Hadley Centre Technical Note No. 99

Assessment of the UKCP09 probabilistic land scenarios, including comparison against IPCC CMIP5 multi-model simulations.

David Sexton, James Murphy, Katy Richardson, Glen Harris, Simon Brown, and Jonathan Tinker (Met Office), and Ambarish Karmalkar (Univ. Massachusetts)

February 2016

Foreword

This Climate Science Technical Note is an abridged version of a report from the DECC and Defra funded Met Office Hadley Centre Climate Programme for 2012-15 (GA01101). The material below is a major component of the original report devoted to a scientific assessment of the UKCP09 probabilistic land projections. This assessment was designed to evaluate their continued suitability for use in the light of later climate model results contributed to the IPCC Fifth Assessment. The scientific material is unchanged from the original report, and is presented to allow interested readers to review the detailed evidence on which the top-level conclusions of the evaluation were based. The probabilistic projections were provided on a 25 km grid (and also for two sets of aggregated (river-basin and administrative) regions, and formed the core of the UKCP09 land scenarios. A weather generator tool was provided to derive local information from samples of the probabilistic future changes. A subsidiary set of spatially coherent projections (SCPs) was also provided. The SCPs were derived from an ensemble of eleven regional climate model simulations that formed part of the UKCP09 climate modelling strategy. The weather generator and SCPs are not assessed here, but information on their performance, and how they relate to the probabilistic scenarios, is available from existing reports on the UKCP09 website (see http://ukclimateprojections.metoffice.gov.uk/). Similarly, information on the UKCP09 observed trends report, and the marine and coastal projections (also not assessed here), is available from the website.

Executive Summary

The current national scenarios for 21st century climate, UKCP09, have been exploited by a wide range of users since their launch, and were an important tool in the first UK climate change risk assessment (CCRA). This report assesses the extent to which the core probabilistic component of the UKCP09 land scenarios remains suitable for adaptation planning in the light of new modelling results from the latest (CMIP5) multi-model archive of climate projections used in the IPCC Fifth Assessment Report.

The main findings are:

- The fifth IPCC Working Group 1 assessment concludes that models have in general improved since the previous IPCC report (AR4), especially in the way they represent some aspects of climate variability and extremes. However, there also remain challenges in achieving realistic simulations of some key climate processes.
- Based on a consultation with experts in the Met Office Hadley Centre and at Reading University, there have been improvements in the modelling of some aspects of the key drivers of UK climate: storm tracks; blocking; and the Atlantic meridional overturning circulation.
- A quantitative evaluation of HadCM3, the Met Office climate model that underpinned the UKCP09 land projections, and ensembles of its perturbed variants, shows these models are competitive with the latest IPCC archive, CMIP5, in terms of their performance in simulating long-term climatological averages. This is because HadCM3 was a top performing climate model in the previous IPCC report, and also because its simulations run for UKCP09 also benefited from the use of flux adjustments, included to reduce biases in sea surface temperature and to facilitate efficient sampling of modelling uncertainties.
- There is little evidence of major shifts in the climate change projections between AR4/UKCP09 and CMIP5. In particular, ranges of future change in average climatological conditions across CMIP5 models were generally found to be consistent with the probabilistic projections from UKCP09. Further work is needed to understand how future inclusion of CMIP5 information might alter UK projections in detail. In particular, the ad hoc design of CMIP5 creates challenges in understanding how best to interpret its results.

- However, some differences were found for projections of UK summer rainfall. While UKCP09 and CMIP5 agree that average summer rainfall is more likely to reduce rather than increase in the future, CMIP5 suggests smaller reductions than UKCP09, and a somewhat larger chance that UK summer rainfall could remain similar or become wetter than it is today.
- We conclude that all aspects of the UKCP09 land scenarios can still be used for adaptation planning, and in the forthcoming second UK Climate Change Risk Assessment (CCRA2). However, we recommend that CMIP5 results should be considered alongside UKCP09 in making decisions that are sensitive to future changes in summer rainfall.

1. Introduction

In June 2009, UKCP09, the latest generation of climate projections for the UK, was published. The projections contain land and marine components, and currently form the UK Government's official scientific basis for planning climate change adaptation activities in the UK. In UKCP09, the core component of the land projections (see Foreword) was, for the first time, presented in probabilistic form, accounting as comprehensively as possible for known uncertainties, in order to allow planners to take a risk-based approach to their decision making. For each of three emissions scenarios, the probabilistic information is available for several variables in each month of the year, for seven 30-year periods during the 21st century, expressed on a national 25km grid, and also as larger scale areal averages for several strategic UK regions.

