A comparison of regional climate projections with a range of climate sensitivities

Clair Barnes[1], Richard E Chandler[2], and Christopher Metcalfe Brierley[3]

[1]Imperial College London
[2]UCL
[3]University College London

December 19, 2023

A comparison of regional climate projections with a range of climate sensitivities

Clair R. Barnes[1,2], Richard E. Chandler[1] and Christopher M. Brierley[3]

[1]Department of Statistical Science, University College London, Gower Street, London WC1E 6BT
[2]Grantham Institute, Imperial College London, Exhibition Road, London SW7 2AZ
[3]Department of Geography, University College London, Gower Street, London WC1E 6BT

Key Points:

- The UKCP and EuroCORDEX regional model ensembles have similar biases, but project very different future climate over the UK
- These differences are driven largely by differences in the climate sensitivity of the GCMs used to force the regional models
- Comparing projections after a specified degree of warming, rather than in specified decades, reduces but does not resolve these differences

Corresponding author: Clair Barnes, c.barnes22@imperial.ac.uk

Abstract

To investigate the extent to which differences in regional model projections can be explained by differences in the warming rates of their driving models, we compare projections of temperature and precipitation over the UK from two regional climate ensembles – the EuroCORDEX multi-model ensemble and UKCP18 perturbed parameter ensemble – along with projections produced by the 'parent' GCMs from which boundary conditions were taken. We evaluate the ensembles in terms of their representation of recent climate, then compare the changes simulated between 1981–2010 and 2050–2079. While both ensembles exhibit seasonal biases with similar magnitudes and spatial patterns during the evaluation period, the UKCP18 ensemble exhibits a somewhat stronger change signal in future simulations, due to a combination of higher climate sensitivity of the driving models, variations in the forcings applied, and - in the regional simulations - the inclusion of time-varying aerosols.

In order to reconcile the two sets of projections, we compare two periods corresponding to fixed global warming levels in the driving models, to constrain the variability within and between the ensembles which can be ascribed to differing rates of global warming: the discrepancy between the ensembles is greatly reduced, although some differences in the local response remain, with the UKCP18 runs slightly warmer and drier than the EuroCORDEX runs, particularly in summer. We also highlight potential pitfalls of comparing warming levels with a reference time period, due to uncertainty about the warming that has already occurred in the driving models prior to the reference period.

Plain Language Summary

We compare temperature and precipitation over the UK from two different collections (known as 'ensembles') of climate model runs: the EuroCORDEX ensemble, consisting of simulations from many combinations of global- and regional-scale models; and the UKCP18 regional ensemble, which uses a single pair of models, but adjusts the model parameters for each run. Both ensembles perform well in the current climate, but future changes in the UKCP18 ensemble are generally larger by 2050–2079 than those in the EuroCORDEX ensemble. This is largely because the UKCP18 global models warm more quickly in response to the greenhouse gases in the atmosphere, and use slightly higher concentrations of greenhouse gases.

To understand the differences between the two ensembles that cannot be explained by differences in the rate of global warming, we look at changes as the models warm from 1°C to 2°C globally above levels in the early 20th century. This reduces the discrepancy between the ensembles, although some differences remain: the UKCP18 ensemble remains slightly warmer and drier than EuroCORDEX, particularly in summer. We highlight issues that arise when comparing simulations at a given warming level against simulations in a fixed decade, due to uncertainty about how much warming has already occurred.

1 Introduction

Adapting to climate change will be one of the great challenges of the twenty-first century. Knowledge of how future changes will impact a locality is an important prerequisite to planning for them. Many available global climate projections do not provide information at local spatial scales, and the use of regional climate models (RCMs) to dynamically downscale their coarse resolution is an important step to provide locally-relevant information (Jacob et al., 2014; Sørland et al., 2018). Uncertainty about the potential range of future changes is often assessed through the use of ensembles (coherent collections) of simulations (von Trentini et al., 2019; Lehner et al., 2020). Large model ensembles, consisting of simulations using the same climate models but initialised with different atmospheric conditions, aim to sample internal variability within a particular model

(Collins et al., 2006, 2011); whilst perturbed parameter ensembles (PPEs), in which multiple simulations are again produced by a single model but now varying the parameters controlling the representation of physical processes in each realisation, examine the effects of the associated uncertainties. In contrast to this, a multi-model ensemble (MME) in which all simulations use the same pathway of future emissions or atmospheric concentrations, but different combinations of global and regional models, can also sample the uncertainty arising from the formulation of the selected models (von Trentini et al., 2019; Christensen & Kjellström, 2020). Ensembles such as the global Coupled Model Intercomparison Project Phase 5 experiment (Taylor et al., 2012, CMIP5) and, at a regional scale, the CoOrdinated Regional Downscaling EXperiment (Giorgi & Gutowski Jr, 2015, CORDEX) are essential tools to understand the potential range of future climate impacts.

Here we compare projections of temperature and precipitation from two regional climate ensembles at the same spatial scale – the CORDEX MME (Jacob et al., 2014) and the UK Climate Projections 2018 PPE (Murphy et al., 2018) – along with projections produced by the global General Circulation Models (GCMs) used to drive the regional models. We note that the GCMs are not designed or expected to capture detailed orographic or coastal effects, and so can only be fairly assessed at the national scale; however, they are included here so that the contributions and capabilities of the downscaling RCMs can be evaluated. Both the EuroCORDEX and UKCP18 ensembles aim to provide plausible climate projections under the RCP8.5 emissions scenario, but they project quite different of outcomes by the late 21st century. To explore the differences between the two ensembles, we compare projections not only between a reference period and 2050–2079, but also between global warming levels of 1°C and 2°C compared with the early 20th century. We focus on the climate of the United Kingdom, a relatively small region for which decision-making and planning are frequently quite localised in comparison to other, larger countries, and for which the spatial resolution of the GCMs is too coarse to provide the local-scale information required for localised adaptation. However, outputs from both of the regional ensembles are also available for a much wider area encompassing most of Europe, and we anticipate that the considerations around the use of Global Warming Levels (GWLs) with regional climate projections, along with the broader points raised in the Discussion, will have broader relevance to users of those and other regional ensembles.

The ensembles of simulations used in the study are described, along with the necessary preprocessing, in Section 2; the methods used to regrid the data and calculate climatologies are described in Sections 2.2 and 2.3. Section 2.4 describes the approach used to calculate the GWL climatologies and highlights some important caveats to be considered when using GWLs with regional climate model output. In Section 3 we evaluate the representation of historical and future surface temperatures in the ensembles; a similar analysis is carried out for projected changes in seasonal precipitation in Section 4. Section 5 concludes with a discussion of potential benefits and drawbacks of the use of GWLs.

Any plots referred to but not shown in the main text can be found in the supplementary material, or — along with plots of other climate indices — by using the EuroCORDEX-UK Plot Explorer tool at `https://github-pages.ucl.ac.uk/EuroCORDEX-UK-plot-explorer/` (Barnes et al., 2023).

2 Methods

2.1 Datasets

The analysis is focused mainly on a comparison of the regional component of UKCP18, the latest suite of national climate projections for the UK (Murphy et al., 2018), with projections produced by the EuroCORDEX project under the RCP8.5 scenario (Van Vuuren et al., 2011; Jacob et al., 2014). UKCP18 provides a range of different products, the regional component of which is a 12-member PPE that uses HadREM3-GA7-05 to downscale output from the global HadGEM3-GC3.05 model at approximately 60km resolution to a resolution of 0.11° over Europe, equivalent to about 12km resolution over the UK (Murphy et al., 2018). Each numbered ensemble member uses the same perturbations at both 60km and 12km resolutions, with the first ensemble member having no perturbations from the standard model. The ensemble members additionally sample a range of future emissions scenarios consistent with the single RCP8.5 pathway used in the CMIP5 experiments, rather than using the RCP8.5 pathway directly. CO_2 pathways were chosen to represent the range of outcomes indicated by the UKCP18 probabilistic projections (Murphy et al., 2018), with most of the pathways falling above the standard RCP8.5 scenario; in addition, some of the perturbed parameters relate to scalings of anthropogenic aerosol emissions (Sexton et al., 2021; Yamazaki et al., 2021). Henceforth, the ensemble of regional runs will be referred to as UKCP regional, and the global PPE as GC3.05-PPE, in line with UKCP documentation.

At the time of writing, the EuroCORDEX project has produced runs driven by RCP8.5 forcings from six of the coupled ocean-atmosphere models run as part of the Coupled Model Intercomparison Project Phase 5 (CMIP5) experiment (Taylor et al., 2012), using thirteen RCMs (Jacob et al., 2014). However, runs have only been produced for a subset of the possible GCM-RCM pairs, and the EuroCORDEX ensemble used in the present analysis consists of 64 climate simulations, shown in Figure 1. Two of the GCMs provided three independent realisations to the project, but each marked GCM-RCM pair contributes a single run to the 64-member ensemble. The EuroCORDEX models are also run at 0.11° resolution over Europe, with the exact spatial extent varying according to the downscaling RCM. Henceforth, the ensemble of runs used to drive the EuroCORDEX simulations will be referred to as CMIP5-EC.

For each of the ten RCMs listed in Figure 1 the EuroCORDEX ensemble also provides a single evaluation run forced by ERA-Interim reanalysis (Dee et al., 2011) rather than by GCM output: these runs allow the performance of the RCMs to be evaluated in the absence of errors or biases inherited from the driving GCMs. The evaluation periods for which these runs were produced differ between models, with only the period from January 1st 1989 – December 31st 2008 covered by all of the runs. Biases in the model output during this period are evaluated against interpolated daily estimates of historical precipitation and daily maximum and minimum temperature – referred to henceforth as the observations – from the HadUK-Grid dataset (Hollis et al., 2019). Where observations of daily mean temperature are required, the mean of the daily maximum and minimum is used (Perry et al., 2009).

