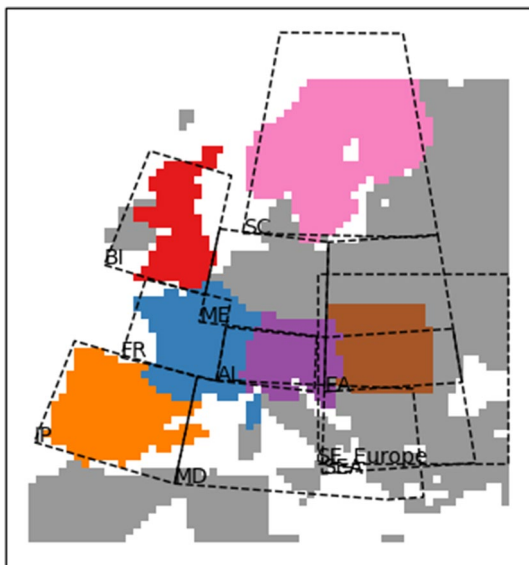
## 2.3 Analysis methods

### 2.3.1 Observed data sets used

To ensure consistency, grid box model results should be compared to area-averaged observational estimates. However the accuracy of grid box average estimates depends

on the underlying station density, interpolation methods, and the spatial homogeneity of the grid box (Herrera et al. 2019). Here we use the E-OBS (v20) dataset of European daily 0.22° resolution gridded data set of surface temperature and precipitation (Cornes et al. 2018). The underlying station density (and thus the accuracy) varies across the domain, with the highest station density in central Europe and Scandinavia, and fewer stations in the South and East of the domain.

The decision to use E-OBS is due to its domain wide coverage, however datasets produced by National Met Services (NMSs) for countries or sub regions of Europe are generally developed using many more station data than are available to E-OBS, and can use an interpolation procedure best suited to the particular region and station density (Cornes et al. 2018). Consequently, NMS data sets would be expected to provide a more accurate estimate than E-OBS. For precipitation the differences are largest in the extremes (Cornes et al. 2018), and hence where available we have used regional data in the evaluation of daily rainfall distributions (Sect. 3.3.1). Figure 1 shows where alternative sources have been used, and references for the sub region datasets are provided in Table 1. It should be noted that in addition to interpolation error, measurement errors in precipitation may be substantial: undercatch can be as much as 20% for rain (Sevron and Harmon 1984), and 80% for snow (Goodison et al. 1997). Consequently, the NMS datasets are still likely to underestimate precipitation.



**Fig. 1** Regions where different source datasets have been used in the analysis of extreme precipitation. Table 1 provides names and references of all sources. The grid is that of RCM-LOWRES which is being used for the analysis. Dotted lines define the subregions that have been used for analysis in this paper

**Table 1** Sources of observed datasets used in Sect. 3.3 for the analysis of extreme precipitation

Dataset name	References	Colour
E-OBS (v20)	Cornes et al. (2018)	Grey
SPAIN02 (v2)	Herrera et al. (2012)	Orange
SAFRAN reanalysis (v2)	Vidal et al. (2010)	Blue
EURO4M-APGD (v1.2)	Isotta et al. (2014)	Purple
CARPATCLIM	Spinoni et al. (2015)	Brown
NGCD_RR (v19.03)	Tveito et al. (2005)	Pink
UKCP09	Perry et al. (2005)	Red

The colour in column 3 refers to the colour this dataset is given in Fig. 1

For cloud and radiation, mean cloud cover and surface radiation estimates from CERES EBAF edition 4 (Clouds and Earth's Radiant Energy System, Energy balanced and Filled, Loeb et al. 2018; Kato et al. 2018) are used. This is a gridded global dataset with 1° resolution covering the years 2000–2020. The sensitivity of our results to choice of observed estimates has been tested by using the CLARA-A2: CM SAF cLoud, Albedo and surface Radiation dataset from AVHRR data—Edition 2 (Karlsson et al. 2017), and our results have found to be generally insensitive.

We use mean sea level pressure from the ERA-Interim reanalysis.

### 2.3.2 Time periods and common grid

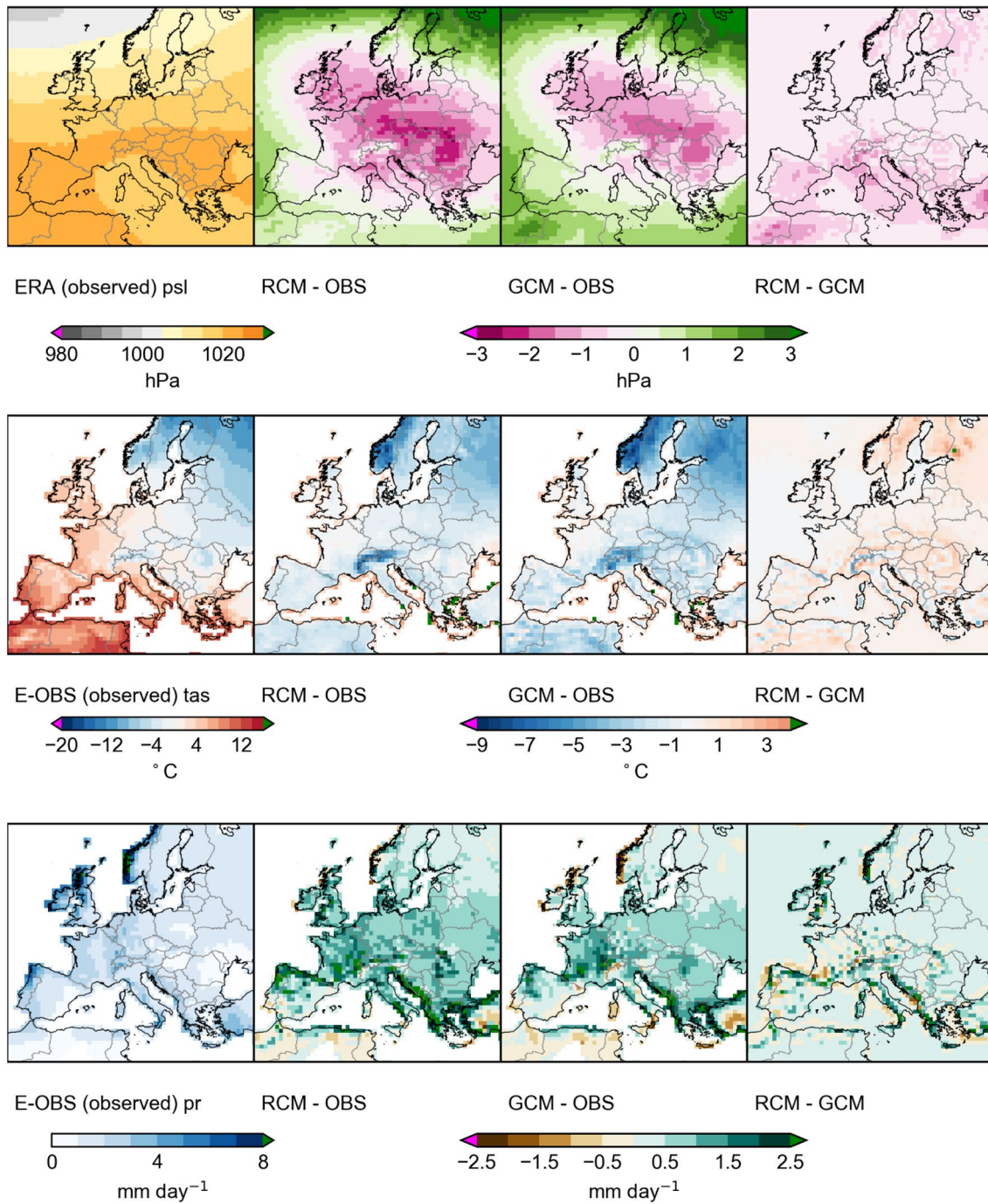
We evaluate the 20 year time period 1982–2002 and analyse winter (DJF) and summer (JJA) seasons. This time period has been chosen to be as close as possible to the 1980–2000 baseline used in the UKCP report, whilst also being fully spanned by the ERAI-RCM-STD simulation. Note data availability means that it was necessary for the CERES cloud and radiation climatologies to cover a different 20 year time period (2000–2020).

All data has been conservatively regridded on to the domain of RCM-LOWRES. This is at a resolution similar to the GCM whilst also providing a quasi-uniform grid spacing.

## 3 Results

### 3.1 Evaluation of precipitation and temperature climatologies

Figure 2 shows DJF ensemble mean sea level pressure (PMSL), surface air temperature and precipitation biases. Plots indicate very good large-scale circulation consistency between the driving global and regional models.



**Fig. 2** Left DJF seasonal mean observed, left-centre t RCM-PPE biases, right-centre GCM-PPE biases, right differences between RCM and GCM. Top row: sea level pressure (ERA-INTERIM reanalysis is

used as observations), middle row: surface air temperature, bottom row: precipitation

RCM-PPE and GCM-PPE ensemble mean biases have similar spatial patterns and magnitudes. In particular both ensembles have a cold bias that reaches 8 K over Scandinavia. The magnitude of the bias over Scandinavia is however slightly reduced in RCM-PPE and is largely removed in ERAI-RCM-STD except over Norway (supplementary figure

S2). It is also worth noting that for mountainous regions such as the Alps and Norway, Cornes et al. (2018) report a discrepancy between the E-OBS values and those from NMSs, with E-OBS warmer on average by more than 5 °C across the high Alps. This discrepancy is likely the result of there being more high-elevation stations available for the

MeteoSwiss interpolation. The ERAI-RCM-STD results suggest that the Scandinavian cold bias is largely inherited from the driving GCM. Reasons for the cold bias in GCM-PPE are discussed in the UKCP report (Murphy et al. 2018), and contributing factors include a strong aerosol forcing. Additionally, McSweeney et al. (2021) find that GCM-PPE members with the largest Scandinavian cold biases also have the weakest circulation over the Atlantic. Consistent with this, McDonald et al. (prep) find that the members with fewer storms over the UK are drier and colder over Europe. Thus, variations in mean circulation seem to play a role in explaining variations in the cold bias across members. Both the GCM-PPE and RCM-PPE ensembles have a wet bias over almost the entire domain, with the bias slightly ( $\sim 0.3$  mm/day) worse in RCM-PPE.

In addition to ensemble mean maps, in supplementary figure (S4) we show boxplots for the different PRUDENCE analysis regions to see the spread among members. These regions are intended to provide homogeneous climatic conditions and have been used to allow easy comparison with results from EURO-CORDEX studies (such as Kotlarski et al. 2014). For reference, a EURO-CORDEX box is also included for the 32  $0.11^\circ$  RCM EURO-CORDEX simulations we were able to plot using ESMValTool (Righi 2020) (a full list of included simulations is included in the figure caption). For Scandinavia, 11 out of 12 UKCP18 RCM-PPE members have a cold bias with strong correlations between RCM and driving GCM-PPE members. The plot shows that the cold bias is not systematic in the EURO-CORDEX ensemble, and furthermore simulations using the MOHC-HadREM3-GA7-05 RCM (same as our RCM-STD but driven by different GCMs) span a significant part of the spread in EURO-CORDEX temperatures. This provides further evidence that the cold bias is inherited from the driving UKCP18 GCM-PPE. It is also of note that the same member is the warmest in winter for all regions in both RCM-PPE and GCM-PPE.

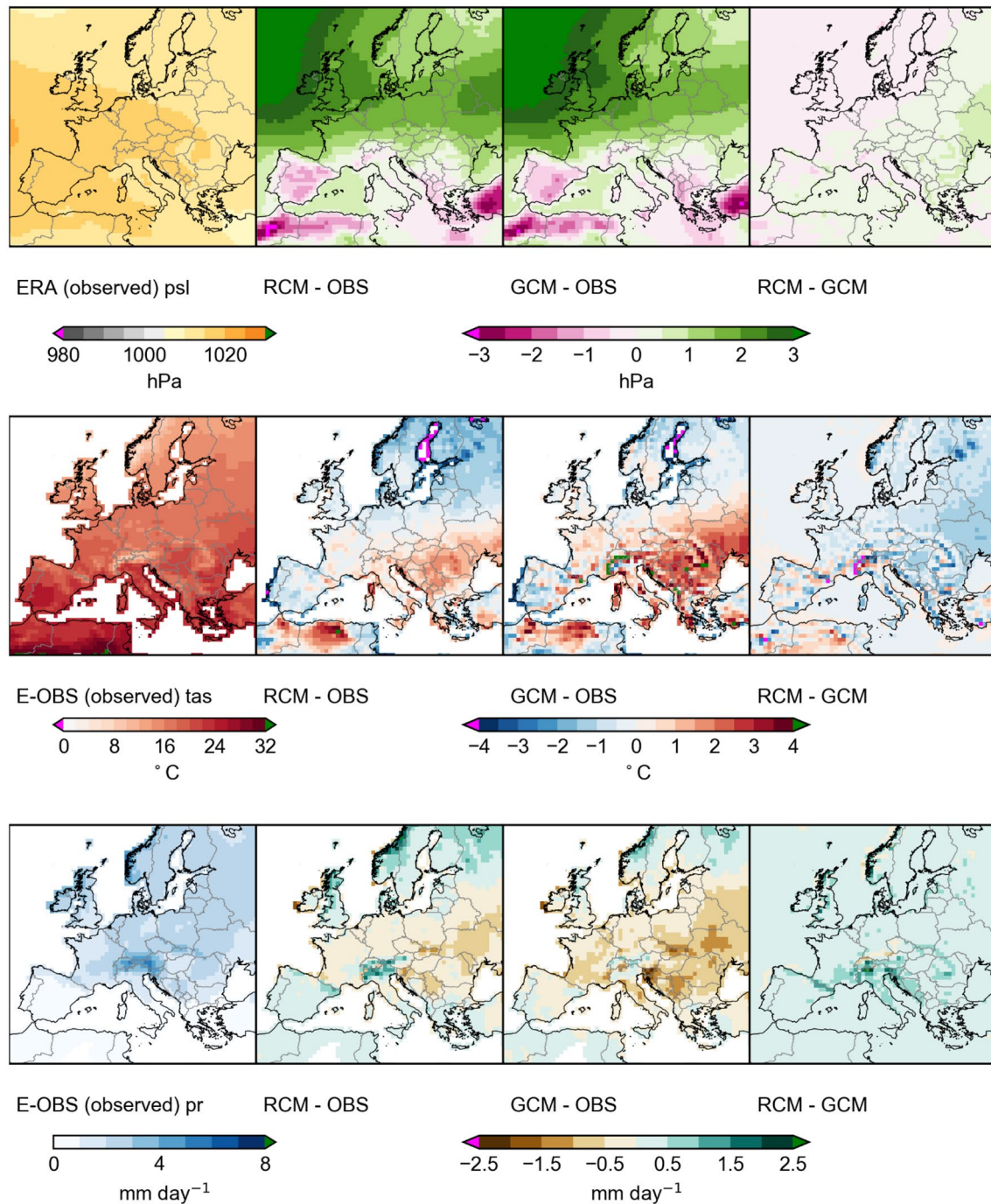
Scatter plots of land only surface air temperature and precipitation (supplementary figure S7 top) show that warmer members tend to be wetter. Both land only surface air temperature (supplementary figure S7 middle) and precipitation (supplementary figure S7 bottom) correlate well with domain averaged 850 hPa specific humidity. We observe that warmer members tend to be wetter because specific humidity is higher and therefore more water is available for precipitation. We also note that there is a high level of consistency between RCM-PPE and GCM-PPE in terms of correlation between RCM members and their driving GCM members (i.e. the warmest and wettest GCM members being also the warmest and wettest RCM members).