The UKCP09 probabilistic projections are based on a combination of observations and two sources of climate simulations: (i) several perturbed parameter ensembles of the Met Office's 3rd generation climate model, HadCM3, and (ii) CMIP3, the model archive generated for the 4th IPCC assessment using other international climate models. The ensembles in (i) are designed to sample a number of key drivers of spread in projections. For a given emissions scenario, this spread comes from three sources of uncertainty: (i) "forcing uncertainty" arising from uncertain knowledge of the emissions of greenhouse gases and other atmospheric pollutants, their removal by physical and biogeochemical processes, and from estimation of the radiative energy imbalance associated with their time-varying concentrations in the atmosphere; (ii) for a given forcing, "modelling uncertainty" arising from imperfect understanding of climate system processes, and inevitable approximations in their representation in climate models; and (iii) "internal climate variability" on a range of time scales, arising from natural phenomena such as synoptic cyclones, atmospheric blocking episodes, ENSO events or the North Atlantic Oscillation. The HadCM3 ensembles explore uncertainties in physical land, atmosphere, and ocean processes, sulphate aerosol chemistry, and the terrestrial carbon cycle. In UKCP09, the "parametric" modelling uncertainties sampled using HadCM3 ensembles are augmented by estimates of "structural"[1] modelling uncertainty derived from CMIP3 ensembles. Observational metrics of model quality are used to constrain the projections by weighting realisations according to their ability to simulate historical mean climate (Sexton et al 2012) and large scale temperature trends (Harris et al 2013).

But climate science does not stand still. In particular, new or improved climate models are developed regularly. For the IPCC Fifth Assessment a new archive of climate model experiments was constructed called CMIP5, hence we can now assess the UKCP09 projections against this latest generation of climate modelling information. These latest models benefit from recent improvements in physical understanding of the climate system, revised estimates of the emissions and concentrations of greenhouse gases and aerosols, and from more supercomputing power. Most of the CMIP5 coupled ocean-atmosphere models used for long term projections feature atmospheric horizontal resolutions in the range 1-2^{o}, compared to 2^{o} or coarser in most CMIP3 models. Some CMIP5 models also feature better vertical resolution of the stratosphere. These shifts towards higher resolution allow the equations of motion to be solved with an improved level of precision (though CMIP5 models remain well short of the resolutions typically used in operational weather prediction). Along with improved parameterisations of sub grid-scale processes such as cloud formation and convection, these developments should provide an improved basis for the realistic

[1] Constructing a climate model involves making a number of structural choices (e.g. spatial resolution, or basic assumptions underpinning the calculations of physical phenomena or dynamical transports), and also the specification of a large number of input parameters which control the representations of physical and biogeochemical processes within a given model structure. Parametric modelling uncertainty within a single model can be estimated by varying these parameters within expert-specified uncertainty ranges. The structural component of modelling uncertainty is typically estimated by the spread of a multimodel ensemble which samples different ways to build climate models, with the caveat that this approach inevitably cannot account for the unknown effects of errors common to all current models.

simulation of climatalogical characteristics such as mean temperature or rainfall (e.g. Reichler and Kim, 2008), or high-impact events such as storms, heatwaves or floods. However, since such emergent properties invariably result from a complex balance of processes, it cannot be guaranteed that such model developments will always lead to reduced biases in specific output variables of interest to users. The new climate models have also grown in complexity, with the addition of more processes. This is particularly true regarding the representation of aerosol species, and aerosol-cloud interactions. Indeed, the CMIP5 archive features a number of new earth system models, containing explicit representations of vegetation, ocean biogeochemistry and the earth's carbon cycle, in addition to atmosphere, ocean, land, sea-ice and aerosol modules. Such simulations provide a more complete basis for the representation of uncertainties, but create additional sources of potential model bias, since certain properties which were previously prescribed from observations, such as vegetation distributions, are now allowed to respond interactively to simulated physical drivers. In some models, increased complexity may therefore offset some of the improvements in output variables that might be expected from developments in the physical components of the models.

These CMIP5 earth system models have been deployed in both concentration-driven and emissions-driven simulations, the latter accounting for the interactive effects of physical and carbon cycle feedbacks on future CO_2 concentrations and temperature changes (Booth et al, 2013). This information may be useful in updating the sampling of carbon cycle feedbacks included in a future update of UKCP09. However, the largest of the CMIP5 future climate change experiments features simulations driven by prescribed historical and future CO_2 changes, in which the impacts of carbon cycle uncertainties on future temperature changes are not accounted for. This larger concentration-driven experiment is the focus of the comparisons here.