2.2 Regridding onto a common grid

The various ensembles considered in this paper include models run at different spatial resolutions (e.g. the outputs from the RCMs as well as from the GCMs used to drive them) and with different native grids. To facilitate direct comparison across all of the ensembles, each model's outputs are interpolated from the native grid onto a common grid. In this paper, all data are presented on the same 12km grid used in the HadUK-Grid data set and UKCP regional over the UK land surface.

Indices are first computed on each model's native grid, then interpolated to the 12km grid using a conservative area-weighting scheme (Jones, 1999). When regridding the regional model outputs, only grid cells falling within the UK land surface are used: this is to avoid introducing bias by interpolating across the land-sea boundary. When regridding the lower-resolution CMIP5-EC and GC3.05-PPE output however, this approach is not used: removing cells flagged as belonging to the sea surface before regridding the

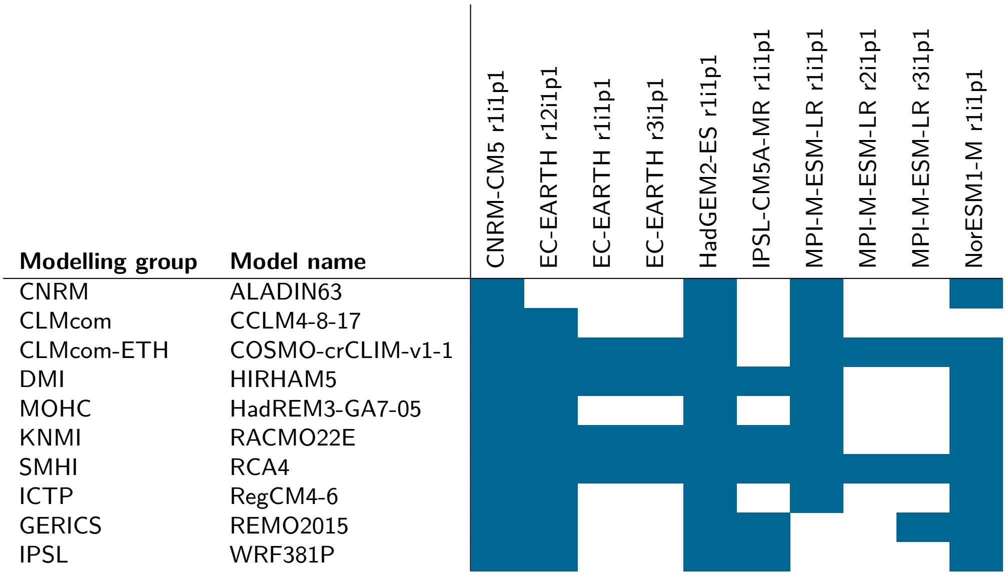

Figure 1. The 64 GCM-RCM pairs included in the EuroCORDEX ensemble.

data would result in an absence of data in large areas of the UK. Instead, the low-resolution data are regridded directly onto the land surface 12km grid, and the effect of any resulting blurring of land and sea surface variables is highlighted when discussing the results below. This choice was made to keep the focus of this paper on the 12km resolution of the regional climate models; if GCM performance was of direct interest then it would be more informative to compare the GCMs to observations on a coarser grid, for example the 60km version of HadUK-Grid (Hollis et al., 2019). However, as noted above, results for GCMs are presented here primarily to illustrate which aspects of the RCM performance are largely inherited from the driving models, and which arise from the downscaling models themselves.

2.3 Calculating climatologies

Model biases are calculated as the difference between the model climatology and the equivalent HadUK-Grid observed climatology during the common evaluation period from January 1st 1989 to December 31st 2008. Changes in temperature-based indices are calculated as the difference between the aggregated value of the index during the future period (December 1st 2049 – November 30th 2079) and the reference period (December 1st 1980 – November 30th 2010). For precipitation indices, biases are presented as relative (percentage) differences with respect to observed precipitation, and changes as relative differences with respect to the reference climatology. UKCP18 users should note that this is not the same reference period as that used in the original UKCP18 analysis, which considered twenty-year periods (Murphy et al., 2018): instead, the present paper focuses on the thirty-year time-slices recommended by the World Meteorological Organisation (WMO, 2017).

2.4 Changes between global warming levels

The GCMs used to drive the EuroCORDEX and UKCP regional ensembles have very different climate sensitivities (Flato et al., 2014; Yamazaki et al., 2021) and, as noted in Section 2.1, the GC3.05-PPE runs also use a variant of the standard RCP8.5 emissions scenario. This translates into rather different rates of warming — illustrated in Table 1, which shows the change in GMST in each GCM between 1900–1950 and 2050–2079 — which might be expected to propagate into systematic differences between the regional outputs, obscuring potentially interesting differences in local responses.

Table 1. GMST change (°C) from 1900-1950 to 2050–2079 in the CMIP5-EC and GC3.05-PPE runs. CMIP5-EC values are taken from the IPCC Interactive Atlas (Iturbide et al., 2021); GC3.05-PPE values were calculated directly from the area-weighted model output.

CMIP5-EC		**GC3.05-PPE**			
GCM	**°C**	***Member***	**°C**	***Member***	**°C**
CNRM-CM5	2.7	*01*	3.9	*09*	4.3
EC-EARTH	2.9	*04*	4.2	*10*	3.7
HadGEM2-ES	3.4	*05*	3.7	*11*	3.9
IPSL-CM5A-MR	3.6	*06*	3.8	*12*	3.2
MPI-ESM-LR	3.1	*07*	3.6	*13*	3.8
NorESM1-M	2.6	*08*	3.3	*15*	3.8

One approach to controlling the variability associated with both the rates of warming exhibited by different models and the choice of emissions scenario is to compare changes in model climatology not at particular time periods but at periods centred on the year in which the change in global mean surface temperature since preindustrial levels exceeds a particular threshold of interest, known as the global warming level (GWL) (James et al., 2017; Hausfather et al., 2022). This approach was adopted in the IPCC's AR6 (Chen et al., 2021), and ensemble means of the CORDEX projections at specified GWLs are available through the IPCC's Interactive Atlas (`https://interactive-atlas.ipcc.ch`). By fixing the GWL in this way, inter-model variation arising from the choice of forcing scenario and from differences between the driving models' global responses to greenhouse gases is reduced: the remaining differences between the runs may therefore be attributed with greater confidence to differences in the local climate response and natural variability (James et al., 2017).

Particular care must be taken when using the GWL approach to evaluate changes in regional model output, although the authors are not aware of any case in the literature where this has previously been highlighted. This is because, while GWLs are typically calculated with respect to a preindustrial reference, regional climate model output is typically only available from the late twentieth century onward. As a result, changes are commonly reported with respect to a reference period beginning no earlier than 1980: for example, the IPCC Interactive Atlas presents changes of climate indices computed from CORDEX regional model output at GWLs of 1.5, 2, 3 and 4°C with respect to three reference periods beginning later than 1980 (1981–2010, 1986–2005, and 1995–2014). However, due partly to the differences in climate sensitivity that GWL selection is intended to mitigate, the driving runs have already warmed by different amounts between the preindustrial and reference periods. Figure 2a, showing the change in GMST between 1900–1950 (used in place of a preindustrial baseline due to unavailability of earlier GC3.05-PPE output) and the reference period of 1981–2010, illustrates this. The observed GMST increase during this period was approximately 0.6°C (calculated from HadCRUT.5.0, Morice et al. (2021)), and as Figure 2a shows, more than half of the CMIP5-EC driving runs have already exceeded this threshold before the start of the regional model output. In contrast, all but one of the GC3.05-PPE runs have warmed by less than 0.5°C prior to the reference period. This systematic difference can largely be attributed to GC3.05-PPE's strong cooling response to increased aerosol concentrations during the second half of the twentieth century (Murphy et al., 2018; Tucker et al., 2021), which may mean a strong warming response to greenhouse gas forcing emerges during model development (Nijsse

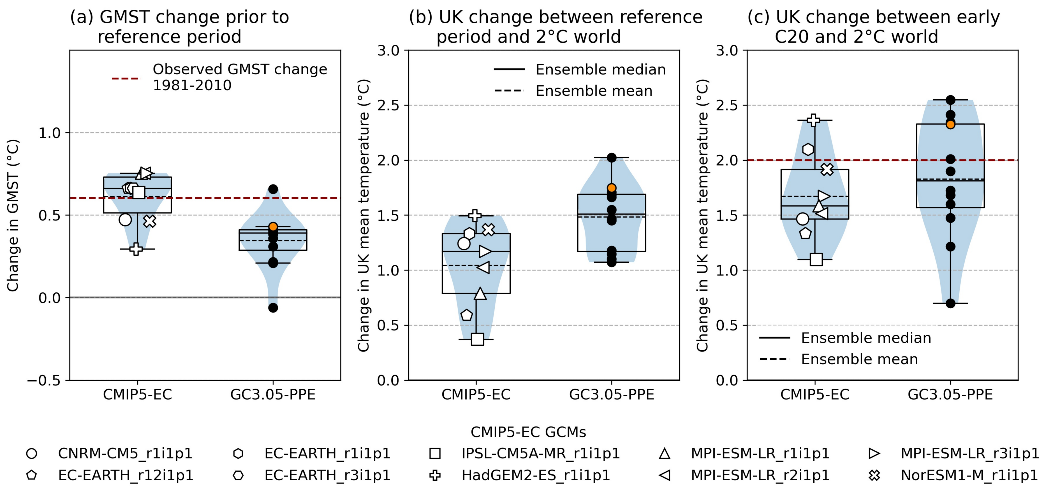

Figure 2. Boxplots of annual temperature changes in the CMIP5-EC and GC3.05-PPE runs: (a) global temperature change from the early 20th century (1900–1950) to 1981–2010; and changes in mean UK land near-surface temperature from (b) 1981–2010 and (c) the early 20th century to the 30-year time period centred on the year in which the driving model exceeded a 2°C increase in GMST with respect to early 20th century climate.

et al., 2020). While the difference is not particularly problematic in the EuroCORDEX ensemble (in part due to the relatively low warming rates of most of the driving models), preliminary analysis suggests that the CMIP6 models simulate an even wider spread of historical changes, ranging from -0.05 to 1.06 degrees: it is therefore very that the range of temperature changes observed in any representative CMIP6-driven CORDEX ensemble prior to the reference period would be somewhat wider than in the current EuroCORDEX and UKCP regional ensembles, although we note that a balanced ensemble design is planned for the next CORDEX phase to sample the range of climate sensitivities in CMIP6 more systematically (Sobolowski et al., 2023).