Figure 3 shows similar results for summer, where both ensembles have a cold bias over Scandinavia, but a warm and dry bias over South East Europe. In terms of differences

between the two ensembles the RCM-PPE is around 0.5 K cooler than the GCM-PPE, and like winter the RCM-PPE is  $\sim 0.3$  mm/day wetter. The temperature difference means that the South East European warm bias is smaller in RCM-PPE, whilst the cold bias in the North East of the domain is larger.

Warm dry summer biases in South East Europe are also found in EURO-CORDEX evaluation (reanalysis driven) simulations (Kotlarski et al. 2014), and are believed to be related to local soil moisture feedbacks in a soil moisture-controlled evaporative regime. Specifically, where evaporation from the surface is limited by available soil moisture, the amount of energy used for latent heat flux decreases and sensible heat flux increases, leading to an increase of air temperature (e.g., Seneviratne et al. 2010). Supporting this, Knist et al. (2017) investigate land atmosphere coupling strength in the same set of simulations and find it tends to be too strong in this region. We note as an aside from boxplots of the Mediterranean (MD) and East Europe (EA) region (Supplementary figure S7) that warm and dry and wet biases are not present consistently in the GCM driven EURO-CORDEX simulations, suggesting that errors other than too strong land atmosphere coupling may be compensating. Support for a soil moisture feedback explanation in our ensembles is provided by supplementary figure S8, the top panel shows a scatter plot of South East Europe surface air temperature and evaporation, that shows negative correlations between these variables in both RCM-PPE and GCM-PPE. Further the middle panel show that evaporation is correlated with 850 hPa specific humidity, so reduced evaporation reduces the amount of moisture available for precipitation (and thus the soil stays dry). Supplementary figure S8 bottom panel shows that all GCM-PPE members are warmer and drier than observations, nine out of twelve RCM-PPE members are warmer than the observed, and nine out of twelve RCM-PPE members are wetter than the observed. Additionally we note from the top panel of Fig. 3 that PMSL biases are small (less than 3 hPa) suggesting that large scale circulation differences are unlikely to be the cause of the large temperature biases, although we do note that in both ensembles there appears to be a northward shift in the Azores high that may cause more blocking conditions over central Europe.

The large-scale circulation consistency between the driving global and regional models is further assessed at the daily time scale. For each day the spatial correlation of sea level pressure between the RCM and GCM has been calculated, supplementary table S1 shows the 25th, 50th and 75th percentile of the correlations for each member. In winter the median correlation is above 0.97 for all members, and in summer above 0.90. The correlations found here are generally higher than those found in EURO-CORDEX simulations (Vautard et al. 2020), which may be due to each RCM-PPE



**Fig. 3** As Fig. 2, but for JJA

and GCM-PPE member having the same atmosphere and land surface configuration. These results provide further evidence that precipitation and temperature differences are not due to large scale circulation differences.

Supplementary figure S3 shows that ERAI-RCM-STD is warmer than RCM-STD over land despite having cooler SSTs. This is an interesting result that requires further investigation, but which must be due to either different lateral

boundary conditions and/or aerosol differences. ERAI-RCM-STD temperature is a clear improvement over Scandinavia, but over SE Europe the warm bias is enhanced.

In summary, there is generally good large scale consistency between RCM-PPE and GCM-PPE in both seasons, both in terms of correlations between members (domain average correlations above 0.9 for all variables and seasons) and spatial patterns of the ensemble mean. In winter



this means that RCM-PPE inherits a large cold bias over Scandinavia. Despite the generally good large scale consistency, RCM-PPE JJA ensemble mean is cooler by 0.5 K on average, with all RCM members being cooler than their driving GCM. Additionally, in both seasons RCM-PPE is wetter than GCM-PPE in both seasons by  $\sim 0.3$  mm/day. We investigate these differences in the next section.

### 3.2 Causes of differences in temperature and precipitation

Surface temperature is ultimately determined by the net effect of many drivers. Such drivers can include remote influences on atmospheric heat and moisture convergence into the region, plus local effects such as clouds, precipitation, aerosols, boundary layer mixing, snow cover and soil properties including soil moisture content. These drivers can both respond to and modify the state variables, and impact surface air temperature via the surface energy budget (radiative fluxes, conductance into the deep soil, sensible heat flux and latent heat flux).

To illustrate some of the terms in the surface energy budget, Fig. 4 shows JJA downward surface radiative flux (shortwave plus longwave), surface upward radiative flux, latent heat flux and sensible heat flux. In GCM-PPE there is too much downward (i.e. incoming) surface radiation over the UK and central Europe. In terms of differences between RCM-PPE and GCM-PPE, RCM-PPE has less downward (and net) radiation, which may be due to differences in cloud cover. The latent heat flux is higher over parts of the Mediterranean Sea, and South East Europe. For the latter region this is consistent with the idea that soil moisture feedbacks are enhancing the temperature differences between the two ensembles. Both upward radiation and sensible heat flux is lower than in GCM-PPE, consistent with the temperature being lower in the RCM-PPE.

Full analysis of the temperature differences between ensembles would involve analysing all the various drivers and their interactions. The focus of this subsection however, is the analysis of clouds and their radiative effects in order to isolate their contribution to the summertime temperature differences between RCM-PPE and GCM-PPE. This is partly motivated by the differences in downward (shortwave plus longwave) radiation above, and also the fact that cloud cover is known to be influenced by resolution. We find that cloud cover is considerably higher in RCM-PPE than in GCM-PPE. In general, increased cloud cover causes a decrease in downward surface shortwave radiation and an increase in downward longwave radiation. However, the cloud radiative effect is also dependent on a number of factors including the height of the cloud, and the cloud thickness. In Sect. 3.2.1, we analyse how cloud cover and surface radiation diagnostics compare to observations. In Sect. 3.2.2, we use the extra

simulations to try and unpick what configuration differences cause the differences between the RCM and GCM.

#### 3.2.1 Clouds and radiation

Figure 5 shows DJF cloud cover, and shortwave and longwave cloud radiative forcing at the top of the atmosphere. Both ensembles have more cloud cover than observed (domain average bias 13% in RCM-PPE and 10% in GCM-PPE), with the bias being greater over land than sea. Despite the positive cloud cover bias, both ensembles have a positive shortwave cloud forcing bias, and negative longwave cloud forcing bias. This suggests that other errors in the representation of clouds (possibly clouds being too thin) are offsetting the cloud cover biases.

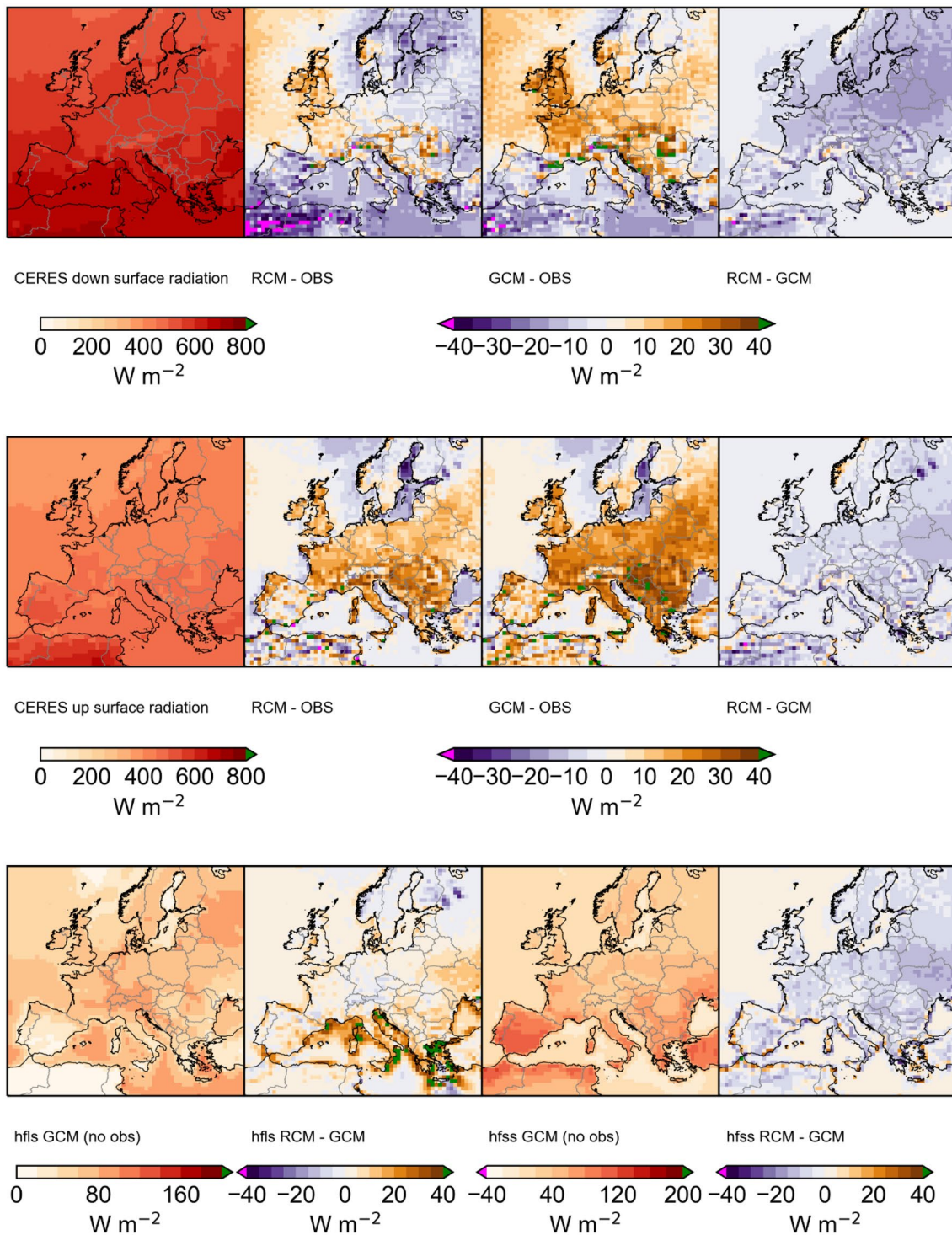
Figure 6 is for JJA. Cloud cover is increased in RCM-PPE compared to GCM-PPE (domain average absolute difference of 6.5%). As both ensembles under simulate cloud cover for the majority of the domain, this represents an improvement in RCM-PPE. Unlike in DJF, radiative forcing biases seem to simply follow the cloud cover differences, this means that cloud forcing biases are reduced in RCM-PPE over the parts of the domain where cloud cover is under-estimated.

Supplementary figures S8 (DJF) and S9 (JJA) show that cloud cover bias is reduced in ERAI-RCM-STD compared to RCM-STD, suggesting that the bias is partially inherited from GCM-PPE. The magnitudes of the differences are generally small, but it is of note that in the North East of the domain there is a large ( $\sim 20 \text{ W m}^{-2}$ ) difference in JJA shortwave cloud forcing, which are likely contributing to the warmer surface temperatures in ERAI-RCM-STD.

We also note that for DJF, biases in net downward surface radiation are negative for shortwave and positive for longwave (supplementary figure S10). This is the opposite sign to the cloud radiative forcing biases, meaning that the largest radiative forcing biases are coming from clear-sky rather than cloud processes. In contrast, JJA net surface radiation biases (supplementary figure S11) are largely similar to cloud radiative forcing biases.