This report assesses the UKCP09 probabilistic land projections in the light of CMIP5. This includes comparisons between simulations of historical climate from CMIP3-generation models forming the basis of UKCP09 and CMIP5, consisting of a brief recap of the main conclusions of the comprehensive evaluation recently carried out for IPCC Fifth Assessment (section 2a), supported by quantitative analysis of selected basic variables in section 2b. In section 3, the comparison is extended to cover key aspects of future projections in UKCP09 and CMIP5, focusing mainly on global mean temperature changes, and changes in multiyear mean surface temperature and precipitation over the UK. Some specific drivers of UK climate are considered in section 4, reporting assessment of recent research derived from a survey of experts at the Met Office Hadley Centre and the University of Reading, who have analysed relevant aspects of CMIP5.

In section 3, some results compare probability distributions based on UKCP09 against multimodel realisations from CMIP5. While this provides valuable evidence regarding the status of UKCP09, it is important to appreciate that probabilistic climate projections and multimodel ensembles are different scientific entities, implying a need for caution in interpreting differences between the two products. The UKCP09 probability distributions were constructed by combining results from perturbed parameter ensembles with CMIP3 multimodel information and a set of observational constraints using a formal statistical framework (Sexton et al, 2012; Harris et al, 2013). We would not therefore expect UKCP09 to have provided uncertainty ranges identical to those given by the CMIP3 information in isolation. The same applies to the present comparison with CMIP5, which provides a more up-to-date multimodel archive of the type which contributed one aspect of the uncertainty accounted for in UKCP09. For instance, the multimodel ensemble, while essential for the sampling of structural modelling uncertainty, is an ad hoc construct that should not in isolation be interpreted probabilistically, and cannot easily be subjected to the type of rigorous statistical framework used to apply observational constraints in UKCP09 (Knutti et al 2010; Sexton et al 2012).

2. Performance of UKCP09 and CMIP5 models during historical period

2a. Evaluation of CMIP5 models in IPCC Fifth Assessment report

Here, we briefly summarise the assessment of the performance of CMIP5 models given in chapter 9 (Flato et al 2013) of the recently published IPCC Working Group I Fifth Assessment report (AR5), and presented in terms relative to the CMIP3 models evaluated in the fourth IPCC assessment (AR4). As mentioned in the Introduction, climate models have been enhanced in a number of ways since CMIP3, either by including more processes such as biogeochemical cycles and aerosol-cloud interactions, or simply by increases in high performance computing affording higher resolution both horizontally and vertically. Some models (designated as "high-top" models) have been extended to include a fully resolved stratosphere with a model top above the stratopause, and this has allowed for better simulation of stratospheric-tropospheric interactions and their potential effects on future climate change (see section 3). Another development since CMIP3 is that climate models are now routinely evaluated across a wider range of variables, including metrics relating to aerosols, the carbon cycle, extreme events and climate variability on intraseasonal to decadal time scales, as well as long-term averages and historical trends relating to physical properties of the atmosphere and ocean. Some of the key drivers of UK climate fall into the climate variability category, and these are discussed in section 4.

Flato et al (op. cit.) conclude that in general models have improved since CMIP3, though this statement depends on the variables under consideration. For a given variable, there is typically considerable overlap between the two ensembles in terms of the performance spread of constituent individual models. No instances are found where the CMIP5 ensemble as a whole performs significantly worse than the CMIP3 ensemble; for some variables the performance levels are similar, and for others the CMIP5 ensemble performs better.

For example, simulations of large scale patterns of average precipitation have improved somewhat in CMIP5, though significant errors remain at regional scales. Large scale surface temperature patterns are well simulated in both CMIP3 and CMIP5, with some improvements in CMIP5 at regional scales. The simulation of precipitation extremes is also improved in CMIP5; this is partly related to increases in spatial resolution (see Introduction).

However, Flato et al. also conclude that there are still some major challenges for climate modelling, and that in these areas there has only been a modest improvement since CMIP5. The modelling of clouds is one prime example. Similarly, whilst aspects of the Madden-Julian Oscillation (MJO) are better represented in CMIP5 models, their overall ability to reproduce the observed characteristics of this key driver of tropical intraseasonal variability is still considered to be low. For ENSO, the simulation of the tropical Pacific Ocean mean state has improved since the AR4, with a 30% reduction in the spurious westward extension of the cold tongue near the equator, a pervasive bias of coupled models. Whilst there have been improvements in ENSO teleconnections to the whole Pacific region, teleconnections beyond this region are still considered to be poorly represented. Also, whilst increases in resolution and better treatment of cloud physics have both been linked with improvements in the simulation of the diurnal cycle of precipitation, the majority of CMIP5 models, like CMIP3, tend to start moist convection too early in the day over land in the Tropics. Over the UK, Kendon et al (2012) find that a 12km configuration of the Met Office Unified Model gives a similar error on summer days dominated by convective events, whereas a 1.5km configuration in which the standard convection parameterisation is switched off (relying instead on the model dynamics to resolve convective storms) gives an improved diurnal cycle with a midafternoon peak corresponding with observations. In UKCP09, it was not attempted to provide information on how climate change drivers might affect sub-daily precipitation distributions, because the underpinning climate modelling did not include an explicit representation of convective dynamics.