The potential for confusion caused by comparing a GWL with a fixed reference period is illustrated in Figure 2b, which shows the change in UK mean temperature in each of the driving models between the reference period (1981–2010) and the year in which each model's GMST first exceeded 2°C. Consideration only of the changes between the reference period and a particular GWL in this way fails to take into account the effect of the models' differing warming rates prior to the reference period: the GMST of each model during the reference period is unknown, and as a result, it is not clear how to interpret the changes. Furthermore, because the GC3.05-PPE runs were generally cooler during the reference period than the CMIP5-EC ensemble (Figure 2a), UK temperatures appear to have warmed somewhat more in GC3.05-PPE by the time the models reach 2°C than in CMIP5-EC. In Figure 2c, projected changes are instead evaluated against a reference GWL – here, the early twentieth century. This removes almost all variability due to climate sensitivity and to the choice of forcing scenario, leaving only the models' regional response to a defined period of global warming: once this source of variability is accounted for, the local responses of the two ensembles are in fact fairly similar. This makes interpretation of Figure 2c straightforward: in both ensembles, most of the runs simulate between 1.5 and 2°C of warming over the UK land surface in response to a 2°C change in GMST, with ensemble mean changes of 1.7 and 1.8°C, respectively. A similar approach was used by Arnell et al. (2021), who accept the observed rise of 0.61°C between pre-industrial and 1981–2010 as fixed, then use a projected further increase of 1.39°C relative to 1981–2010 to define the 2°C GWL for each model.

Figure 2 uses the driving (global) models to demonstrate the problems inherent in comparing GWLs with a fixed reference period. For regional model output however, the equivalent of Figure 2c often cannot be produced because, as noted above, regional simulations are typically unavailable for time periods before 1980. If the intention of an analysis really is to characterise the change in regional climate between 1981–2010 and some future GWL, then one simple approach would be to replace the fixed 1981–2010 reference period with a model-dependent reference period defined, for each model, as the time at which the driving GCM reaches a GWL equivalent to that observed in the real world by 1980, and to use this as the basis for comparisons. However, the choice of reference GWL is constrained by the available data: one of the CMIP5-EC models has warmed by 0.8°C prior to the start of the reference period, corresponding to the warming actually observed by around 2006. In the following analysis therefore, we explore the range of changes simulated in each of the regional ensembles between periods in which the driving GCMs reached GWLs of 1°C – approximately the level observed by 2015 – and 2°C respectively. GWL climatologies are calculated by identifying the year in which the GMST of the driving GCM first exceeds the GWL of interest; calculating the climatology of the regional model output for the 30-year periods centred on those years; and computing the change between the two. As above, changes in temperature are presented as absolute changes, while changes in precipitation are presented as relative differences with respect to the amount projected after 1°C of global warming.

3 Simulation of UK temperatures

3.1 Historical biases, 1989–2008

Figure 3a shows maps of the HadUK-Grid mean daily temperature in winter and the mean bias in each ensemble during the evaluation period (1989–2008), with corresponding plots for summer temperatures in Figure 3b. The CMIP5-EC runs are on average around 1°C too cold over much of central England but somewhat too warm at high elevations and, in winter, around much of the coast. This pattern can be attributed to an underlying cold bias in many of the GCMs consistent with that observed over much of western Europe by Vautard et al. (2021), offset by local warm biases due to unresolved topography and blurring between land and sea surface temperatures due to the coarse resolution. A similar spatial pattern is seen in the GC3.05-PPE ensemble mean, although with much reduced biases at higher elevations. The RCMs inherit this cold bias but are able to resolve the features causing local warm biases in the driving models, with the EuroCORDEX ensemble as a whole having a fairly uniform bias of between -1 and -2°C across the UK land surface in winter (-1°C in summer). In the evaluation runs driven by ERA-Interim reanalysis, the magnitude of this bias is reduced in both summer and winter, supporting the suggestion that the error is to some extent inherited from the driving GCMs. The UKCP regional ensemble also inherits a slight cold bias from the 60km driving runs, largely attributed by Murphy et al. (2018) and Tucker et al. (2021) to a strong aerosol forcing, moderated by differences in large-scale circulation patterns; in winter this bias is slightly smaller on average than seen in the EuroCORDEX ensemble, with a fairly uniform mean bias of -0.5°C across much of the UK, increasing to -1.5°C over higher elevations in Scotland.

The boxplots in Figure 4 show the distributions of average UK winter and summer temperatures in each ensemble. Average UK temperatures in the regridded CMIP5-EC ensemble are slightly higher than in the corresponding EuroCORDEX runs, with the differences particularly pronounced in winter (panel a); this is largely due to the warm biases at high elevations and in coastal regions mentioned above, which the regional models are able to resolve. Within the EuroCORDEX ensemble, average summer temper-

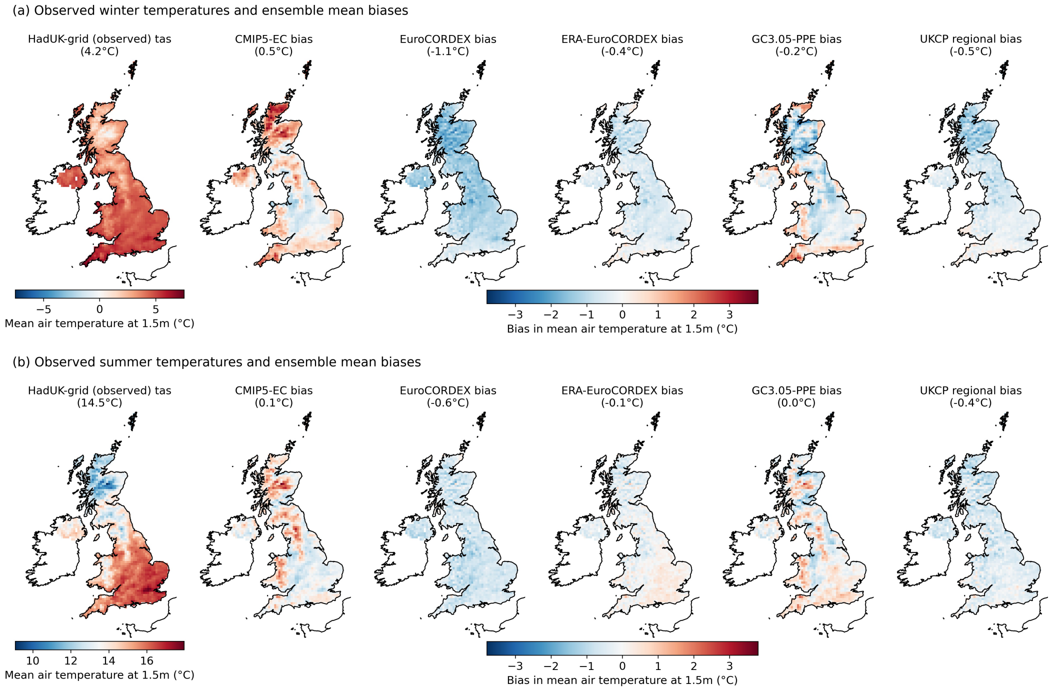

Figure 3. Maps of seasonal averages of HadUK-Grid daily mean temperature (in °C) from 1989 to 2008, and of the mean climatological biases in (a) winter and (b) summer, in each of the ensembles of models. The mean bias over the UK land surface is given in parentheses.

atures also display a degree of clustering by RCM, with the coolest summers simulated by runs downscaled using RACMO22E and RCA4 (coloured yellow and lime green); the same RCM ordering is also seen in the reanalysis-driven ERA-EuroCORDEX ensemble, suggesting that the regional models also contribute systematic differences of their own (Sørland et al., 2018; Vautard et al., 2021). UKCP18 ensemble members display similar biases at both 60km and 12km resolution and, with the exception of the coldest runs in winter, the spread of biases in the UKCP18 ensembles is broadly comparable to that of the EuroCORDEX ensemble. In both the EuroCORDEX and UKCP regional ensembles the biases in mean temperatures are largely due to underestimation of daily maxima arising from large-scale processes driven by the GCMs (Vautard et al., 2021), while daily minima are typically well represented.