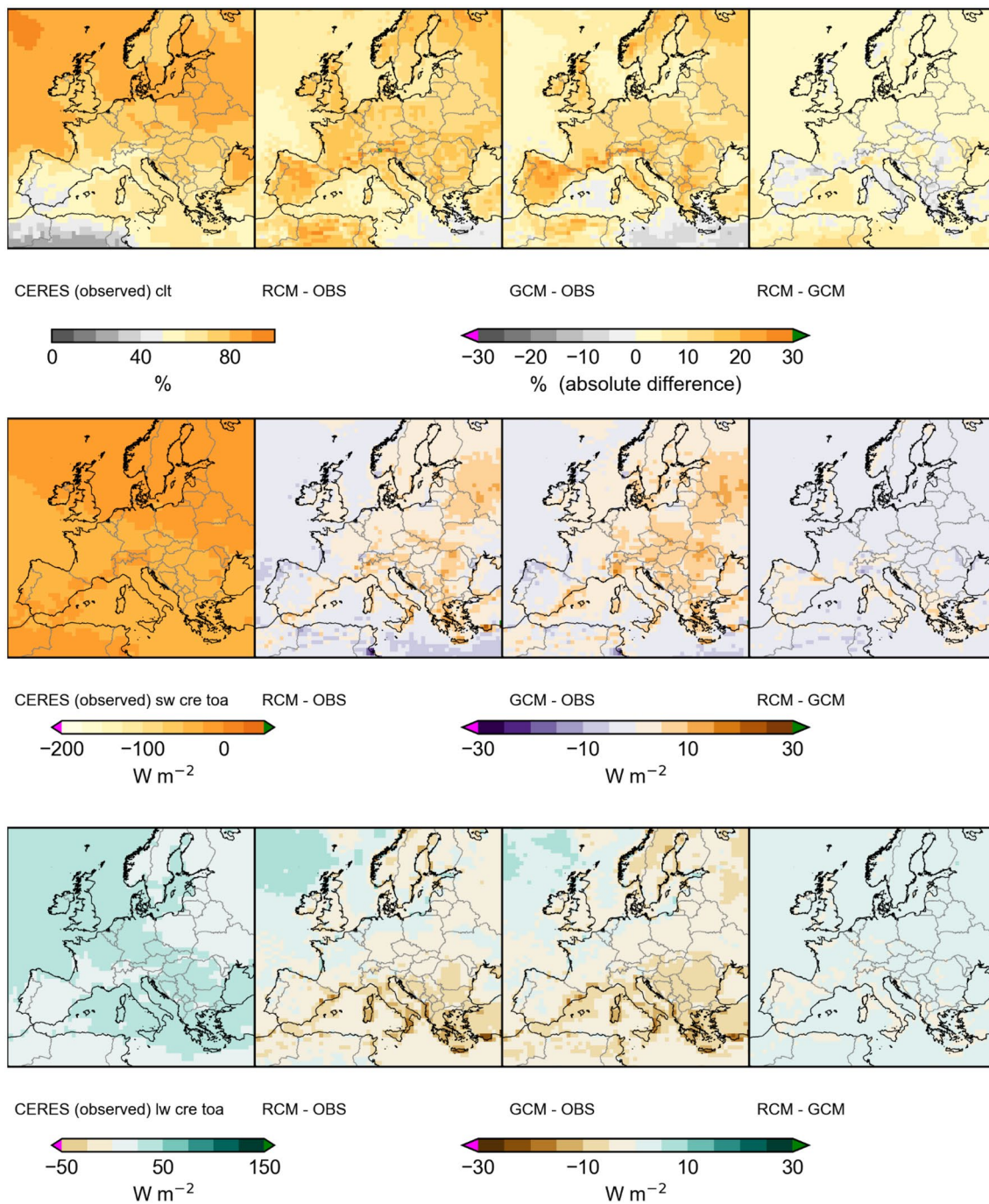
#### 3.2.2 Sensitivity simulations

In terms of what is causing the cloud and radiation differences between the RCM and GCM, model development tests show a cloud sensitivity to resolution in GA7, with higher resolutions having an increased summertime cloud cover of  $\sim 5\%$  around the UK. However, the increased cloud cover goes with a reduced optical depth and so the radiative variation with resolution is smaller. We therefore might expect to see cloud differences with resolution. In terms of other configuration differences, we speculate that the most notable is the treatment of aerosols. In Bellouin and Thornhill (2018), it is found that the simulation



**Fig. 4** JJA. Top row surface downward radiation (longwave + shortwave). Top row left is the observed estimate from CERES, top row second from left is the bias in RCM-PPE, third from left the bias in GCM-PPE and on the right the difference between RCM-PPE and GCM-PPE. Second row is the same as the top row, but for surface

upward radiation. Bottom row left is the latent heat flux in GCM-PPE, bottom row second panel is the difference in latent heat flux between RCM-PPE and GCM-PPE. Bottom row third from left is the sensible heat flux in GCM-PPE, and bottom row right is the difference in sensible heat flux between RCM-PPE and GCM-PPE.



**Fig. 5** First left DJF seasonal mean observed, third from right RCM-PPE biases, second from right GCM-PPE biases, right differences between RCM and GCM. Top row: cloud cover, middle row: shortwave cloud radiative effect, bottom row: downward longwave surface radiation

using EasyAerosol has increased liquid cloud cover, and consequently reduced downward shortwave, and increased downward longwave radiation. In addition, there is also the DLH-error in the prescribing of EasyAerosol shortwave extinction and absorption properties. Any differences due to the use of EasyAerosol, and in particular the day light hours error is undesirable as the scheme is intended to

replicate in the RCM the effects of explicit aerosol modeling included in the driving GCM. In the rest of this subsection we use the RCM-LOWRES and RCM-LOWRES-DLHFIX simulations to determine which differences are due to resolution, which are due to the error, and which are due to other differences. Specifically, we look at:

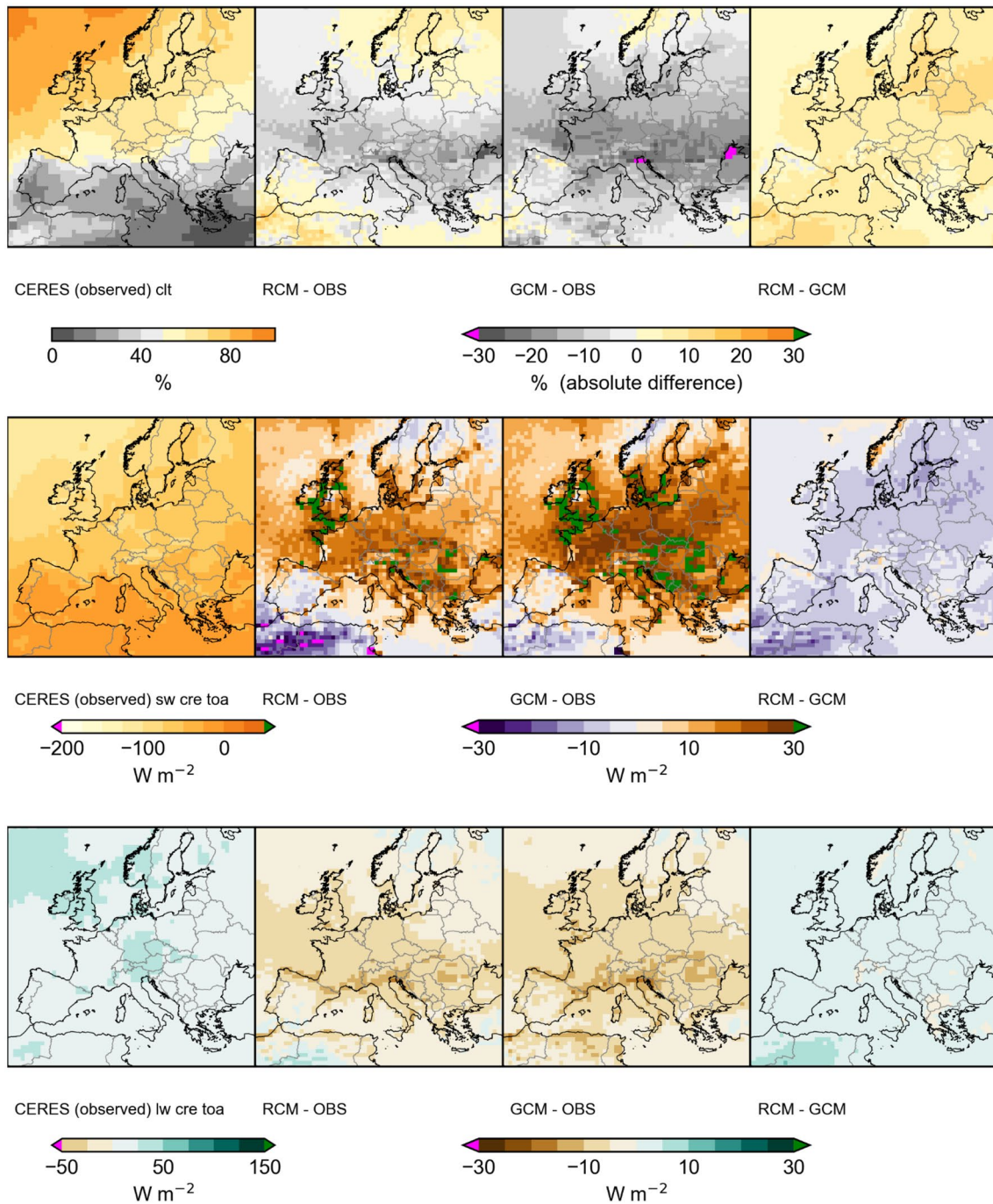


Fig. 6 As Fig. 5, but for JJA

- RCM-STD – RCM-LOWRES to see what differences are due to resolution (RES-CONT)
- RCM-LOWRES – RCM-LOWRES-DLHFIX to see what differences are due to the daylight hours error (DLH-ERR-CONT)
- RCM-LOWRES-DLHFIX – GCM-STD to see what differences are not due to either resolution or the DLH

- error, believed to be dominated by the different treatment of aerosols (OTHER-CONT)
- RCM-STD – GCM-STD showing the total difference between the two simulations (TOTAL-DIFF)

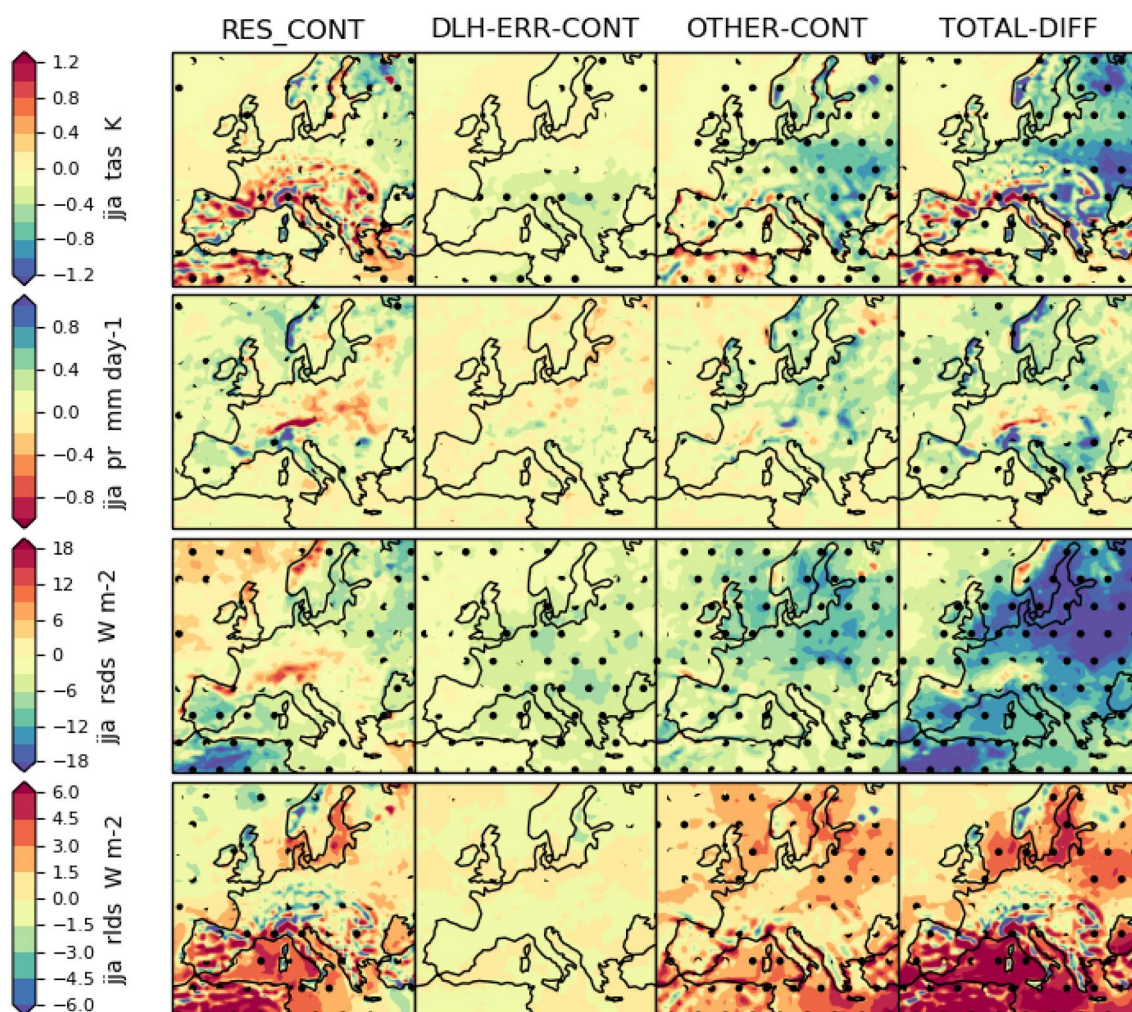
We speculate that the differences between RCM-LOWRES-DLHFIX and GCM-STD (OTHER-CONT) are

dominated by the treatment of aerosols, however other differences between the two runs are described in Sect. 2.2 and summarised as.

1. The RCM-LOWRES-DLHFIX is an atmosphere only model with prescribed SST and sea ice from GCM-STD, whilst GCM-STD includes coupled ocean and sea ice models.
2. The RCM-LOWRES-DLHFIX is a limited area model covering the EURO-CORDEX domain, with information passed from GCM-STD at the lateral boundaries. Additionally, the land sea mask has been produced from a different source and so will differ from the GCM mask. Most notably lakes Onega and Ladoga stand out in the plots.