2b. Quantitative comparisons between core UKCP09 simulations and CMIP5

In this sub-section we provide some quantitative comparisons between CMIP5 models and one of several perturbed parameter ensembles of variants of the HadCM3 coupled ocean-atmosphere model that were run for the UKCP09 land projections and other applications. This ensemble was run from 1860-2100, and consisted of 17 variants sampling multiple perturbations to 30 surface and atmospheric parameters. The perturbations were designed to sample a wide range of uncertainty in global climate feedbacks, while providing credible simulations of historical climate (Collins et al, 2011). It was derived from a larger 280-member ensemble of simulations of the equilibrium response to doubled CO_2, carried out using the "slab"[2] configuration of HadCM3. One of its key roles in the UKCP09 methodology was to provide the basis for the calibration of pattern-scaling relationships allowing estimates of idealised equilibrium climate change to be converted into corresponding estimates of time-dependent ("transient") changes. This allowed estimates of equilibrium change featuring comprehensive sampling of the atmospheric model parameter space to be converted into samples of transient change forming the kernel of the probabilistic projections of changes during different 21st century periods (Harris et al, 2013). This 17-member HadCM3 ensemble (termed QUMP, following Harris et al, 2013) also provided driving data for an 11-member ensemble of regional climate model simulations for the UK and Europe, another key component of UKCP09. We necessarily limit our analysis of ensemble performance to comparison of QUMP against CMIP5 coupled ocean-atmosphere simulations, as the CMIP5 experiment protocol did not include production of simulations based on slab model configurations.

We focus mainly on global performance metrics relating to the simulation of long-term climatological averages of a set of physical variables commonly used to evaluate climate models. This is consistent with the approach used to apply weights to alternative model variants in UKCP09 (Sexton et al, 2012), which was based on an assumption that since influences on climate change in some region can in general arise from a complex mix of local and remote dynamic and thermodynamic influences, the use of metrics designed to estimate model quality on a global basis, and across a broad basket of variables, is appropriate. However, we note that various approaches are conceivable for the quantitative assessment of model suitability for use in future projections, and there is currently no clear basis for identifying an optimum method (Knutti et al, 2010; Flato et al, 2013). For example, some studies have utilised approaches focused more on the use of region-specific metrics (see McSweeney et al (2012) for one recent example).

[2] A slab model is an atmosphere model coupled to a very simple thermodynamic ("slab") ocean, rather than to a fully dynamic ocean model. This provides a relatively inexpensive configuration with which to sample parametric model uncertainties efficiently, in this case via large ensembles of the equilibrium response to an idealised doubling of CO_2 concentrations.