3.2 Projected changes in temperature, 2050–2079 relative to 1981–2010

Maps of ensemble mean changes (shown in Figure S1 in the supplementary material, and also available from the accompanying Plot Explorer as detailed in Section 1), indicate a fairly uniform increase in temperature across the whole of the UK, although the UKCP regional ensemble warms somewhat more over higher elevations in winter and in southern England in summer. The distributions of the average changes in seasonal temperature across the UK projected by individual runs are shown in the boxplots in Figure 5. In both winter and summer, EuroCORDEX runs denoted by the same symbol (indicating that they were driven by the same GCM) are closely grouped together, with the average changes in the EuroCORDEX runs generally of similar magnitude to the changes in the driving GCM runs, indicating that the dominant contribution arises from the driving models: the CMIP5-EC and EuroCORDEX ensembles warm by, on average, around 2°C in both winter and summer. This is also the case in the UKCP18 en-

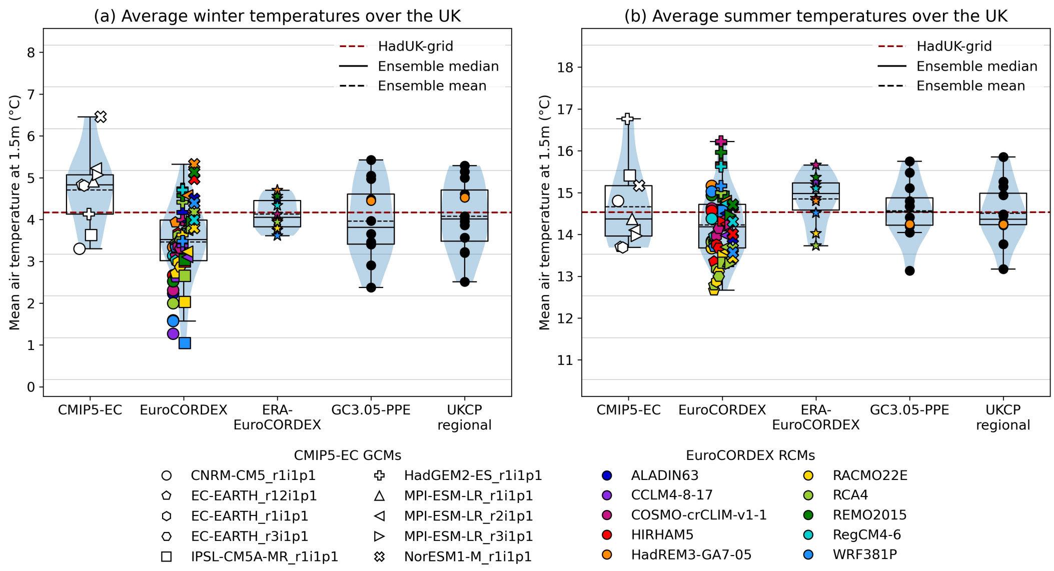

Figure 4. Boxplots showing the distribution of UK-averaged daily mean temperatures in each ensemble during the evaluation period (1989-2008) during (a) winter and (b) summer months. The boxes indicate the central 50% of the distribution; the whiskers of the boxplot extend to values lying 1.5 times the interquartile range beyond the upper and lower quartiles. The shaded region behind each boxplot shows a kernel density estimate of the empirical distribution of the values. Members of the CMIP5-EC and EuroCORDEX ensembles are represented by coloured symbols, with the shape indicating the GCM used to force the run, and the colour indicating the downscaling RCM; points corresponding to the output of a single GCM are jittered horizontally for ease of viewing. The unperturbed UKCP18 ensemble member, corresponding to HadREM3-GA7-05 in the regional ensemble, is shaded orange.

sembles, where each ensemble member warms by the same amount at both 60km and 12km resolutions: in winter, by around 0.6°C more on average than the EuroCORDEX ensemble, and in summer, by around 1.7°C more. Similar differences between the GC3.05-PPE and CMIP5 projections have been discussed by Yamazaki et al. (2021), who attributed them partly to greater climate sensitivity in the UKCP18 members than in most of the CMIP5-EC models, and partly to the fact that the CO_2 pathways sampled by GC3.05-PPE tend to lie above the standard RCP8.5 pathway used to drive the CMIP5 runs, as discussed in Section 2.1. Boé et al. (2020) and Taranu et al. (2023) also note that the absence of time-varying aerosols from most of the EuroCORDEX RCM simulations may also suppress the range of future projections in that ensemble.

Readers may note that GC3.05-PPE is derived from a model descended from HadGEM2-ES (represented by a cross in the CMIP5-EC and EuroCORDEX ensembles), which also projects a strong warming trend. While the similarity between their projected warming levels suggests that these two GCMs share a similar degree of climate sensitivity, the differences between the two models are substantial (Williams et al., 2018; Murphy et al., 2018), and the intermediate variant HadGEM3-GC2 introduced changes that reduced climate sensitivity (Senior et al., 2016), before changes to parameterisation (notably in the aerosol and cloud microphysics schemes) increased the sensitivity of the GC3 generation of models (Bodas-Salcedo et al., 2019). This greater sensitivity appears to be compounded during the summer months by the use of HadREM3-GA7-05 (indicated by orange symbols in the EuroCORDEX ensemble), which produces the warmest run in the

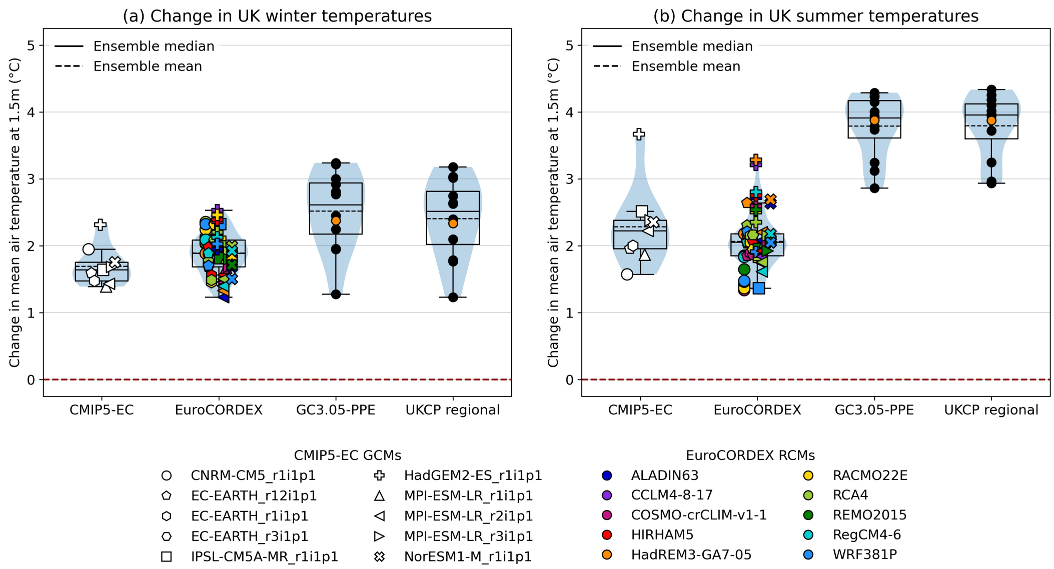

Figure 5. Boxplots showing the average change in daily mean (a) winter and (b) summer temperatures over the UK land surface between the reference period (1981–2010) and the future period (2050–2079). For details of the plot elements see the caption to Figure 4.

EuroCORDEX ensemble for every GCM with which it is paired. This difference is driven by particularly large increases in summer daily maxima, which are typically 1°C higher across the UK than the corresponding increases in summer minima in all UKCP regional and HadREM3-GA7-05 runs (Figure S2), probably largely due to the inclusion of aerosol forcing in the regional model (Boé et al., 2020; Tucker et al., 2021). As noted by Lo et al. (2020) and Keat et al. (2021), UKCP regional also exhibits a particularly strong urban heat island effect, with summer daily minima in London increasing by around 0.2°C more than in the rest of south-eastern England. Maps and boxplots of biases and changes in seasonal temperature maxima and minima are also available through the online Plot Explorer tool.

3.3 Projected changes in temperature, 2°C relative to 1°C global warming

Figure 6 shows boxplots of the seasonal changes in temperature simulated by the two regional ensembles in response to an increase of GMST from 1°C to 2°C in the driving models. During the winter months both the CMIP5-EC/EuroCORDEX and UKCP18 ensembles warm by slightly less than 1°C over the UK (0.8°C and 0.6°C respectively, Figure 6a). The UKCP18 ensembles both have a strongly bimodal distribution, with four of the runs warming very little, and the remaining eight runs warming by 0.7-1°C, roughly in line with the central 50% of the EuroCORDEX distribution. This is a contrast to Figure 5a, where more than half of the UKCP18 runs exceeded the 75th percentile of the EuroCORDEX ensemble. Even after removing variation associated with different global warming rates, EuroCORDEX runs with the same symbol (denoting the same driving model) are still loosely grouped together, reflecting the importance of large-scale processes in determining daily temperatures over the UK (Pope et al., 2022).

The pattern of changes in summer temperatures in Figure 6b bears more resemblance to that seen in Figure 5b: in UKCP18 the UK warms slightly more rapidly than the global mean in summer, with most members simulating increases of 1–1.2°C, while the range of responses simulated by the EuroCORDEX ensemble is rather wider and slightly lower, and broadly similar to the range of responses in winter temperatures. As noted previously, this larger change in summer temperatures in the UKCP18 runs is driven primarily by an increase in daily maxima, with ten of the twelve runs simulating increases of 1.2–1.7°C, while daily minima warm by 0.8–1°C. This difference between the two ensembles can no longer be attributed to a difference in climate sensitivity or in the CO2 pathways sampled, but is still largely determined by the GC3.05-PPE global models.

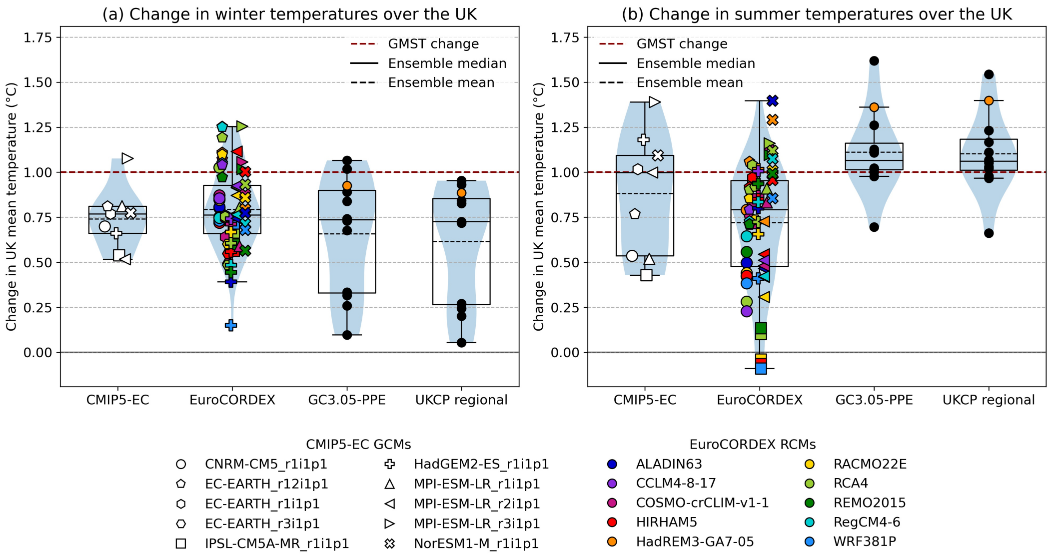

Figure 6. Boxplots of changes in (a) winter and (b) summer UK mean temperature in response to an increase of GMST from 1°C to 2°C. For details of the plot elements see the caption to Figure 4.