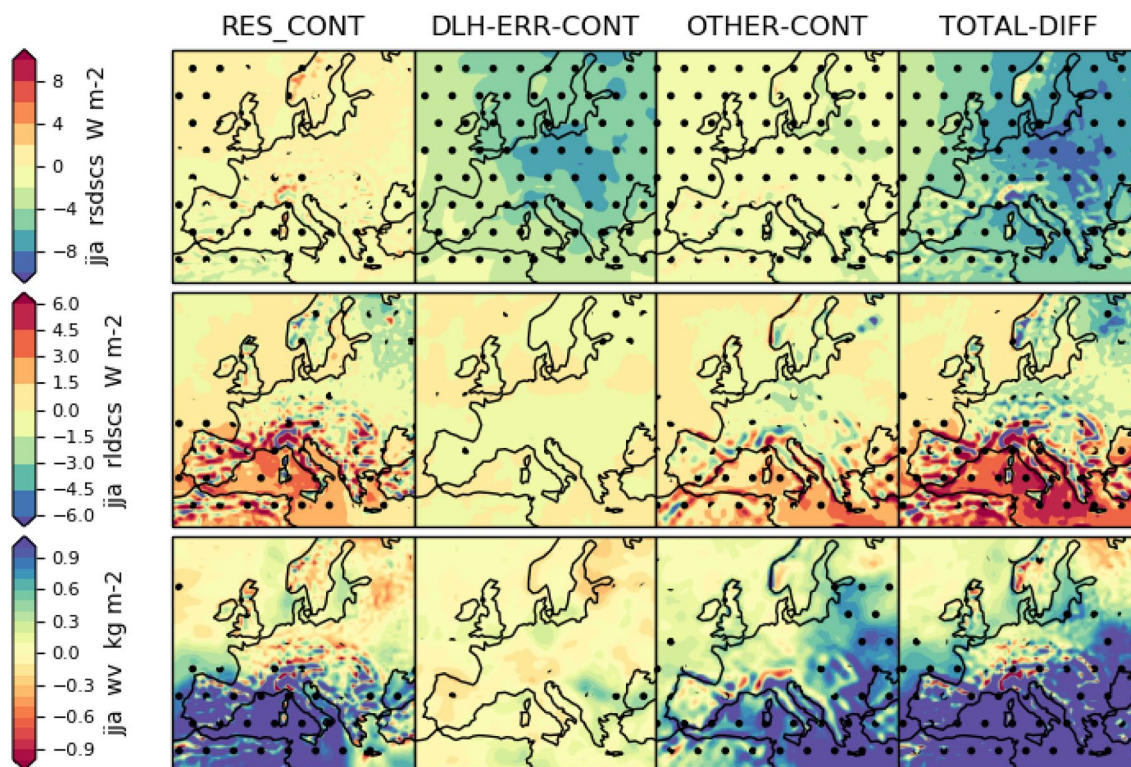
3. The RCM has a lower (~40 km) upper lid.
4. A couple of schemes are only available in global models, namely a stochastic physics package (Sanchez et al. 2016), and the TRIP (Total Runoff Integrating Pathways model: Oki and Sud 1998) river routing scheme.

We show figures for JJA with corresponding figures for DJF in supplementary material (figures S12, S13 and S14). Figures 7, 8 and 9 the left hand column shows RES-CONT, second from left column shows DLH-ERR-CONT, third from left column shows OTHER-CONT and the right hand side shows TOTAL-DIFF. Due to only having 10 years of data for RCM-LOWRES, a paired t-test has been performed to determine whether differences are



**Fig. 7** JJA. Left column shows differences between RCM-STD and RCM-LOWRES (RES-CONT), second from left column differences between RCM-LOWRES and RCM-LOWRES-DLHFIX (DLH-ERR-CONT), third from left differences between GCM-STD and RCM-LOWRES-DLHFIX (OTHER-CONT), and right column shows differences between RCM-STD and GCM-STD (TOTAL-DIFF). Top

row: surface air temperature, second row: precipitation, third row: downward shortwave surface radiation, bottom: downward longwave surface radiation. All plots are seasonal means for the 10 year period that RCM-LOWRES was run for. Dots indicate areas where differences are statistically significant at the 5% level. Note that the longwave and shortwave plots have different scales



**Fig. 8** JJA clear sky differences. Left column shows differences between RCM-STD and RCM-LOWRES (RES-CONT), second from left column differences between RCM-LOWRES and RCM-LOWRES-DLHFIX (DLH-ERR-CONT), third from left differences between GCM-STD and RCM-LOWRES-DLHFIX (OTHER-CONT), and right column shows differences between RCM-STD and GCM-STD (TOTAL-DIFF). Top row: clear sky downward shortwave

surface radiation, second row clear sky downward longwave surface radiation, bottom: total atmospheric water vapour. All plots are seasonal means for the ten year period that RCM-LOWRES was run for. Dots indicate areas where differences are statistically significant at the 5% level. Note that the longwave and shortwave plots have different scales

significant: the areas where the differences are statistically significant at the 5% level are marked with dots.

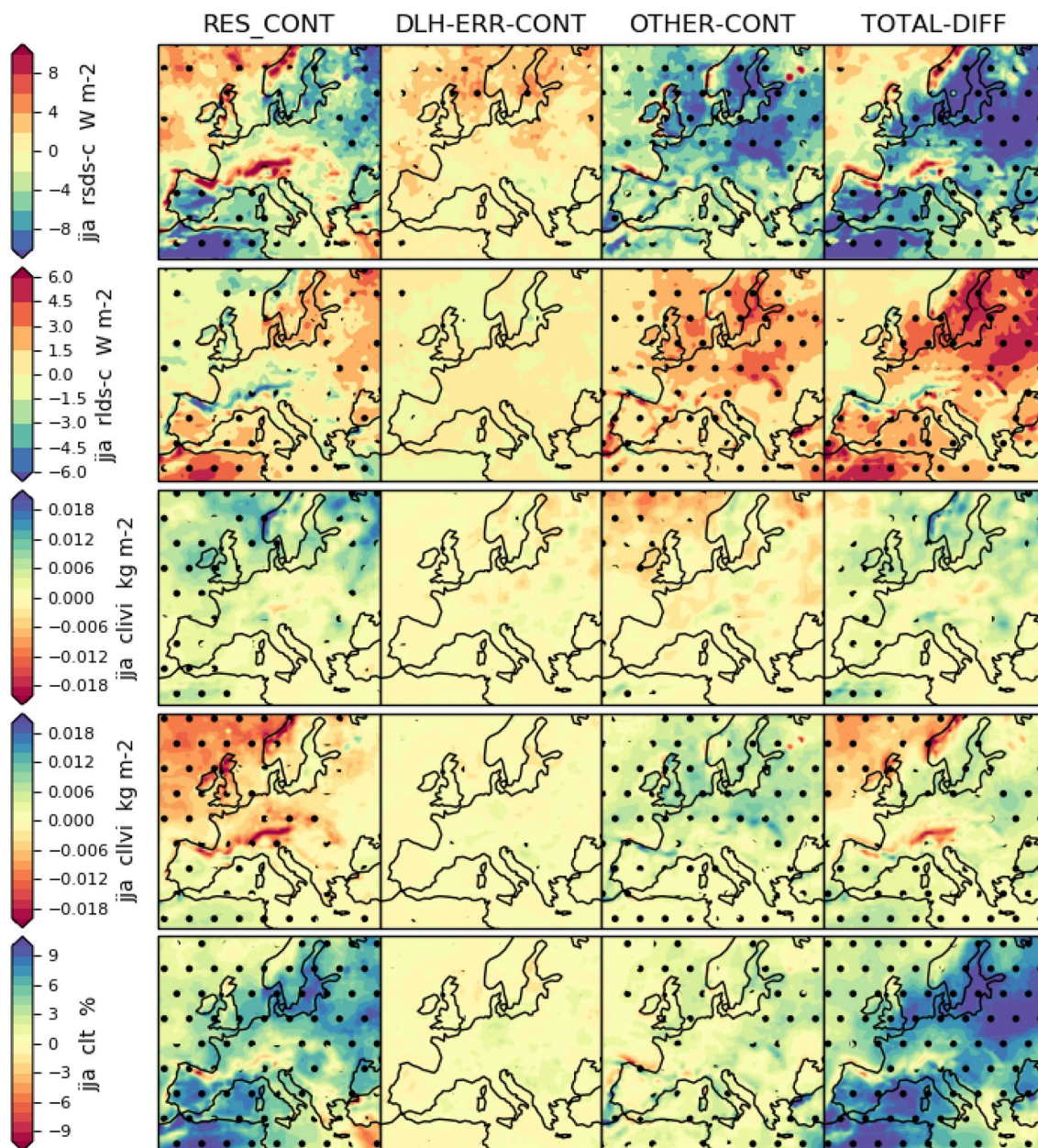
The top row of Fig. 7 shows temperature differences, where RES-CONT shows a clear orography footprint with the Pyrenees, Alps, Carpathian, Scandinavian and Scottish mountain ranges clearly identifiable. Correspondingly high elevation areas have increased downward shortwave (third row), and decreased downward longwave (bottom row). In addition to orographic features, higher resolution is causing a cooling over Finland. As expected, the day light hours error is causing a reduction in downward shortwave radiation and no impact on longwave, as a consequence of this DLH-ERR-CONT is cooler over large areas of central and Eastern Europe by up to 0.4 K. It should be noted that the day light hours error is typically contributing less than a third of the total surface air temperature difference, and OTHER-CONT is the largest contributor to the total summertime cold difference.

Precipitation differences are on the second row. Based on previous experience (e.g. Jones et al. 1995, 1997; Prein et al. 2016) we expect to see more precipitation at higher

resolution due to a stronger hydrological cycle as well as local effects in the vicinity of mountains/coast due to the better representation of topography at high resolution. Our results bear this out, with resolution causing an increase in precipitation over the Atlantic, and local effects in the vicinity of mountains. Reassuringly the DLH-error is having little impact on precipitation differences. Other differences are also contributing to increased precipitation over large areas of central and eastern Europe.

TOTAL-DIFF radiation differences are of opposite signs for longwave and shortwave, however the negative shortwave differences are of a larger magnitude than longwave differences leading to a net negative downward radiation difference consistent with RCM-STD being cooler. The same reverse pattern between long and short wave differences can be seen in RES-CONT and OTHER-CONT, with both contributing to the total difference.

In the next two figures we further split the radiation differences into cloud and clear sky differences. Bellouin and Thornhill (2018) attribute radiative imbalances resulting from the use of EasyAerosol as due to the interactions of



**Fig. 9** JJA cloud differences. Left column shows differences between RCM-STD and RCM-LOWRES (RES-CONT), second from left column differences between RCM-LOWRES and RCM-LOWRES-DLHFIX (DLH-ERR-CONT), third from left differences between GCM-STD and RCM-LOWRES-DLHFIX (OTHER-CONT), and right column shows differences between RCM-STD and GCM-STD (TOTAL-DIFF). Top row: downward shortwave surface radiation due

to clouds ( $rsds-rsds_c$ ), second row: downward longwave surface radiation due to clouds ( $rlds-rlds_c$ ), third row: total cloud ice content, fourth row: total cloud water content, bottom: cloud area fraction. All plots are seasonal means for the 10 year period that RCM-LOWRES was run for. Dots indicate areas where differences are statistically significant at the 5% level. Note that the longwave and shortwave plots have different scales

aerosols and clouds, consequently this partitioning allows us to check whether our findings are consistent. Additionally, looking at cloud properties allows us to explore the cloud sensitivity to resolution observed in GA7 development.

Figure 8 top row shows differences in downward clear sky surface radiation. It can be seen that around 80% of the negative clear sky radiation differences are due to

DLH-ERR-CONT (ignoring locations where there is a land sea mask difference). Once the error is removed the remaining clear sky differences largely correspond to where there are differences in atmospheric water vapour (bottom row). RES-CONT is showing orographic features as well as a large positive difference in the South West of the domain. A possible explanation for the increased evaporation in RES-CONT

could be stronger surface winds, possibly due to the RCM generating more frequent or more intense storms in an area that includes two cyclogenesis regions (Trigo 1999). OTHER-CONT shows positive differences in the South East of the domain.

Figure 9 shows differences in downward surface cloud radiative effects (total downward short/long wave radiation – clear sky downward short/long wave radiation). Typically, where shortwave differences are negative/positive, longwave differences are positive/negative, with the magnitude of the longwave differences roughly half that of the shortwave.

RES-CONT has an increase in cloud cover over practically the entire domain (bottom row), with an absolute difference exceeding 5% over the North East of the domain (everything North East of a line drawn from Oslo to Istanbul), the West Mediterranean and North Africa. We might therefore expect a decrease in downward shortwave radiation, and an increase in longwave radiation, and indeed this is largely the case, for example over Spain there is a negative shortwave difference of over  $6 \text{ W m}^{-2}$ , and positive longwave difference of around  $3 \text{ W m}^{-2}$ . However over other parts of the domain it is more complicated. For example, over the Atlantic in the North West of the domain, despite a small increase in cloud cover, there is a positive shortwave difference (of  $\sim 4 \text{ W m}^{-2}$ ), and a negative longwave difference (of  $0\text{--}2 \text{ W m}^{-2}$ ). In this region we see a widespread reduction in liquid cloud amount, but an increase in cloud ice. Since liquid clouds tend to be lower and more reflective, this will act to decrease the downward longwave and increase the downward shortwave. The main exception to increased cloud cover, is mountainous regions with the Alps, Pyrenees, and highland areas of Scotland and Norway standing out. In these regions a reduction in cloud thickness is causing a reduction in the cloud radiative effect.

OTHER-CONT sees negative shortwave differences over the majority of the domain, with a strong spatial correlation between these and increases in liquid cloud amount. Easy-Aerosol does not directly interact with ice clouds, and for most of the domain we do not see any ice cloud differences, however in the Northern part over the Atlantic there is a small decrease. To further illustrate the changes in the vertical distributions of clouds, vertical profiles are shown in supplementary figure S15, OTHER-CONT shows an increased cloud amount at low levels (peak difference around 2 km), whilst RES\_CONT shows a reduction in low level clouds, but increase in high clouds.

Overall, in summer there is an increased cloud radiative effect in the RCM, with negative shortwave differences being larger than the positive longwave differences. The exception being over the North West of the domain over the Atlantic. The differences over the Atlantic largely cancel each other out, with RES-CONT and OTHER\_CONT

contributions being of opposite sign. Additionally, over the Atlantic, there are differences in cloud properties between the RCM and GCM, with increases in ice cloud amount and cloud area, but decreases in liquid cloud, which act to compensate for each other in terms of radiative effects.