4 Simulation of UK precipitation

4.1 Historical biases, 1989–2008

Figure 7 shows maps of the mean winter and summer daily precipitation rates in HadUK-Grid, together with the relative mean biases in each ensemble. Observed precipitation is highest in west-facing areas of high elevation throughout the year, with the heaviest rainfall concentrated in western Scotland. In winter, with the exception of the UKCP regional ensemble, the models tend to overestimate precipitation in the drier lower-lying areas, and to underestimate it in the wetter areas and at higher elevations. Although the biases are most acute in CMIP5-EC and GC3.05-PPE – indicating that this is due to unresolved features – the fact that they persist in the EuroCORDEX ensembles, both those driven by GCM outputs and by reanalyses, suggests that the RCMs also do not fully resolve the driving processes despite better representing the local topography. For the EuroCORDEX and UKCP regional ensembles, separate analyses of precipitation frequency and wet-day intensity (Figures S3 and S4, also available through the online Plot Explorer tool: a wet day is one with at least 1mm precipitation) reveals that the wet biases in lower-lying, drier regions tend to correspond to simulation of too many wet days, with dry biases at higher elevations the result of lower rates of wet-day precipitation, although the signs are more mixed in winter: similar results were presented in Kendon et al. (2021). The UKCP regional ensemble does not suffer from this dry elevation bias, and

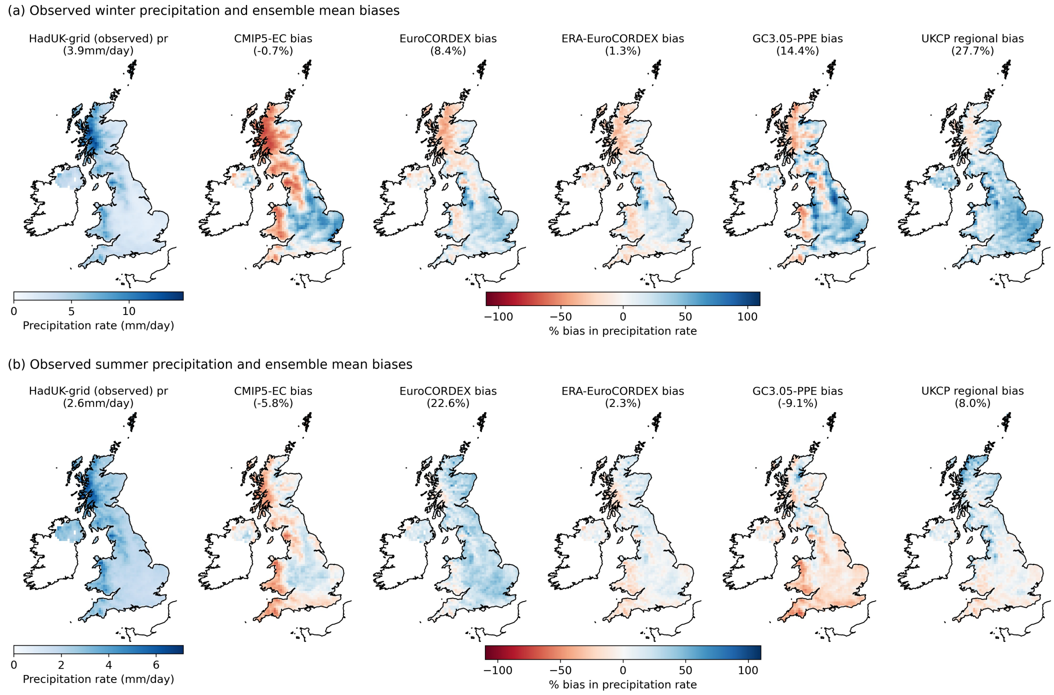

Figure 7. Maps of HadUK-Grid daily mean precipitation rates from 1989 to 2008, and of the relative biases in the means of each ensemble, during (a) winter and (b) summer.

is too wet across much of the UK in winter, with the exception of the far northwest where observed precipitation is highest.

In summer the CMIP5-EC and GC3.05-PPE ensembles are also too dry at higher elevations – again, predominantly the result of unresolved topography – while the main EuroCORDEX ensemble is, on average, too wet across most of the UK; the average bias in the reanalysis-driven runs is similar to that seen in winter, but the magnitude is rather smaller. Aside from the elevation-induced biases, the GC3.05-PPE runs are slightly too wet in rain shadows, but too dry across much of England: the UKCP regional runs are slightly too dry in England but again slightly too wet across much of Scotland. Again, inspection of plots available on the Explorer tool indicates that wet-day precipitation rates are typically underestimated across the UK, with too many wet days simulated on average in all ensembles except for the GC3.05-PPE ensemble.

The distributions of the relative biases in UK mean summer and winter precipitation within each ensemble are shown in Figure 8. The CMIP5-EC runs underestimate precipitation over the UK in both summer and winter, with most of the models underestimating winter precipitation by more than 20%. This pattern of biases is not directly reflected in the EuroCORDEX biases, which vary widely, ranging from -40% to +50% in winter and -30% to +70% in summer. Biases in the EuroCORDEX runs tend to be more closely grouped by colour (denoting the RCM) than by shape (denoting the GCM), and to be similar to those of the corresponding reanalysis-driven ERA-EuroCORDEX runs; this suggests that biases in precipitation are determined to a greater extent by the choice of RCM than the choice of GCM, which in turn implies that the differences between the properties of RCM and GCM output are due to more than just the differing spatial resolutions of the models. Biases in average UK precipitation are more closely correlated with biases in the wet-day precipitation rate in winter, and with biases in the number of wet days in summer.

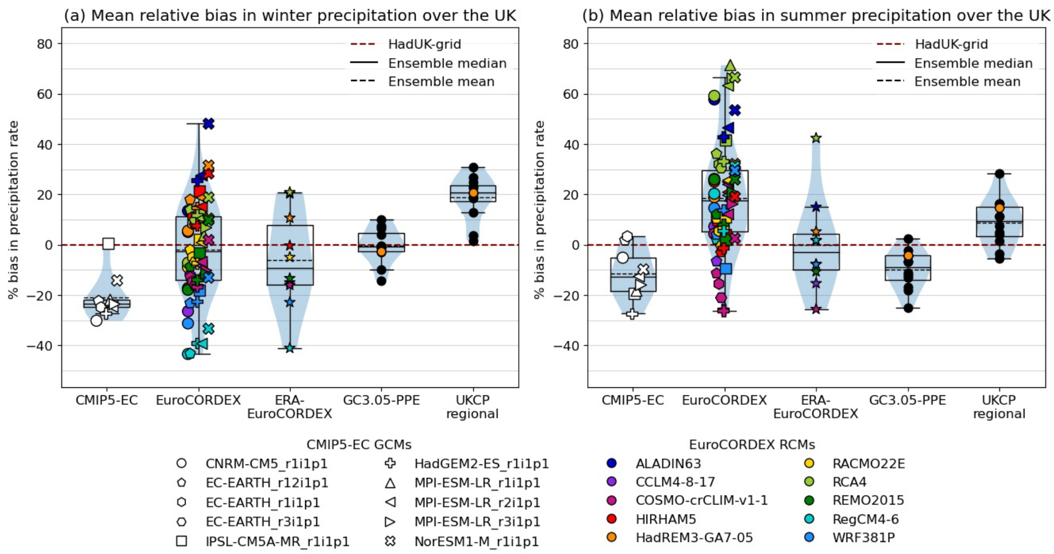

Figure 8. Boxplots showing the distribution of relative biases in UK-averaged winter and summer precipitation in each ensemble during the evaluation period (1989-2008). For details of the plot elements see the caption to Figure 4.

Both of the UKCP18 ensembles have a much smaller range of relative biases than the EuroCORDEX ensemble, with the 12km runs being somewhat wetter overall than their 60km driving runs, largely due to compensation of local wet and dry biases in the lower-resolution runs. This difference is primarily due to the ability of the regional models to represent orographic processes at finer resolution, and so to reduce the dry bias at higher elevations, although it is that the inclusion of time-varying aerosols in the UKCP regional runs also plays a part (Boé et al., 2020; Tucker et al., 2021). In both summer and winter the relative bias in the mean of the UKCP regional ensemble is similar to the biases in EuroCORDEX runs using HadREM3-GA7-05, the model from which the UKCP regional PPE was constructed, further reinforcing the role of the choice of RCM in determining biases in precipitation.

4.2 Projected changes in precipitation, 2050–2079 relative to 1981–2010

Maps of the relative changes in the ensemble means of precipitation between 1981–2010 and 2050–2079 (Figure S5, also available from the accompanying Plot Explorer tool) show an overall increase in winter and a decrease in summer across nearly the entire territory: in winter, the CMIP5-EC and EuroCORDEX ensembles the mean increases are around 10% across much of the UK, while the two UKCP18 ensembles exhibit slightly larger increases on average and a slight gradient, with the 60km ensemble simulating 15-25% more precipitation in south-west England and little or no change in the north-east and Scotland. This pattern is also apparent in the 12km ensemble mean, with the increases slightly damped at higher elevations. In summer, means of the the CMIP5-EC and EuroCORDEX ensembles project around 10% less precipitation across the UK on average, and in the UKCP18 ensembles, an average of 25% less. The UKCP18 ensembles again display a pronounced northeast-southwest gradient, with as much as 45% less precipitation in the southwest of England; in the 12km ensemble, there is additional drying on western-facing elevations and slightly less in rain shadows.