Similar plots for winter are shown in supplementary material (figures S13, S14 and S15). We see that Easy-Aerosol introduces a warm difference in winter (around  $0.4\text{--}0.8 \text{ K}$ ) over France and Germany that is not detected as statistically significant in the total difference between the RCM and GCM, presumably due to increased variability in the RCM. Resolution generally causes a similar change in cloud properties in winter as it did in summer, although some parts of Southern Europe, such as Northern Spain, Italy, and the Balkans are an exception: in these regions resolution causes a reduction in cloud, and thus a reduction in downwards longwave and increase in downwards shortwave.

In conclusion, the summertime temperature difference between the RCM and GCM comes predominantly from other differences, believed to be EasyAerosol. Other differences also introduce a warm difference in winter over France and Germany. Our findings are consistent with those of in Bellouin and Thornhill (2018), who found that the simulation using EasyAerosol increased liquid cloud cover and thickness, and consequently reduced downward shortwave, and increased downward longwave radiation. These differences are further traced to the prescribing of monthly mean cloud droplet number concentrations. As the intention of EasyAerosol was to replicate the effects of the aerosol modelling in the GCM, these differences are undesirable.

It is of note that the RCM's summertime cold, and winter warm difference over Scandinavia is a genuine consequence of resolution. As is increased rainfall in both seasons over the Atlantic, and over Spain in JJA. These differences may be related to different cloud properties in the RCM, that include a change in vertical distribution, an increase in cloud cover, but reduced cloud thickness. These differences may represent an improvement, but to establish this and to gain further understanding of cloud sensitivity to resolution requires further investigation, which is beyond the scope of the current study.

### 3.3 Evaluation of daily distributions

In this subsection, we analyse daily precipitation (Sect. 3.3.1) and temperature (Sect. 3.3.2) distributions, two variables that are important from an impacts perspective. As before we restrict ourselves to analysing data on the low resolution grid, to allow a fair comparison between RCM-PPE and GCM-PPE. We note however that RCM-PPE may have additional skill at smaller spatial scales.

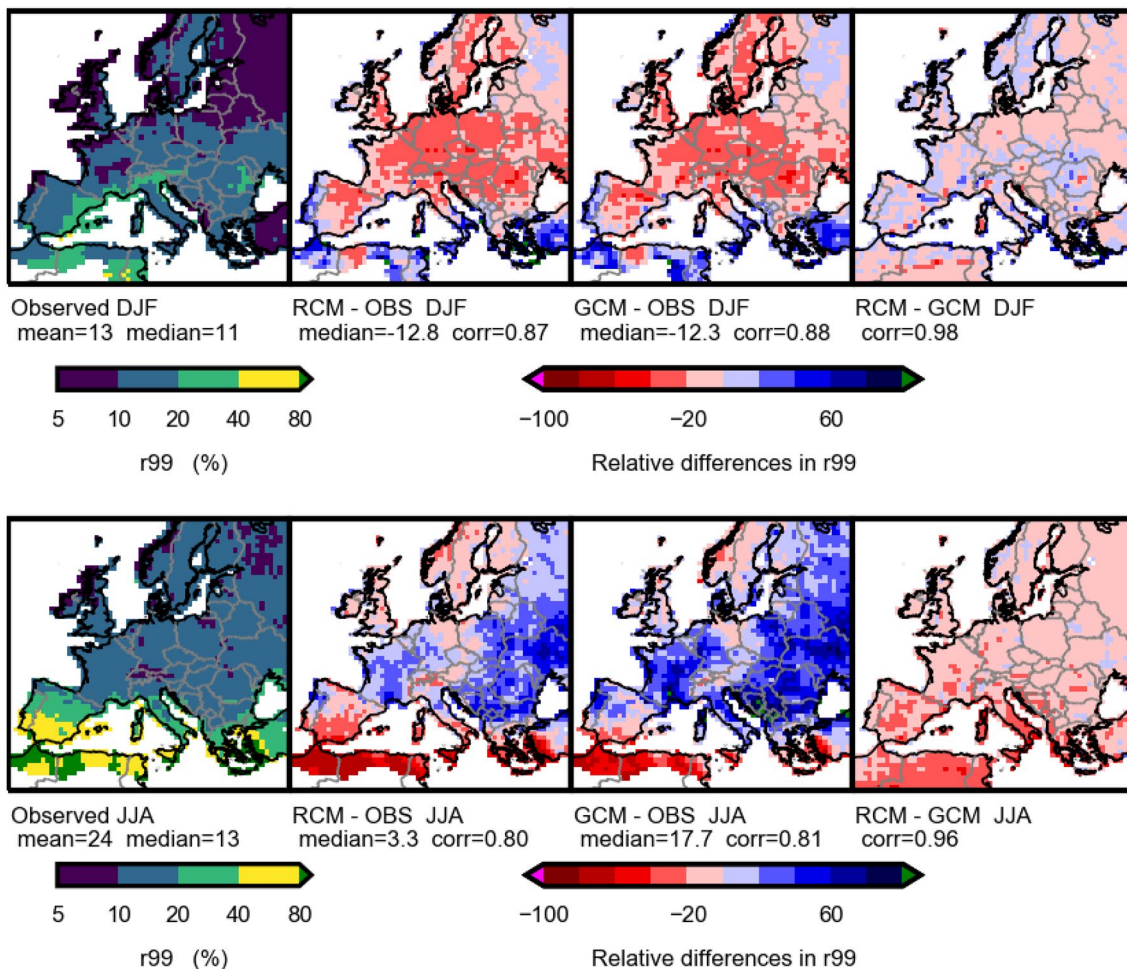


### 3.3.1 Precipitation

Figure 10 shows the percentage of seasonal mean precipitation that falls on days that have precipitation above the 99th percentile of daily precipitation for that season ( $r_{99}$ ). This statistic has been chosen as we already know that RCM-PPE is wetter on average, and instead here we want a measure of the shape of the rainfall distribution. Reassuringly there are typically no obvious changes in observed values along the boundaries where the source data has been switched. An exception may be the Russian border, that stands out in winter difference plots. This suggests that a lack of observations over Russia might be leading to an underestimate of the contribution of extremes to the total rainfall. The observed spatial distribution is very non uniform, particularly in summer. The median summer value is 13%, but there are Southern Mediterranean

regions in which it typically rains on fewer than 5% of days, and the majority of the total rainfall falls on days where the rainfall exceeds the 99th percentile. Due to the large spatial variation in this statistic, relative rather than absolute error has been used to compare the different datasets.

In winter both ensembles underestimate the proportion of rainfall falling as extremes, over the majority of the domain. The performance of RCM-PPE and GCM-PPE is remarkably similar. In summer both ensembles have too much precipitation falling as extremes over the majority of the domain, but the magnitude of the error is reduced in RCM-PPE (median absolute error reduced from 2.4 to 0.4, and relative error from 17.7 to 3.3%). RCM-PPE consistently has a lower contribution from extreme events than GCM-PPE over the whole domain, and this is largely an improvement except over Southern Mediterranean regions.



**Fig. 10** Left: the observed percentage of seasonal mean precipitation that falls on days that have precipitation above the 99th percentile of daily precipitation for that season ( $R_{99}$ ). Third from right: relative RCM-PPE biases of  $R_{99}$ , second from right: relative GCM-PPE biases of  $R_{99}$ , right: relative differences between RCM-PPE and

GCM-PPE. Top row DJF, bottom row JJA. In the labels, ‘mean’ refers to the spatial average value, ‘median’ is the spatial median of the field being shown, and ‘corr’ is the spatial correlation of the datasets being differenced

In order to gain a more in depth understanding of the differences in rainfall distributions, supplementary material (figures S16–S26) shows the fractional contribution of different rainfall intensities to the total rainfall amount, for each region that we have NMS observations for. As in Berthou et al. (2019), a logarithmic axis has been used on the intensity ( $x$ ) axis, with the graph constructed from exponentially sized bins, so the total rainfall is proportional to the area under the curve as visualised. The following statistics are also reported:

$r99$ , the percentage of total rainfall falling above the 99th percentile

$S$ , the common area of the model and observed distributions as defined in Perkins et al. (2007), (expressed here as a percentage)

$dav = 100(S_{RCM} - S_{GCM})/S_{GCM}$ , the distribution added value as used by Soares and Cardoso (2018) to quantify added value in downscaling

$dry\_freq$ , the percentage of days with less than 0.1 mm of rain.

The fractional contribution plots reveal errors in the distribution of rainfall intensity, whilst dry day frequency allows us to examine errors in rainfall occurrence. In winter the dry day frequency is underestimated in both RCM-PPE and GCM-PPE by between a third and a half depending on region. In Scandinavia the under-simulation of dry days leads to a wet bias in the models, despite a very good distribution ( $S_{RCM} = S_{GCM} = 96\%$ ) of rainfall intensities. In general, up until around the 99th percentile, the RCM-PPE winter distributions for all regions are similar to that of GCM-PPE, just slightly shifted towards higher intensities. However high intensities contribute more to the RCM's total rainfall, resulting in the RCM having higher values of  $r99$  in five out of six regions. These higher values of  $r99$  are an improvement for Great Britain, Scandinavia, France, and the Carpathians, but not for Iberia.

In summer the under-simulation of dry days is reduced to less than 20% for mainland European regions (Spain, France, the Alps and the Carpathians). For all regions however the dry day frequency in RCM-PPE is worse than in GCM-PPE. Just like in winter, the intensity distribution for Scandinavia has excellent skill in both RCM-PPE ( $S = 97$ ) and GCM-PPE ( $S = 99$ ), but both ensembles have a wet bias due to raining too frequently. However, in all other regions the skill is lower in summer than in winter, with too much rainfall being contributed from low intensities (less than  $\sim 3$  mm), and not enough from moderate intensities (around 10 mm). This is true in both ensembles, but to a lesser extent in RCM-PPE as reflected in positive  $dav$  scores. The positive summer  $r99$  biases over mainland Europe appear due to differences in the bulk of the distribution, rather than in the contribution from

high intensity events. Specifically, GCM-PPE has too many days with less than 3 mm of rainfall.

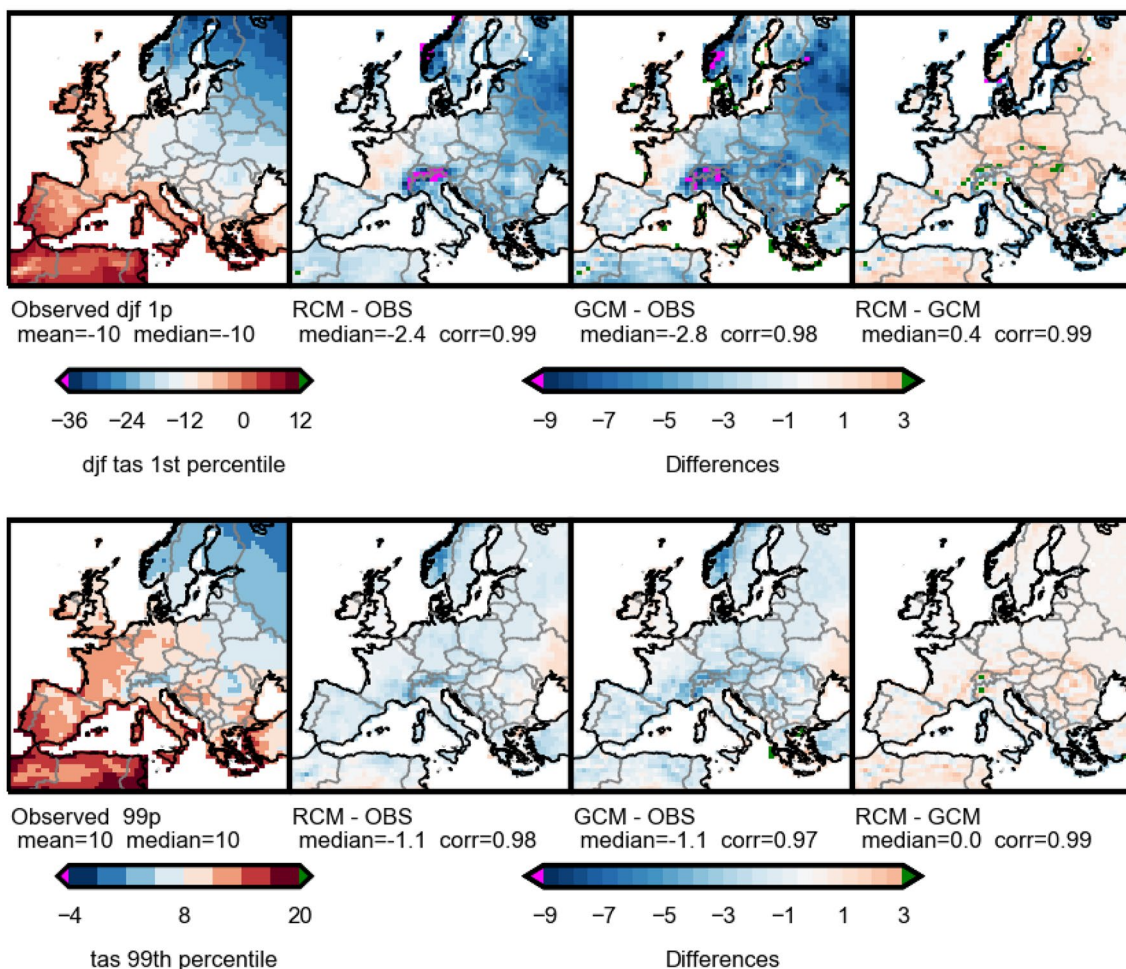
To summarise, in winter both ensembles underestimate dry day frequency by 30–50% depending on the region but do a generally good job at reproducing the observed intensity distribution. There is little difference in RCM-PPE and GCM-PPE performance except for extreme intensities, where RCM-PPE performs better. Skill is generally lower in summer than winter, with both ensembles having too much rainfall coming from low intensity days, however downscaling leads to an improvement in the intensity distribution. We have used  $r99$  as a measure of the shape of the precipitation distribution and have found that RCM-PPE has higher values in five out of six regions in winter, but lower values in the summer. This is an improvement in 5 out of six regions in winter, and 3 regions in summer. To conclude, we find that downscaling generally adds value, but not everywhere, and for summertime the dry day underestimation is actually worse in RCM-PPE.