The boxplots in Figure 9 show the distribution of the percentage changes in UK-averaged winter and summer precipitation within each ensemble. The trend of increasing mean winter precipitation is fairly consistent across the ensembles: all of the CMIP5-EC runs and most of the EuroCORDEX runs project increases of 5-15% in winter precipitation (Figure 9a), although a handful of runs from HIRHAM5 and HadREM3-GA7-05 simulate less precipitation than during the reference period; all but one of the UKCP18 runs simulate increases of 5-20% at both resolutions. This change is driven primarily by an increase in the wet-day precipitation rate, with very little change in the average number of wet days simulated on average across the UK (see also Kendon et al. (2021), who investigate the issue in more detail for the UKCP local ensemble). Within the EuroCORDEX ensemble, points are loosely grouped by shape, indicating that the GCMs are dominant in determining the change in winter precipitation: however, within these groups the points are also ordered by colour, suggesting that the choice of RCM also plays a fairly significant part. Again, it is interesting to note that the unperturbed member in the UKCP regional ensemble – produced by the same parametrisation of HadREM3-GA7-05 as the EuroCORDEX runs, and highlighted in orange in the plots – simulates one of the smallest increases in winter precipitation in that ensemble, suggesting that this may be a characteristic of that particular RCM.

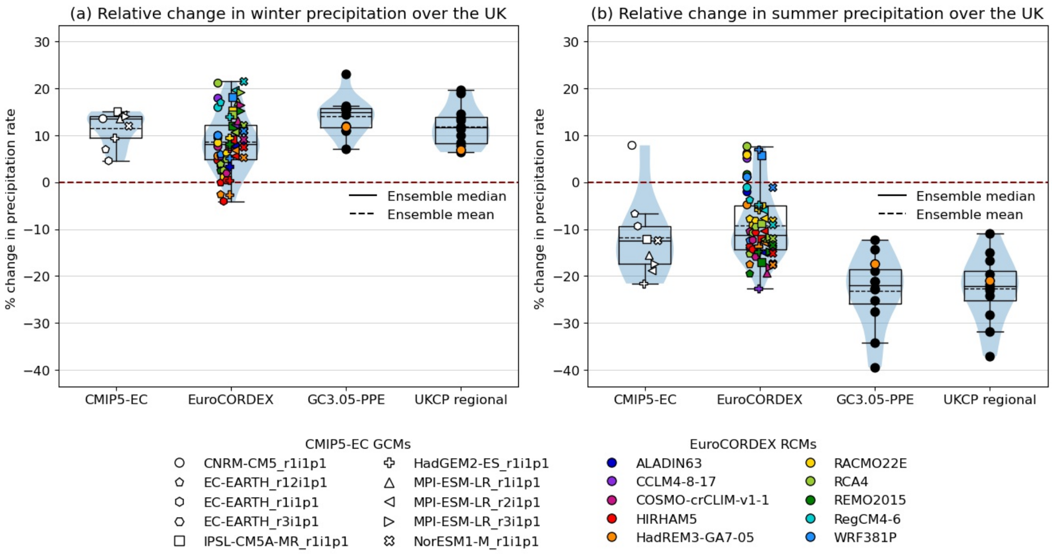

Figure 9. Boxplots showing the relative changes in accumulated UK precipitation in each ensemble during (a) the winter and (b) the summer months between the reference period (1981–2010) and the future period (2050–2079). For details of the plot elements see the caption to Figure 4.

The distributions of changes in mean summer precipitation in the CMIP5-EC and EuroCORDEX ensembles (Figure 9b) are fairly skewed, with most of the GCMs simulating 5-20% less precipitation but with one outlying model – CNRM-CM5 – simulating 7.5% more precipitation across the UK than in the reference period: six of the eight EuroCORDEX runs that simulate an increase in summer precipitation across the UK are driven by this GCM, which is the only one in the ensemble to simulate an increase in the number of wet summer days, suggesting that this may be the result of changes in large-scale circulation patterns. The largest reduction in summer precipitation in the CMIP5-EC ensemble is produced by HadGEM2-ES, the model that also simulated the largest increase in summer temperatures (Figure 5b). However, this tendency is not inherited directly by the runs driven by that model, which produce a wide spread of changes in precipitation, including both the largest decrease and the second largest increase in the EuroCORDEX ensemble: as already noted, the choice of RCM also contributes significantly to the differences between individual runs. All of the UKCP18 runs project reductions of at least 10% in summer precipitation across the UK – a stronger drying trend than the EuroCORDEX ensemble mean or median – with an average reduction of 22.5%. Changes in average precipitation during the summer months are driven by a reduction in the number of wet days simulated, with the effect slightly mitigated by small increases in the wet-day precipitation rate in the CMIP5-EC and EuroCORDEX ensembles, but compounded in several of the UKCP18 runs by small decreases in the rate of wet-day precipitation. This is again consistent with the findings of Kendon et al. (2021) and of Pope et al. (2022), who note a projected increase in occurrences of large-scale circulation patterns associated with dry, settled weather over the UK during the summer months in GC3.05-PPE.

4.3 Projected change in precipitation, 2° relative to 1° global warming

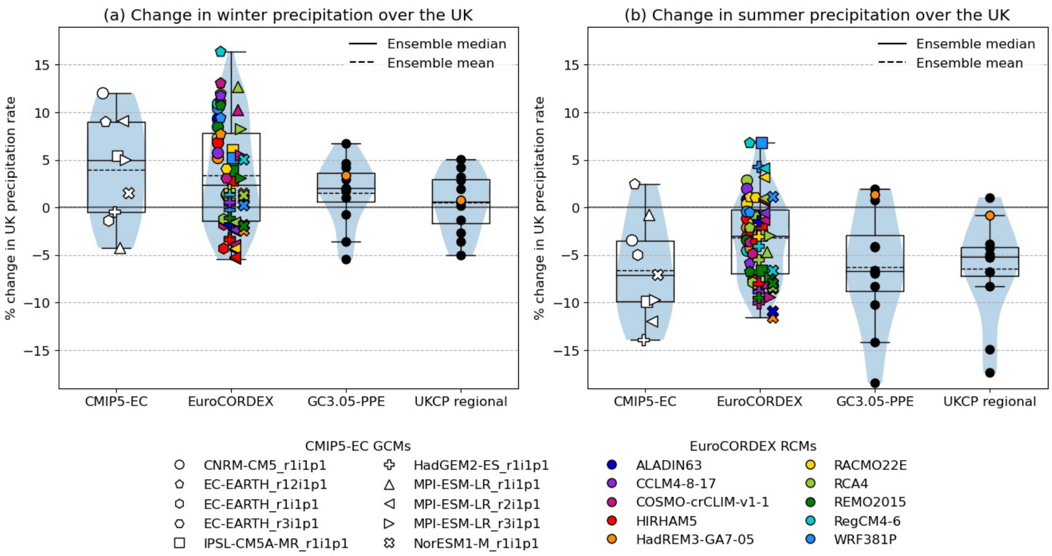

Figure 10. Boxplots of changes in (a) winter and (b) summer UK mean precipitation in response to an increase of GMST from 1°C to 2°C. For details of the plot elements see the caption to Figure 4.

The relative changes in mean UK winter and summer precipitation in response to an increase in GMST from 1°C to 2°C are shown in Figure 10. The UKCP18 runs simulate very little change in winter precipitation on average (panel a), with individual runs projecting between ±5%; more than 25% of the EuroCORDEX members also project a reduction of up to 5%, with the remainder projecting increases of up to 16%, slightly lower than the increases seen in Figure 9a. Overall, both ensembles simulate an increase of around 2.5% in wet-day precipitation rates in response to 1°C of continued warming; the EuroCORDEX runs simulate 1-2% more wet days on average, while the UKCP18 runs simulate 2.5% fewer, leading to very little net change in precipitation.

During the summer months the trends are again similar to those seen in Figure 9b, with EuroCORDEX runs projecting changes from -12% to +8%, and the UKCP18 ensemble slightly more intense drying (Figure 10b). However, this difference between the two ensembles is largely driven by just two of the UKCP18 runs, with the remainder spanning the central 70% of the EuroCORDEX ensemble. As noted previously, these changes are largely driven by a reduction in the number of wet days simulated in both ensembles (by about 5%), with the UKCP18 runs also simulating about 2% less precipitation on wet days: plots illustrating these changes can be found in Figures S6 and S7 or using the aforementioned Plot Explorer tool. Again we find that, although variability attributable to differing warming rates has been removed, runs driven by the same GCM (denoted by the same symbol) still tend to be grouped together, indicating that the choice of GCM still determines the simulated climate to a large extent.

5 Discussion and conclusions

Sections 3 and 4 present an analysis of biases and changes in summer and winter temperatures and precipitation over the UK: while the results presented are specific to the local climate, they offer a useful illustration of the insights that can be gained by considering changes over GWLs alongside those over fixed time periods. We anticipate that this analysis could be used as a template for regional ensemble comparisons more widely, providing a framework by which the effect of regional responses to global warming might be assessed alongside projected changes at a given time period, in order to disentangle the drivers.

Both the EuroCORDEX and UKCP18 regional model ensembles were found to exhibit a similar range of biases in temperature in both summer and winter; a persistent cold bias is inherited by all runs from the driving models but the regional models reduce this tendency somewhat, and the resulting temperature biases are small on average. The UKCP18 runs tend to be too wet on average in winter, while most of the EuroCORDEX runs are too wet in summer, with wet biases generally associated with the simulation of too many wet days.