### 3.3.2 Temperature

Figure 11 shows the ensemble mean of the 1st and 99th percentile of DJF daily temperature differences. In both ensembles the 1st percentile cold bias is generally larger than the mean bias (Fig. 2), which in turn is larger than the cold bias in the 99th percentile. In other words, in addition to being too cold on average, the spread in the daily temperature distributions is too large. The first percentile of RCM-PPE is warmer than GCM-PPE by 0.4 K on average, with some locations exceeding 3 K. This means the RCM-PPE has reduced the cold day bias present in the GCM-PPE, and furthermore RCM-PPE is showing an improvement in cold day extremes that goes beyond a simple shift of the entire daily temperature distribution.

Figure 12 shows the equivalent plot for JJA. The 99th percentile biases show a similar spatial pattern to the mean biases, namely that both ensembles are too warm in South East Europe, but too cool over Scandinavia. However, the 99th percentile warm bias is of greater magnitude and extends further into central Europe than the mean bias, a finding consistent with our suggestion of soil moisture feedbacks. Just like the seasonal mean temperature, RCM-PPE is cooler than GCM-PPE everywhere which represents a bias reduction in the 99th percentile for South and central Europe. Additionally, the spatial correlation with observations is higher in RCM-PPE (0.95) than in GCM-PPE (0.89).

Downscaling appears to be improving the distribution of daily temperature in winter, and hot extremes in summer. We know from Sect. 3.2.2 that there are some differences in mean temperature between the RCM and GCM that are not directly due to resolution. Such effects are also likely to influence the tail of the distribution, however additional



**Fig. 11** Left: the observed 1st (top) and 99th (bottom) percentile of daily DJF surface air temperature. Third from right: RCM-PPE biases, second from right: GCM-PPE biases, right: differences between RCM-PPE and GCM-PPE. Top row 1st percentile, bottom

row 99th percentile. In the labels, ‘mean’ refers to the spatial average value, ‘median’ is the spatial median of the field being shown, and ‘corr’ is the spatial correlation of the datasets being differenced

processes that control the occurrence and intensity of extremes may benefit more from improved resolution. An attempt to determine the source of the differences in the extremes between the simulations has been made using the 10 year simulations from Sect. 3.2.2 (see supplementary figure S27). Differences in extremes are not detected as statistically significantly (assessed using a bootstrapping procedure), possibly due to 10 years of data is not enough to assess differences in extremes. There are however suggestions that the 1st percentile winter differences are due to a combination of both resolution and other configuration differences. For JJA there are suggestions that on extreme hot days, the influence of EasyAerosol is reduced and the influence of the DLH error increased, compared on the seasonal mean contributions. One possible explanation for this is that summer temperature extremes occur on days with little cloud cover, with aerosol-radiation forcing having more influence. Figure S27 suggests that reduced biases in daily temperature

extremes may not solely be a genuine improvement due to higher resolution, with the role of other effects depending on region and season.

### 4 Discussion and conclusions

In this study, we have evaluated the performance of a new 12-member ensemble of regional climate simulations over Europe, carried out as part of the UKCP project. We have focussed on surface air temperature and precipitation and have found that there is generally good large scale consistency between RCM-PPE and GCM-PPE: both in terms of correlation between RCM members and their driving GCM members (domain average correlations above 0.9 for all variables and seasons) and spatial patterns (with median daily mean sea level correlations above 0.9 in both seasons). In winter this means that RCM-PPE inherits a large cold bias

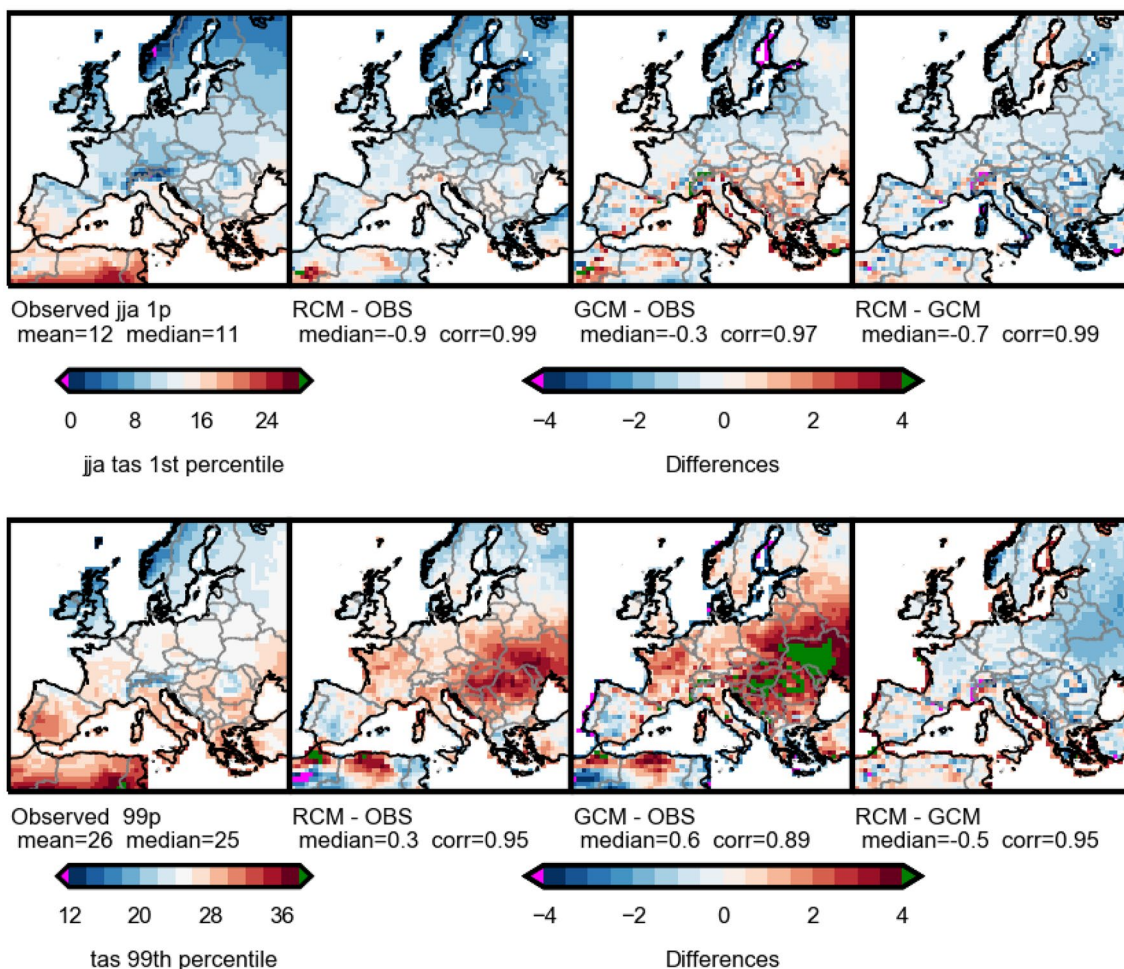


Fig. 12 as Fig. 11, but for JJA

over Scandinavia. Reasons for the cold bias in GCM-PPE, which is common to all members but varies in magnitude, are discussed in the UKCP report (Murphy et al. 2018). Contributing factors include a strong aerosol forcing, negative biases in long-wave cloud radiative forcing, a weak AMOC, and too low westerly mid-latitude wind speed. Consistent with this, we find a negative longwave cloud radiative forcing bias in the North East of the RCM domain, but note that the net DJF longwave surface radiation bias is actually positive. The cold bias is reduced slightly in RCM-PPE, and over Scandinavia we show here that this is a consequence of increased resolution (due in part to more cloud), rather than due to the use of EasyAerosol or the day light hours error. It is argued in Murphy et al. (2018) that the cold bias may mean that there is too much snow coverage in the present day, with consequent potential to cause too large an albedo feedback in the climate change projections.

In summer we find that both RCM-PPE and GCM-PPE have a cold bias over Scandinavia, and a warm bias over South East Europe. Warm dry summer biases in South East

Europe are also found in EURO-CORDEX RCMs (Kotlarski et al. 2014), and are believed to be related to local soil moisture feedbacks. The South East Europe warm dry bias is reduced in RCM-PPE compared to GCM-PPE, however we show that this may be due to the use of EasyAerosol leading to increased cloudiness, and to a lesser extent the day light hours error making the shortwave aerosol forcing stronger in RCM-PPE than GCM-PPE.

The use of the ERAI-RCM-STD simulation further confirms that RCM-PPE's Scandinavian winter cold bias is largely inherited from GCM-PPE. In summer ERAI-RCM-STD is warmer than RCM-STD over land, despite having cooler SSTs for most of the domain. As in winter, we find that ERAI-RCM-STD shows improvements in cloud and radiation biases compared to RCM-STD. In South East Europe there is an indication of compensating errors in RCM-PPE as radiation bias improvements in ERAI-RCM-STD contribute to an increased surface temperature warm bias. Differences between ERAI-RCM-STD and RCM-STD could either be due to the use of observed boundary data, or

due to different EasyAerosol inputs. However, further analysis is required to identify the relative roles of these.

Despite the generally good large-scale consistency, RCM-PPE JJA ensemble mean is cooler by 0.5 K on average, with all RCM members being cooler than their driving GCM. Additionally, in both seasons RCM-PPE is wetter than GCM-PPE by  $\sim 0.3$  mm/day. We have used two additional 10 year simulations to try and identify to what extent these differences are due to resolution, or other differences in the configuration of RCM-PPE and GCM-PPE. Where we have found that configuration differences between RCM-PPE and GCM-PPE are contributing, we have assumed that these are predominately due to the use of EasyAerosol. We recognise however that there are other configuration differences as well as EasyAerosol that could be contributing, for instance the stochastic physics package that is only in GCM-PPE can impact mean climate via noise induced drift (Sanchez 2016). The summertime temperature differences over Eastern Europe are predominantly due to configuration differences (EasyAerosol plus other). Configuration differences also introduce a warm difference in winter (around 0.4–0.8 K) over France and Germany that is not detected as statistically significant in the total difference between RCM and GCM, presumably due to resolution adding increased variability. Increased precipitation is also seen in the same regions where configuration differences are impacting temperature. Reinforcing our view that the predominant configuration difference is EasyAerosol, our results are consistent with those of Bellouin and Thornhill (2018), where it is found that the simulation using EasyAerosol has increased liquid cloud cover and thickness, and consequently reduced downward shortwave, and increased downward longwave radiation. These differences were linked to the prescription of monthly mean cloud droplet number concentrations. EasyAerosol prescribes monthly mean properties from the driving GCM. As such there is no sub-monthly variability, meaning some differences are inevitable where there are non-linearities, as is the case for aerosol cloud interactions.

The use of present day aerosol climatologies in some EURO-CORDEX RCM simulations are found to be the reason for diverging projections of surface solar radiation (Gutiérrez et al. 2020), and JJA surface air temperature and precipitation (Boé et al. 2020). The difficulty is that full aerosol modelling in RCMs is often prohibitively expensive, particularly as we move to convection permitting resolutions. Consequently, the use of a scheme like EasyAerosol, that is intended to represent the effects of aerosols on radiation and clouds at lower cost, is arguably essential. Following the findings here and in Bellouin and Thornhill (2018), we recommend further consideration on the treatment of CDNC within EasyAerosol. Ideas being trialled include applying a scaling factor to CDNC, and a scheme enhancement that uses both the mean and variance of CDNC

(Bellouin and Thornhill 2018). We note, however, that the use of EasyAerosol even though imperfect remains a big step forward for regional modelling, which until now has typically used aerosol climatologies for the present day.

The findings here are also a reminder that reduced RCM biases with respect to those of the GCMs, as is the case in JJA over South East Europe, should not automatically be considered an improved representation of processes due to increased resolution. A similar point is made in García-Díez et al. (2015) who argue for evaluating more variables than temperature and precipitation in order to reduce the risk of compensation of errors between variables.

In this study, we have also looked at the distribution of daily precipitation and temperature, two variables commonly used in impacts assessments. For precipitation there is little difference in RCM-PPE and GCM-PPE winter performance, both ensembles underestimate dry day frequency by 30–50% depending on the region but do a generally good job at reproducing the observed intensity distribution. Skill is generally lower in summer than winter, and both ensembles have too much rainfall coming from low intensity days, however downscaling shows a skill increase. We have used r99 as a measure of the shape of the precipitation distributions and have found that RCM-PPE has higher values in five out of six regions in winter, but lower values in the summer. Very similar findings are found when comparing precipitation distributions in pairs of 50 km and 12 km Euro-CORDEX simulations (Prein et al. 2016), in particular that increased resolution causes a summertime reduction in dry day frequency, increased mean and heavy precipitation. Prein et al. (2016) suggest that the differences are due to the improved representation of orography, and in summer having the larger scales of convection captured by the resolved-scale dynamics. Given the similarity of these findings, we conclude that the differences in daily rainfall distribution beyond an increase in mean precipitation in Central/Eastern Europe in winter/summer, are likely due to resolution rather than the use of EasyAerosol. For winter temperature both ensembles are too cold on average, and the spread of daily temperatures being too large. However, RCM-PPE 1st percentile is warmer than GCM-PPE, with the warm difference being greater than the warm difference in the seasonal mean, meaning the improvement goes beyond a simple shift of the entire temperature distribution.

In addition to providing climate change projections for Europe, RCM-PPE also provides boundary data to a convection permitting ensemble (CPM-12) over the UK (Kendon et al. 2019). In testing the configuration of CPM-12, it was found that the use of a 12 km nest improved CPM performance and was also cost effective as it reduced the required domain size (Fosser et al. 2020). Differences in precipitation are found to be much larger between CPM-12 and RCM-PPE, than between RCM-PPE and GCM-PPE.

For instance, Kendon et al. (2019) find that the rms error of dry day frequency in CPM-12 is half that of RCM-PPE for all seasons. Additionally, CPM-12 sees large improvements in the simulation of hourly rainfall. Thus the CPM-12 data provides additional opportunities for climate assessments over the UK, alongside the RCM-PPE data, but this is beyond the scope of the current paper.

Finally, we comment on our relative confidence in future climate projections from RCM-PPE and GCM-PPE, based on the results from this current study. Overall, RCM-PPE is generally consistent with the GCM-PPE meaning that it shares many of its strengths and deficiencies. RCM-PPE does however add local details and improves the representation of cold winter days, winter precipitation extremes, and summertime daily precipitation intensity, but not dry day frequency. We also note that cloud properties are considerably different in the RCM, with a change in vertical distribution, an increase in cloud cover, but reduced cloud thickness. These changes lead to RCM-PPE reducing GCM-PPE's winter cold bias over Scandinavia. Thus, we would recommend that users interested in projections of daily extremes (temperature and precipitation) or Scandinavian winter temperature use the RCM-PPE projections in preference to the GCM-PPE projections. On the other hand, we have an inconsistency with the driving GCM in Eastern and Central Europe that is not a direct consequence of resolution. Users interested in changes to mean climate over Eastern or Central European regions will therefore need to balance this inconsistency with their requirements for higher resolution data. Although not shown here, we also comment that for some seasons and regions, RCM-PPE climate change projections span a smaller range of changes than those in EURO-CORDEX. It has also been shown that the EURO-CORDEX projections do not include the warmest and driest summertime projections that are present in CMIP5 (Boé 2020; Coppola et al. 2020). In the UKCP18 report (Murphy et al. 2018) projections from the global perturbed parameter ensemble were augmented with 13 models selected from the CMIP5 ensemble (McSweeney 2018). A similar approach of augmenting the regional perturbed parameter ensemble may be required for users who require a broader uncertainty context.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00382-021-05941-3>.

**Acknowledgements** S.O.T, E.J.K. and J.M. gratefully acknowledge funding from the Joint UK BEIS/Defra Met Office Hadley Centre Climate Programme Grant no. (GA01101). We acknowledge the E-OBS dataset from the EU-FP6 project UERRA (<http://www.uerra.eu>) and the Copernicus Climate Change Service, and the data providers in the ECA&D project (<https://www.ecad.eu>)

## Declarations

**Conflicts of interest** The authors declare that they have no competing interests.

**Data availability** Model data is publicly available. Met Office Hadley Centre (2018): UKCP18 Global Projections at 60 km Resolution for 1900–2100. Centre for Environmental Data Analysis, last access: 5 February 2021). <https://catalogue.ceda.ac.uk/uuid/97bc0c622a24489aa105f5b8a8efa3f0> Met Office Hadley Centre (2018): UKCP18 Regional Projections at 12 km Resolution for 1980–2080. Centre for Environmental Data Analysis, (last access: 5 February 2021). <https://catalogue.ceda.ac.uk/uuid/8c6c0ae2c25947168826a70d2241b797>

**Code availability** Due to intellectual property rights restrictions, we cannot provide either the source code or documentation papers for the UM or JULES. Obtaining the UM. The Met Office Unified Model is available for use under licence. A number of research organisations and national meteorological services use the UM in collaboration with the Met Office to undertake basic atmospheric process research, produce forecasts, develop the UM code, and build and evaluate Earth system models. For further information on how to apply for a licence, see <http://www.metoffice.gov.uk/research/modelling-systems/unified-model> (last access: 5 February 2021). Obtaining JULES. JULES is available under licence, free of charge. For further information on how to gain permission to use JULES for research purposes see [http://jules-lsm.github.io/access\\_req/JULES\\_access.html](http://jules-lsm.github.io/access_req/JULES_access.html) (last access: 5 February 2021). Any scripts used in analysing the data are available on request.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Baek HJ, Lee J, Lee HS, Hyun YK, Cho C, Kwon WT, Marzin C, Gan SY, Kim MJ, Choi DH, Lee J, Lee J, Boo KO, Kang HS, Byun YH (2013) Climate change in the 21st century simulated by HadGEM2-AO under representative concentration pathways. *Asia-Pac J Atmos Sci* 49(5):603–618. <https://doi.org/10.1007/s13143-013-0053-7>
- Bellouin N, Thornhill G (2018) PProcess-based climate sIMulation: AdVances in high resolution modelling and European climate Risk Assessment Deliverable D3.1 Quantification of robustness of aerosol-radiation-cloud interactions across models and resolution. [https://www.primavera-h2020.eu/primavera/static/media/uploads/d3.1\\_aerosol\\_rad\\_cloud\\_interactions.pdf](https://www.primavera-h2020.eu/primavera/static/media/uploads/d3.1_aerosol_rad_cloud_interactions.pdf)
- Berthou S, Kendon EJ, Rowell DP, Roberts MJ, Tucker S, Stratton RA (2019) Larger future intensification of rainfall in the West African Sahel in a convection-permitting model. *Geophys Res Lett* 46(22):13299–13307. <https://doi.org/10.1029/2019GL083544>
- Boé J, Somot S, Corre L, Nabat P (2020) Large discrepancies in summer climate change over Europe as projected by global and

- regional climate models: causes and consequences. *Clim Dyn* 54(5–6):2981–3002. <https://doi.org/10.1007/s00382-020-05153-1>
- Burke EJ, Perry RHJ, Brown SJ (2010) An extreme value analysis of UK drought and projections of change in the future. *J Hydrol* 388(1–2):131–143. <https://doi.org/10.1016/j.jhydrol.2010.04.035>
- Casanueva A, Kotlarski S, Herrera S, Fernández J, Gutiérrez JM, Boberg F, Colette A, Christensen OB, Goergen K, Jacob D, Keuler K, Nikulin G, Teichmann C, Vautard R (2016) Daily precipitation statistics in a EURO-CORDEX RCM ensemble: added value of raw and bias-corrected high-resolution simulations. *Clim Dyn* 47(3–4):719–737. <https://doi.org/10.1007/s00382-015-2865-x>
- Coppola E, Nogherotto R, Ciarlò JM, Giorgi F, Meijgaard E, Kadyrov N, Iles C, Corre L, Sandstad M, Somot S, Nabat P, Vautard R, Levavasseur G, Schwingshackl C, Sillmann J, Kjellström E, Nikulin G, Aalbers E, Lenderink G, Wulfmeyer V (2020) Assessment of the European climate projections as simulated by the large EURO-CORDEX regional and global climate model ensemble. *J Geophys Res Atmos*. <https://doi.org/10.1029/2019jd032356>
- Cornes RC, van der Schrier G, van den Besselaar EJM, Jones PD (2018) An ensemble version of the E-OBS temperature and precipitation data sets. *J Geophys Res Atmos* 123(17):9391–9409. <https://doi.org/10.1029/2017JD028200>
- Davies T (2014) Lateral boundary conditions for limited area models. *Q J R Meteorol Soc* 140(678):185–196. <https://doi.org/10.1002/qj.2127>
- Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Vitart F (2011) The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol Soc* 137(656):553–597. <https://doi.org/10.1002/qj.828>
- Edwards JM, Slingo A (1996) Studies with a flexible new radiation code I: choosing a configuration for a large-scale model. *Q J R Meteorol Soc* 122(531):689–719. <https://doi.org/10.1002/qj.49712253107>
- Fantini A, Raffaele F, Torma C, Bacer S, Coppola E, Giorgi F, Ahrens B, Dubois C, Sanchez E, Verdecchia M (2018) Assessment of multiple daily precipitation statistics in ERA-Interim driven Med-CORDEX and EURO-CORDEX experiments against high resolution observations. *Clim Dyn* 51(3):877–900. <https://doi.org/10.1007/s00382-016-3453-4>
- Fosser G, Kendon E, Chan S, Lock A, Roberts N, Bush M (2020) Optimal configuration and resolution for the first convection-permitting ensemble of climate projections over the United Kingdom. *Int J Climatol* 40(7):3585–3606. <https://doi.org/10.1002/joc.6415>
- 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. <https://doi.org/10.1007/s00382-015-2529-x>
- Goodison B, Louie P, Yang D (1997) WMO solid precipitation measurement intercomparison final report
- Gutiérrez C, Somot S, Nabat P, Mallet M, Corre L, Van Meijgaard E, Perpiñán O, Gaertner M (2020) Future evolution of surface solar radiation and photovoltaic potential in Europe: investigating the role of aerosols. *Environ Res Lett* 15(3):034035. <https://doi.org/10.1088/1748-9326/ab6666>
- 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(1):74–85. <https://doi.org/10.1002/joc.2256>
- Herrera S, Kotlarski S, Soares PMM, Cardoso RM, Jaczewski A, Gutiérrez JM, Maraun D (2019) Uncertainty in gridded precipitation products: influence of station density, interpolation method and grid resolution. *Int J Climatol* 39(9):3717–3729. <https://doi.org/10.1002/joc.5878>
- Hurttt GC, Chini LP, Frolking S, Betts RA, Feddema J, Fischer G, Fisk JP, Hibbard K, Houghton RA, Janetos A, Jones CD, Kindermann G, Kinoshita T, Klein Goldewijk K, Riahi K, Shevliakova E, Smith S, Stehfest E, Thomson A, Wang YP (2011) Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Clim Change* 109(1):117–161. <https://doi.org/10.1007/s10584-011-0153-2>
- Isotta FA, Frei C, Weigluni V, Perčec Tadić M, Lassègues P, Rudolf B, Pavan V, Cacciamani C, Antolini G, Ratto SM, Munari M, Micheletti S, Bonati V, Lussana C, Ronchi C, Panettieri E, Marigo G, Vertačnik G (2014) The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. *Int J Climat* 34(5):1657–1675. <https://doi.org/10.1002/JOC.3794>
- 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, Yiou P (2014) EURO-CORDEX: new high-resolution climate change projections for European impact research. *Reg Environ Change* 14(2):563–578. <https://doi.org/10.1007/s10113-013-0499-2>
- Jacob D, Teichmann C, Sobolowski S, Katragkou E, Anders I, Belda M, Benestad R, Boberg F, Buonomo E, Cardoso RM, Casanueva A, Christensen OB, Christensen JH, Coppola E, De Cruz L, Davin EL, Dobler A, Domínguez M, Fealy R, Wulfmeyer V (2020) Regional climate downscaling over Europe: perspectives from the EURO-CORDEX community. *Reg Environ Change* 20(2):1–20. <https://doi.org/10.1007/s10113-020-01606-9>
- Jones RG, Murphy JM, Noguer M (1995) Simulation of climate change over Europe using a nested regional-climate model I: assessment of control climate, including sensitivity to location of lateral boundaries. *Q J R Meteorol Soc* 121(526):1413–1449. <https://doi.org/10.1002/qj.49712152610>
- Jones RG, Murphy JM, Noguer M, Keen AB (1997) Simulation of climate change over Europe using a nested regional-climate model II: comparison of driving and regional model responses to a doubling of carbon dioxide. *Q J R Meteorol Soc* 123(538):265–292. <https://doi.org/10.1002/qj.49712353802>
- Jones CD, Hughes JK, Bellouin N, Hardiman SC, Jones GS, Knight J, Liddicoat S, O'Connor FM, Andres RJ, Bell C, Boo KO, Bozzo A, Butchart N, Cadule P, Corbin KD, Doutriaux-Boucher M, Friedlingstein P, Gornall J, Gray L, Zerroukat M (2011) The HadGEM2-ES implementation of CMIP5 centennial simulations. *Geosci Model Dev* 4(3):543–570. <https://doi.org/10.5194/gmd-4-543-2011>
- Karlsson K-G, Anttila K, Trentmann J, Stengel M, FokkeMeirink J, Devasthale A, Hanschmann T, Kothe S, Jääskeläinen E, Sedlar J, Benas N, van Zadelhoff G-J, Schlundt C, Stein D, Finkensieper S, Håkansson N, Hollmann R (2017) CLARA-A2: the second edition of the CM SAF cloud and radiation data record from 34 years of global AVHRR data. *Atmos Chem Phys* 17(9):5809–5828. <https://doi.org/10.5194/acp-17-5809-2017>
- Kato S, Rose FG, Rutan DA, Thorsen TJ, Loeb NG, Doelling DR, Huang X, Smith WL, Su W, Ham SH (2018) Surface irradiances of edition 4.0 Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) data product. *J Clim* 31(11):4501–4527. <https://doi.org/10.1175/JCLI-D-17-0523.1>
- Kay AL, Jones DA (2012) Transient changes in flood frequency and timing in Britain under potential projections of climate change. *Int J Climatol* 32(4):489–502. <https://doi.org/10.1002/joc.2288>
- Kendon E, Fosser G, Murphy J, Chan S, Clark R, Harris G, Lock A, Lowe J, Martin G, Pirret J, Roberts N, Sanderson M, Tucker