When considering changes in local climate over time, the CMIP5-EC/EuroCORDEX and GC3.05-PPE/UKCP regional ensembles generally agree on the sign of the changes in average temperatures and precipitation over the UK; however, a stronger signal is observed in the UKCP18 runs at both 60km and 12km resolutions, which project much larger temperature increases and larger drying (wettening) effects in the summer (winter) months. This is, in part, due to the fact that the CMIP5-EC runs used to drive the EuroCORDEX simulations do not include the warmest and driest of the CMIP5 projections (Boé et al., 2020; Coppola et al., 2021), while GC3.05-PPE is derived from a model known to exhibit a high rate of warming in response to greenhouse gas emissions (Murphy et al., 2018; Andrews et al., 2019), as illustrated in Table 1. This greater sensitivity is compounded by the use of perturbed CO_2 pathways to force the runs, which resulted in a higher effective forcing than the standard RCP8.5 scenario in the majority of ensemble members (Sexton et al., 2021; Yamazaki et al., 2021).

The effect of these differences is substantial: between 1981–2010 and 2049–2079, the GC3.05-PPE runs warm by 3-4°C globally; of the ten EuroCORDEX driving runs, HadGEM2-ES and IPSL-CM5A-MR warm by around 3°C, while the remainder warm by around 2.25°C in the same period. Taking into account the number of replicates of each GCM in the EuroCORDEX ensemble, the average global warming across the UKCP regional ensemble will be around 1°C more than the corresponding average for EuroCORDEX during this time: given that changes in many key climate indices have been found to increase monotonically with GMST change (James & Washington, 2013; Seneviratne & Hauser, 2020), the UKCP regional ensemble should be expected to display a correspondingly stronger change signal.

This divergence between the two ensembles poses a problem for anyone wishing to use these climate projections to support effective planning and decision making: how should the two sets of projections be interpreted? The results presented here indicate that there is no direct relationship between the biases exhibited during the evaluation period and future rates of warming, so simple bias correction methods are unlikely to be able to reconcile the two ensembles. By comparing model outputs at fixed warming levels as in Figures 6 and 10, rather than at fixed time periods as in Figures 5 and 9, differences between the EuroCORDEX and UKCP regional ensembles attributable to the varying rates of GMST change in the driving models have largely been removed. As a result the two ensembles, when taken together, present a more coherent picture of plausible local changes in response to global warming, with the UKCP regional ensemble exploring the warmer, drier scenarios that are known to be absent from the EuroCORDEX ensemble (Boé et al., 2020): this complementary information may be important in the context of climate change in western Europe and the UK, where models have been found to underrepresent observed trends in warming on the warmest summer days (Vautard et al., 2023).

Given the current focus on the adaptation to a world 1.5°C or more warmer than the preindustrial climate, this GWL-based analysis has potential applications in separating analysis of the local and regional changes that are to be expected at a given level of global warming from consideration of rates of GMST change. The GWL approach can help to answer the question of why the two ensembles indicate different climate futures, but could also be used to investigate broader questions around adaptation: for example, to what extent are local responses to global warming dependent on the emissions scenario used, the climate sensitivity of the driving models, or the absolute level of warming reached?

Further work is also required to evaluate the sensitivity of the GWL approach to the time periods compared. Time slices spanning a fixed number of years either side of a given threshold exceedance will contain different ranges of GMSTs depending on the climate sensitivity of the driving models, which may introduce biases, particularly in any indices measuring extrema or variability. Alternative approaches might be to select a symmetric GMST interval centred on the year of interest; or more sophisticated approaches based on detrended residuals during the chosen time period, following an approach similar to that used by Sexton et al. (2012) in a slightly different context. Furthermore, although the GWL approach reduces some of the discrepancies between the ensembles, it does not fully reconcile them in all respects (see Figures 6 and 10, for example). Some of the reasons for this are outlined above, but these results nonetheless serve as a note of caution that the approach cannot be regarded as a universal panacea, and users should assess the advantages and disadvantages of the approach relative to other frameworks for addressing variability and biases within ensembles of climate projections.

In contexts where timescales are important, the information provided by analysis of changes between GWLs may be less directly relevant. Similarly, for indices of quantities that are less directly dependent on global temperature change – for example, some indices of precipitation, which may be more sensitive to changes in atmospheric circulation and composition than those determined by temperature – the GWL approach may be less effective in reducing inter-model differences: since different models reach the same GWL at different CO_2 levels, they do so under potentially quite different atmospheric compositions, although some studies have found a monotonic or even linear relationship between regional changes and increasing GMST (James & Washington, 2013; Seneviratne & Hauser, 2020; Arnell et al., 2021). Whether the GWL approach is appropriate or not in a given application, there is still useful information to be gained by comparing the outputs of more than one ensemble of models.

One perspective is that ensembles, like the UKCP regional ensemble, with higher warming rates explore the upper tails of the distribution of plausible outcomes, providing a set of storylines of low-likelihood but high-impact futures (Zappa & Shepherd, 2017) for use in risk-averse decision making. However, neither the EuroCORDEX ensemble nor the combined EuroCORDEX-UKCP regional ensemble systematically samples a range of climate sensitivities, so neither should be interpreted as representative of the possible distribution of future scenarios, although the two ensembles taken together are arguably more representative than either one in isolation. To gain a fuller understanding of the uncertainty about projected changes, it may be instructive to place the regional model output within the context of the UKCP18 probabilistic models, which are designed to more fully reflect the spread of potential future outcomes, or the full UKCP18 global ensemble, which includes not only the PPE but also a subset of thirteen CMIP5 models chosen to reflect a wider range of plausible futures (Murphy et al., 2018); both of these products provide global data, although projections are available for fewer climate variables and at coarser spatial and temporal resolution than the regional model output, and may therefore not provide sufficient detail for some applications. Recent work has shown that observational constraints accounting for the rate of warming in recent decades can resolve much of the difference between the rates in CMIP5 and CMIP6 (Brunner et al., 2020; Ribes et al., 2021), suggesting that similar approaches might be applied to resolve the differences between CMIP5 and GC3.05-PPE, although the method has not been applied to maps of the outputs from regional climate models. This problem of how to interpret and extract relevant information from ensembles that include models with a wide range of climate sensitivities is to become increasingly important, given the known preponderance of high-sensitivity models in the CMIP6 ensemble (Zelinka et al., 2020).

Open Research

All plots and data used in this analysis can be downloaded from the EuroCORDEX-UK plot explorer tool at `https://github-pages.ucl.ac.uk/EuroCORDEX-UK-plot-explorer/`, along with plots of other climate indices (Barnes et al., 2023).

CMIP5 and EuroCORDEX climate simulations can be obtained from the Earth System Grid Federation portals (e.g., `https://esg-dn1.nsc.liu.se/search/cordex/`, `https://esgf-node.llnl.gov/search/cmip5/`).

All UKCP18 climate simulations can be downloaded from the Centre for Environmental Data Analysis at `https://catalogue.ceda.ac.uk` (Met Office Hadley Centre, 2018), along with EuroCORDEX simulations regridded to the 12km OSGB grid used in this report (Barnes, 2023).

Acknowledgments

This research was funded under the UK Climate Resilience Programme, which is supported by the UKRI Strategic Priorities Fund. The programme is co-delivered by the Met Office and NERC on behalf of UKRI partners AHRC, EPSRC, and ESRC.

The authors acknowledge the World Climate Research Programme's Working Group on Regional Climate, and the Working Group on Coupled Modelling, former coordinating body of CORDEX and responsible panel for CMIP5. We also thank the climate modelling groups (listed in Figure 1 of this paper) for producing and making available their model output. We also acknowledge the Earth SystemGrid Federation infrastructure, an international effort led by the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison, the European Network for Earth System Modelling and other partners in the Global Organisation for Earth System Science Portals (GO-ESSP).

The authors would also like to thank the two anonymous reviewers for their constructive feedback.

References

Andrews, T., Andrews, M. B., Bodas-Salcedo, A., Jones, G. S., Kuhlbrodt, T., Manners, J., ... others (2019). Forcings, feedbacks, and climate sensitivity in HadGEM3-GC3.1 and UKESM1. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4377–4394.

Arnell, N. W., Freeman, A., Kay, A. L., Rudd, A. C., & Lowe, J. A. (2021). Indicators of climate risk in the UK at different levels of warming. *Environmental Research Communications*, *3*(9), 095005.

Barnes, C. (2023). *EuroCORDEX-UK: Regional climate projections for the UK domain at 12 km resolution for 1980-2080* [dataset]. NERC EDS Centre for Environmental Data Analysis. Retrieved from `https://catalogue.ceda.ac.uk/uuid/b109bd69e1af425aa0f661b01c40dc51` doi: 10.5285/b109bd69e1af425aa0f661b01c40dc51

Barnes, C., Alegre, R., Ho-Lyn, A., Quinn, J., Chandler, R., & Brierley, C. (2023, August). *EuroCORDEX-UK plot explorer.* Retrieved from `https://github-pages.ucl.ac.uk/EuroCORDEX-UK-plot-explorer/`

Bodas-Salcedo, A., Mulcahy, J., Andrews, T., Williams, K., Ringer, M., Field, P., & Elsaesser, G. (2019). Strong dependence of atmospheric feedbacks on mixed-phase microphysics and aerosol-cloud interactions in HadGEM3. *Journal of Advances in Modeling Earth Systems*, *11*(6), 1735–1758.

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. *Climate Dynamics*, *54*(5), 2981–3002.

Brunner, L., Pendergrass, A. G., Lehner, F., Merrifield, A. L., Lorenz, R., & Knutti, R. (2020). Reduced global warming from CMIP6 projections when weighting models by performance and independence. *Earth System Dynamics*, *11*(4), 995–1012.

Chen, D., Rojas, M., Samset, B., Cobb, K., Diongue Niang, A., Edwards, P., ... Tréguier, A.-M. (2021). Framing, Context, and Methods [Book Section]. In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (p. 147–286). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. doi: 10.1017/9781009157896.003

Christensen, O. B., & Kjellström, E. (2020). Partitioning uncertainty components of mean climate and climate change in a large ensemble of European regional climate model projections. *Climate Dynamics*, 1–16.

Collins, M., Booth, B. B., Bhaskaran, B., Harris, G. R., Murphy, J. M., Sexton, D. M., & Webb, M. J. (2011). Climate model errors, feedbacks and forcings: a comparison of perturbed physics and multi-model ensembles. *Climate Dynamics*, *36*(9), 1737–1766.

Collins, M., Booth, B. B., Harris, G. R., Murphy, J. M., Sexton, D. M., & Webb, M. J. (2006). Towards quantifying uncertainty in transient climate change. *Climate dynamics*, *27*(2), 127–147.

Coppola, E., Nogherotto, R., Ciarlo', J. M., Giorgi, F., van Meijgaard, E., Kadygrov, N., ... others (2021). Assessment of the European climate projections as simulated by the large EURO-CORDEX regional and global climate model ensemble. *Journal of Geophysical Research: Atmospheres*, *126*(4), e2019JD032356.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., ... others (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., ... others (2014). Evaluation of climate models. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 741–866). Cambridge University Press.

Giorgi, F., & Gutowski Jr, W. J. (2015). Regional dynamical downscaling and the CORDEX initiative. *Annual review of environment and resources*, *40*, 467–490.

Hausfather, Z., Marvel, K., Schmidt, G. A., Nielsen-Gammon, J. W., & Zelinka,

M. (2022). *Climate simulations: Recognize the 'hot model' problem.* Nature Publishing Group.

Hollis, D., McCarthy, M., Kendon, M., Legg, T., & Simpson, I. (2019). HadUK-Grid – a new UK dataset of gridded climate observations. *Geoscience Data Journal*, *6*(2), 151–159.

Iturbide, M., Fernández, J., Gutiérrez, J., Bedia, J., Cimadevilla, E., Díez-Sierra, J., … others (2021). *Repository supporting the implementation of FAIR principles in the IPCC-WG1 Atlas.* Zenodo. Retrieved from https://github.com/IPCC-WG1/ATLAS/tree/main/datasets-aggregated-regionally/data/CMIP5/CMIP5_tas_landsea doi: 10.5281/zenodo.3691645

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., … others (2014). EURO-CORDEX: new high-resolution climate change projections for European impact research. *Regional environmental change*, *14*(2), 563–578.

James, R., & Washington, R. (2013). Changes in African temperature and precipitation associated with degrees of global warming. *Climatic change*, *117*(4), 859–872.

James, R., Washington, R., Schleussner, C.-F., Rogelj, J., & Conway, D. (2017). Characterizing half-a-degree difference: a review of methods for identifying regional climate responses to global warming targets. *Wiley Interdisciplinary Reviews: Climate Change*, *8*(2), e457.

Jones, P. W. (1999). First-and second-order conservative remapping schemes for grids in spherical coordinates. *Monthly Weather Review*, *127*(9), 2204–2210.

Keat, W. J., Kendon, E. J., & Bohnenstengel, S. I. (2021). Climate change over UK cities: the urban influence on extreme temperatures in the UK climate projections. *Climate Dynamics*, *57*(11-12), 3583–3597.

Kendon, E., Short, C., Pope, J., Chan, S., Wilkinson, J., Tucker, S., … Harris, G. (2021). *Update to UKCP Local (2.2km) projections* (Tech. Rep.). Met Office, Exeter. Retrieved from https://www.metoffice.gov.uk/research/approach/collaboration/ukcp/science/science-reports

Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E. M., Brunner, L., … Hawkins, E. (2020). Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. *Earth System Dynamics*, *11*(2), 491–508.

Lo, Y. E., Mitchell, D. M., Bohnenstengel, S. I., Collins, M., Hawkins, E., Hegerl, G. C., … Stott, P. A. (2020). UK climate projections: summer daytime and nighttime Urban Heat Island changes in England's major cities. *Journal of Climate*, *33*(20), 9015–9030.

Met Office Hadley Centre. (2018). *UKCP18 Regional Climate Model Projections for the UK* [dataset]. NERC EDS Centre for Environmental Data Analysis, 2023. Retrieved from http://catalogue.ceda.ac.uk/uuid/b4d24b3df3754b9d9028447eb3cd89c6

Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J., Hogan, E., Killick, R., … Simpson, I. (2021). An updated assessment of near-surface temperature change from 1850: the HadCRUT5 data set. *Journal of Geophysical Research: Atmospheres*, *126*(3), e2019JD032361.

Murphy, J., Harris, G., Sexton, D., Kendon, E., Bett, P., Clark, R., … others (2018). UKCP18 land projections: science report. *Met Office*.

Nijsse, F. J., Cox, P. M., & Williamson, M. S. (2020). Emergent constraints on transient climate response (TCR) and equilibrium climate sensitivity (ECS) from historical warming in CMIP5 and CMIP6 models. *Earth System Dynamics*, *11*(3), 737–750.

Perry, M., Hollis, D., & Elms, M. (2009). *The generation of daily gridded datasets of temperature and rainfall for the UK* (Tech. Rep.). National Climate Information Centre, Met Office, Exeter.

Pope, J. O., Brown, K., Fung, F., Hanlon, H. M., Neal, R., Palin, E. J., & Reid,

A. (2022). Investigation of future climate change over the British Isles using weather patterns. *Climate Dynamics*, *58*(9), 2405–2419.

Ribes, A., Qasmi, S., & Gillett, N. P. (2021). Making climate projections conditional on historical observations. *Science Advances*, *7*(4), eabc0671.

Seneviratne, S. I., & Hauser, M. (2020). Regional climate sensitivity of climate extremes in CMIP6 versus CMIP5 multimodel ensembles. *Earth's future*, *8*(9), e2019EF001474.

Senior, C. A., Andrews, T., Burton, C., Chadwick, R., Copsey, D., Graham, T., ... others (2016). Idealized climate change simulations with a high-resolution physical model: HadGEM3-GC2. *Journal of Advances in Modeling Earth Systems*, *8*(2), 813–830.

Sexton, D. M., McSweeney, C. F., Rostron, J. W., Yamazaki, K., Booth, B. B., Murphy, J. M., ... Karmalkar, A. V. (2021). A perturbed parameter ensemble of HadGEM3-GC3.05 coupled model projections: Part 1: selecting the parameter combinations. *Climate Dynamics*, *56*(11), 3395–3436.

Sexton, D. M., Murphy, J. M., Collins, M., & Webb, M. J. (2012). Multivariate probabilistic projections using imperfect climate models part I: outline of methodology. *Climate dynamics*, *38*(11), 2513–2542.

Sobolowski, S., Somot, S., Fernandez, J., Evin, G., Maraun, D., Kotlarski, S., ... Brands, S. (2023, February). *EURO-CORDEX CMIP6 GCM selection & ensemble design: Best practices and recommendations.* Zenodo. Retrieved from `https://doi.org/10.5281/zenodo.7673400` doi: 10.5281/zenodo.7673400

Sørland, S. L., Schär, C., Lüthi, D., & Kjellström, E. (2018). Bias patterns and climate change signals in GCM-RCM model chains. *Environmental Research Letters*, *13*(7), 074017.

Taranu, I. S., Somot, S., Alias, A., Boé, J., & Delire, C. (2023). Mechanisms behind large-scale inconsistencies between regional and global climate model-based projections over Europe. *Climate Dynamics*, *60*(11-12), 3813–3838.

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498.

Tucker, S. O., Kendon, E. J., Bellouin, N., Buonomo, E., Johnson, B., & Murphy, J. M. (2021). Evaluation of a new 12km regional perturbed parameter ensemble over Europe. *Climate Dynamics*, 1–25.

Van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., ... others (2011). The representative concentration pathways: an overview. *Climatic change*, *109*(1), 5–31.

Vautard, R., Cattiaux, J., Happé, T., Singh, J., Bonnet, R., Cassou, C., ... others (2023). Heat extremes in Western Europe increasing faster than simulated due to atmospheric circulation trends. *Nat Commun*, *14*. doi: https://doi.org/10.1038/s41467-023-42143-3

Vautard, R., Kadygrov, N., Iles, C., Boberg, F., Buonomo, E., Bülow, K., ... Wulfmeyer, V. (2021). Evaluation of the Large EURO-CORDEX Regional Climate Model Ensemble. *Journal of Geophysical Research: Atmospheres*, *126*(17), e2019JD032344. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JD032344` doi: https://doi.org/10.1029/2019JD032344

von Trentini, F., Leduc, M., & Ludwig, R. (2019). Assessing natural variability in RCM signals: comparison of a multi model EURO-CORDEX ensemble with a 50-member single model large ensemble. *Climate Dynamics*, *53*(3), 1963–1979.

Williams, K., Copsey, D., Blockley, E., Bodas-Salcedo, A., Calvert, D., Comer, R., ... others (2018). The Met Office Global Coupled Model 3.0 and 3.1 (GC3.0 and GC3.1) configurations. *Journal of Advances in Modeling Earth Systems*, *10*(2), 357–380.

WMO. (2017). *WMO guidelines on the calculation of climate normals.* World Meteorological Organization Geneva, Switzerland.

Yamazaki, K., Sexton, D. M., Rostron, J. W., McSweeney, C. F., Murphy, J. M., & Harris, G. R. (2021). A perturbed parameter ensemble of HadGEM3-GC3.05 coupled model projections: Part 2: global performance and future changes. *Climate Dynamics*, *56*(11), 3437–3471.

Zappa, G., & Shepherd, T. G. (2017). Storylines of atmospheric circulation change for European regional climate impact assessment. *Journal of Climate*, *30*(16), 6561–6577.

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., ... Taylor, K. E. (2020). Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, *47*(1), e2019GL085782.