- S, Hoskins B, Kjellström E, Schär C, Van Den Hurk B (2019) UKCP convection-permitting model projections: science report. [www.metoffice.gov.uk](http://www.metoffice.gov.uk)
- Knist S, Goergen K, Buonomo E, Christensen OB, Colette A, Cardoso RM, Fealy R, Fernández J, García-Díez M, Jacob D, Kartios S, Katragkou E, Keuler K, Mayer S, Van Meijgaard E, Nikulin G, Soares PMM, Sobolowski S, Szepszo G, Simmer C (2017) Land-atmosphere coupling in EURO-CORDEX evaluation experiments. *J Geophys Res* 122(1):79–103. <https://doi.org/10.1002/2016JD025476>
- 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. *Geosci Model Dev* 7(4):1297–1333. <https://doi.org/10.5194/gmd-7-1297-2014>
- Lean JL (2009) Calculations of solar irradiance: monthly means from 1882 to 2008, annual means from 1610 to 2008
- Loeb NG, Doelling DR, Wang H, Su W, Nguyen C, Corbett JG, Liang L, Mitrescu C, Rose FG, Kato S (2018) Clouds and the earth's radiant energy system (CERES) energy balanced and filled (EBAF) top-of-atmosphere (TOA) edition-4.0 data product. *J Clim* 31(2):895–918. <https://doi.org/10.1175/JCLI-D-17-0208.1>
- Loveland TR, Reed BC, Ohlen DO, Brown JF, Zhu Z, Yang L, Merchant JW (2000) Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *Int J Remote Sens* 21(6–7):1303–1330. <https://doi.org/10.1080/014311600210191>
- Mann GW, Carslaw KS, Spracklen DV, Ridley DA, Manktelow PT, Chipperfield MP, Pickering SJ, Johnson CE (2010) Description and evaluation of GLOMAP-mode: a modal global aerosol microphysics model for the UKCA composition-climate model. *Geosci Model Dev* 3(2):519–551. <https://doi.org/10.5194/gmd-3-519-2010>
- Manners J, Edwards J, Hill P, Thelen J-C (2015) SOCRATES technical guide suite of community RAdiative transfer codes based on Edwards and Slingo. [https://code.metoffice.gov.uk/trac/socrates/attachment/wiki/WikiStart/socrates\\_techguide.pdf](https://code.metoffice.gov.uk/trac/socrates/attachment/wiki/WikiStart/socrates_techguide.pdf). Accessed 2021
- McCull L, Palin EJ, Thornton HE, Sexton DMH, Betts R, Mylne K (2012) Assessing the potential impact of climate change on the UK's electricity network. *Clim Change* 115(3–4):821–835. <https://doi.org/10.1007/s10584-012-0469-6>
- McDonald R, Sexton D, Mcsweeney C, Yamazaki K, Rostron J (n.d.) Evaluation of storms in an ensemble of coupled model simulations and the sensitivity to parameter perturbations. In: Preparation for climate dynamics
- Mcsweeney CF (2018) Selection of CMIP5 members to augment a perturbed-parameter ensemble of global realisations of future climate for the UKCP18 scenarios. Hadley Centre Technical Note. [https://digital.nmla.metoffice.gov.uk/IO\\_28f93601-3178-4c48-8621-fb4388ec66a1/](https://digital.nmla.metoffice.gov.uk/IO_28f93601-3178-4c48-8621-fb4388ec66a1/)
- McSweeney C, Sexton D, Bett P, McDonald R, Thornton H, Drouard M, Rostron J, Yamazaki K, Murphy J, Woollings T (2021) Evaluating regional drivers of mean UK temperature and precipitation in a coupled perturbed parameter ensemble (in preparation)
- Moss RH, Edmonds JA, Hibbard KA, Manning MR, Rose SK, Van Vuuren DP, Carter TR, Emori S, Kainuma M, Kram T, Meehl GA, Mitchell JFB, Nakicenovic N, Riahi K, Smith SJ, Stouffer RJ, Thomson AM, Weyant JP, Wilbanks TJ (2010) The next generation of scenarios for climate change research and assessment. *Nature* 463(7282):747–756. <https://doi.org/10.1038/nature08823>
- Murphy JM, Sexton DMH, Jenkins GJ, Boorman PM, Booth BBB, Brown CC, Clark RT, Collins M, Harris GM, Kendon EJ, Betts RA, Brown SJ, Howard TP, Humphrey KA, McCarthy MP, McDonald RE, Stephens A, Wallace C, Warren R, Wilby RW (n.d.) UK climate projections science report: climate change projections. [http://cedadocs.ceda.ac.uk/1320/1/climate\\_projections\\_full\\_report.pdf](http://cedadocs.ceda.ac.uk/1320/1/climate_projections_full_report.pdf)
- Murphy JM, Harris GR, Sexton DMH, Kendon EJ, Bett PE, Clark RT, Eagle KE, Fosse G, Fung F, Lowe JA, McDonald RE, McInnes RN, Mcsweeney CF, Mitchell JFB, Rostron JW, Thornton HE, Tucker S, Yamazaki K (2018) (*No Title*). <https://www.theccc.org.uk/tackling-climate>
- Oki T, Sud YC (1998) Design of total runoff integrating pathways (TRIP)—a global river channel network. *Earth Interact.* [https://doi.org/10.1175/1087-3562\(1998\)002<0001:DOTRIP>2.3.CO;2](https://doi.org/10.1175/1087-3562(1998)002<0001:DOTRIP>2.3.CO;2)
- Palin EJ, Thornton HE, Mathison CT, McCarthy RE, Clark RT, Dora J (2013) Future projections of temperature-related climate change impacts on the railway network of Great Britain. *Clim Change* 120(1–2):71–93. <https://doi.org/10.1007/s10584-013-0810-8>
- Perkins SE, Pitman AJ, Holbrook NJ, McAneney J (2007) Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *J Clim* 20(17):4356–4376. <https://doi.org/10.1175/JCLI4253.1>
- Perry M, Hollis D (2005) The generation of monthly gridded datasets for a range of climatic variables over the UK. *Int J Climatol* 25(8):1041–1054. <https://doi.org/10.1002/joc.1161>
- Prein AF, Gobiet A, Truhetz H, Keuler K, Goergen K, Teichmann C, Fox Maule C, van Meijgaard E, Déqué M, Nikulin G, Vautard R, Colette A, Kjellström E, Jacob D (2016) Precipitation in the EURO-CORDEX 0.11° and 0.44° simulations: high resolution, high benefits? *Clim Dyn* 46(1–2):383–412. <https://doi.org/10.1007/s00382-015-2589-y>
- Prudhomme C, Young A, Watts G, Haxton T, Crooks S, Williamson J, Davies H, Dadson S, Allen S (2012) The drying up of Britain? A national estimate of changes in seasonal river flows from 11 Regional Climate Model simulations. *Hydrol Process* 26(7):1115–1118. <https://doi.org/10.1002/hyp.8434>
- Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang W (2002) An improved in situ and satellite SST analysis for climate. [https://doi.org/10.1175/1520-0442\(2002\)015<1609:AIISAS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2)
- Ridley JK, Blockley EW, Keen AB, Rae JGL, West AE, Schroeder D (2017) The sea ice model component of HadGEM3-GC3.1. <https://doi.org/10.5194/gmd-2017-212>
- Righi M, Andela B, Eyring V, Lauer A, Predoi V, Schlund M, Vegas-Regidor J, Bock L, Brötz B, De Mora L, Diblen F, Dreyer L, Drost N, Earnshaw P, Hassler B, Koldunov N, Little B, Loosveldt Tomas S, Zimmermann K (2020) Earth system model evaluation tool (ESMValTool) v2.0-technical overview. *Geosci Model Dev* 13(3):1179–1199. <https://doi.org/10.5194/gmd-13-1179-2020>
- Rummukainen M (2010) State-of-the-art with regional climate models. *Wiley Interdiscip Rev Clim Change* 1(1):82–96. <https://doi.org/10.1002/wcc.8>
- Rummukainen M (2016) Added value in regional climate modeling. *Wiley Interdiscip Rev Clim Change* 7(1):145–159. <https://doi.org/10.1002/wcc.378>
- Sanchez C, Williams KD, Collins M (2016) Improved stochastic physics schemes for global weather and climate models. *Q J R Meteorol Soc* 142(694):147–159. <https://doi.org/10.1002/qj.2640>
- Sanderson MG, Wiltshire AJ, Betts RA (2012) Projected changes in water availability in the United Kingdom. *Water Resour Res* 48:8. <https://doi.org/10.1029/2012WR011881>
- Sato M, Hansen JE, McCormick MP, Pollack JB (1993) Stratospheric aerosol optical depths, 1850–1990. *J Geophys Res* 98(D12):22987–22994. <https://doi.org/10.1029/93jd02553>



- Scaife AA, Spanghel T, Fereday DR, Cubasch U, Langematz U, Akiyoshi H, Bekki S, Braesicke P, Butchart N, Chipperfield MP, Gettelman A, Hardiman SC, Michou M, Rozanov E, Shepherd TG (2012) Climate change projections and stratosphere-troposphere interaction. *Clim Dyn* 38(9–10):2089–2097. <https://doi.org/10.1007/s00382-011-1080-7>
- Scaife AA, Arribas A, Blockley E, Brookshaw A, Clark RT, Dunstone N, Eade R, Fereday D, Folland CK, Gordon M, Hermanson L, Knight JR, Lea DJ, MacLachlan C, Maidens A, Martin M, Peterson AK, Smith D, Vellinga M, Williams A (2014) Skillful long-range prediction of European and North American winters. *Geophys Res Lett* 41(7):2514–2519. <https://doi.org/10.1002/2014GL059637>
- 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 Sci Rev* 99(3–4):125–161. <https://doi.org/10.1016/j.earscirev.2010.02.004>
- Sevrak B, Hamon W (1984) World meteorological organization instruments and observing methods report no. 17 international comparison of National Precipitation Gauges with a Reference Pit Gauge Secretariat of the World Meteorological Organization
- Sexton DMH, McSweeney CF, Rostron JW, Yamazaki K, Booth BBB, Murphy JM, Regayre L, Johnson JS, Karmalkar AV (2021) A perturbed parameter ensemble of HadGEM3-GC3.05 coupled model projections: part 1: selecting the parameter combinations. *Clim Dyn* 1:3. <https://doi.org/10.1007/s00382-021-05709-9>
- Soares PMM, Cardoso RM (2018) A simple method to assess the added value using high-resolution climate distributions: application to the EURO-CORDEX daily precipitation. *Int J Climatol* 38(3):1484–1498. <https://doi.org/10.1002/joc.5261>
- Spinoni J, Szalai S, Szentimrey T, Lakatos M, Bihari Z, Nagy A, Németh Á, Kovács T, Mihic D, Dacic M, Petrovic P, Kržič A, Hiebl J, Auer I, Milkovic J, Štěpánek P, Zahradníček P, Kilar P, Limanowka D, Vogt J (2015) Climate of the Carpathian Region in the period 1961–2010: climatologies and trends of 10 variables. *Int J Climatol* 35(7):1322–1341. <https://doi.org/10.1002/joc.4059>
- Stevens B, Fiedler S, Kinne S, Peters K, Rast S, Müsse J, Smith SJ, Mauritsen T (2017) MACv2-SP: a parameterization of anthropogenic aerosol optical properties and an associated Twomey effect for use in CMIP6. *Geosci Model Dev* 10(1):433–452. <https://doi.org/10.5194/gmd-10-433-2017>
- Storkey D, Blaker AT, Mathiot P, Megann A, Aksenov Y, Blockley EW, Calvert D, Graham T, Hewitt HT, Hyder P, Kuhlbrodt T, Rae JGL, Sinha B (2018) UK Global Ocean GO6 and GO7: a traceable hierarchy of model resolutions. *Geosc Model Dev* 11(8):3187–3213. <https://doi.org/10.5194/gmd-11-3187-2018>
- Torma C, Giorgi F, Coppola E (2015) Added value of regional climate modeling over areas characterized by complex terrain-Precipitation over the Alps. *J Geophys Res Atmos* 120(9):3957–3972. <https://doi.org/10.1002/2014JD022781>
- Trigo IF, Davies TF, Bigg G (n.d.). Objective climatology of cyclones in the mediterranean region in: *Journal of Climate* Volume 12 Issue 6 (1999). [https://journals.ametsoc.org/view/journals/clim/12/6/1520-0442\\_1999\\_012\\_1685\\_ococit\\_2.0.co\\_2.xml?tab\\_body=fulltext-display](https://journals.ametsoc.org/view/journals/clim/12/6/1520-0442_1999_012_1685_ococit_2.0.co_2.xml?tab_body=fulltext-display). Accessed 14 Dec 2020
- Tveito OE, Bjørndal I, Skjelvåg AO, Aune B (2005) A GIS-based agro-ecological decision system based on gridded climatology. *Meteorol Appl* 12(1):57–68. <https://doi.org/10.1017/S1350482705001490>
- 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, Yiou P (2013) The simulation of European heat waves from an ensemble of regional climate models within the EURO-CORDEX project. *Clim Dyn* 41(9–10):2555–2575. <https://doi.org/10.1007/s00382-013-1714-z>
- Vautard R, Kadyrov N, Iles C, Boberg F, Buonomo E, Bülow K, Coppola E, Corre L, Meijgaard E, Nogherotto R, Sandstad M, Schwingshackl C, Somot S, Aalbers E, Christensen OB, Ciarlò JM, Demory M, Giorgi F, Jacob D, Wulfmeyer V (2020) Evaluation of the large EURO-CORDEX regional climate model ensemble. *J Geophys Res Atmos*. <https://doi.org/10.1029/2019jd032344>
- Vidal J-P, Martin E, Franchistéguy L, Baillon M, Soubeyroux J-M (2010) A 50-year high-resolution atmospheric reanalysis over France with the Safran system. *Int J Climatol* 30(11):1627–1644. <https://doi.org/10.1002/joc.2003>
- Walters D, Baran A, Boutle I, Brooks M, Earnshaw P, Edwards J, Furtado K, Hill P, Lock A, Manners J, Morcrette C, Mulcahy J, Sanchez C, Smith C, Stratton R, Tennant W, Tomassini L, Van Weverberg K, Vosper S, Zerroukat M (2017) The met office unified model global atmosphere 7.0/7.1 and JULES global land 7.0 configurations. *Geosci Model Dev Discuss*. <https://doi.org/10.5194/gmd-2017-291>
- Williams KD, Copsey D, Blockley EW, Bodas-Salcedo A, Calvert D, Comer R, Davis P, Graham T, Hewitt HT, Hill R, Hyder P, Ineson S, Johns TC, Keen AB, Lee RW, Megann A, Milton SF, Rae JGL, Roberts MJ, Xavier PK (2018) The Met Office Global Coupled Model 3.0 and 3.1 (GC3.0 and GC3.1) configurations. *J Adv Model Earth Syst* 10(2):357–380. <https://doi.org/10.1002/2017MS001115>
- Yamazaki K, David, Sexton MH, Rostron JW, Mcsweeney CF, Murphy JM, Harris GR (2021) A perturbed parameter ensemble of HadGEM3-GC3.05 coupled model projections: part 2: global performance and future changes, 1, 3. <https://doi.org/10.1007/s00382-020-05608-5>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.