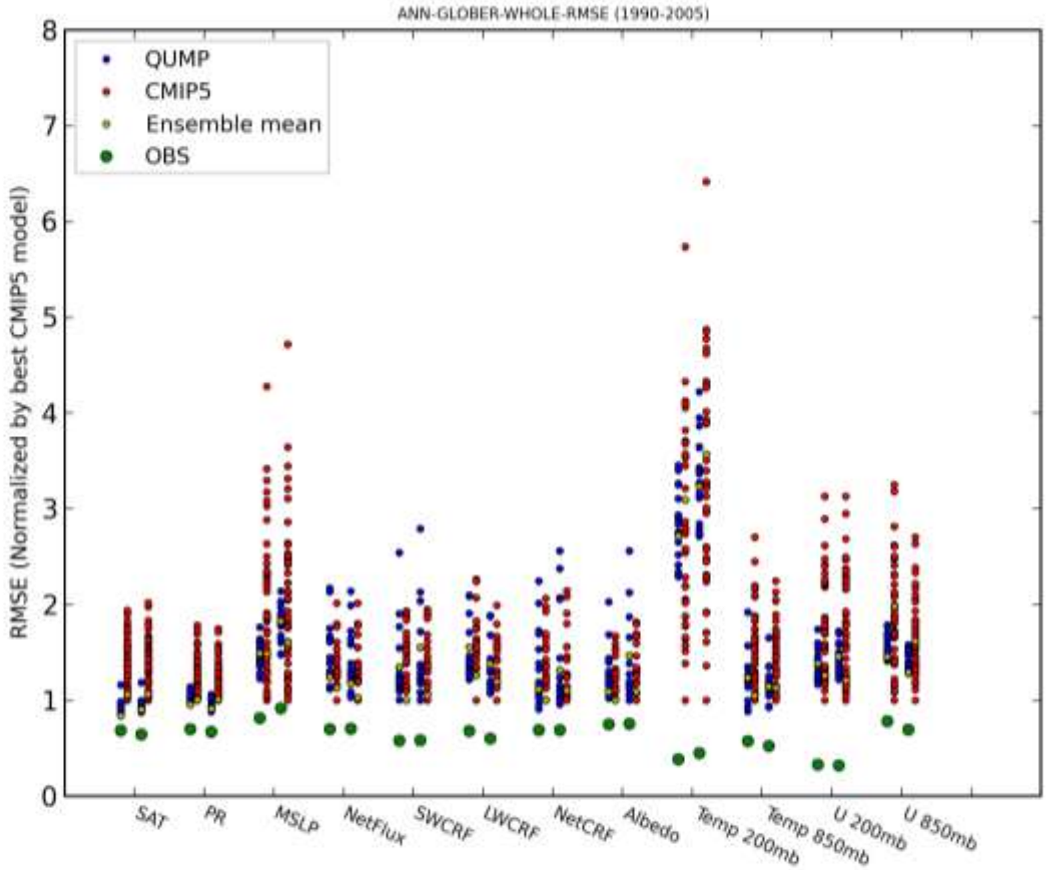

Figure 1. *Comparison between 17 perturbed HadCM3 variants (QUMP) and available CMIP5 coupled ocean-atmosphere models of normalised root mean square errors (RMSE) in simulated 15 year averages of annual values of several climate variables for the period 1990-2005. All model output is regridded to HadCM3 2.5x3.75º latitude-longitude grid prior to the calculation of RMSE values, which are obtained by averaging regional squared errors from 60ºN-60ºS. Polar latitudes are excluded because uncertainties in some of the observational datasets are particularly large in those regions. For each climate variable there are two sets of results, as the analysis is done for two observational data sets (green dots indicate the RMSE between the two observational data sets). To normalise the results, each RMSE value is divided by that of the best performing CMIP5 model for each variable and each observational data set. Therefore for each column of red dots, the lowest value is guaranteed to be 1. The olive dots are the RMSE values of the mean of the relevant ensemble.*

Fig. 1 shows root-mean squared error (RMSE) values for historical long-term averages of spatial fields of a number of key climate variables for the region 60ºN-60ºS, including both land and ocean areas. These consist of air temperature at the surface (SAT) and in the lower and upper troposphere (Temp 850mb, Temp 200mb), precipitation (PR), mean sea level pressure (MSLP), the net radiative flux at the top of the atmosphere (NetFlux), shortwave, longwave and net effects of cloud on NetFlux (SWCRF, LWCRF, NetCRF), the fraction of incoming shortwave solar radiation returned to space (Albedo), and the westerly wind component in the lower and upper troposphere (U 850mb, U 200mb). This set of variables covers several basic emergent properties of climate, all influenced to varying degrees by atmospheric dynamics, radiative transfer, the water cycle and surface exchanges. For the relevant ensemble member and variable, the RMSE value essentially represents the globally-averaged magnitude of regional errors. In each case, values are normalised by the value of the best-performing CMIP5 model (which by definition takes a value of unity) to give a relative measure of error. Observational uncertainties are accounted for by providing two sets of results for each variable, based on comparisons with two alternative verification datasets. The results show that the QUMP ensemble generally compares favourably with CMIP5. For most variables the best QUMP performers would rank amongst the top CMIP5 models. For surface air temperature and precipitation, there are many CMIP5 models which score worse than the least skilful of the QUMP variants, whilst a number of the QUMP members give scores slightly better than any of the best CMIP5 models. For the tropospheric temperature and wind metrics, the QUMP simulations span a narrower range of scores than CMIP5: Temp 200mb is the weakest variable regarding QUMP performance, however the largest biases always come from the CMIP5 ensemble, while several CMIP5 members outperform the best of the QUMP variants (with the exception of Temp 850mb). One or two QUMP members do score worse than all CMIP5 models for some of the radiative flux diagnostics, however even in these cases the two ensembles span similar overall ranges of performance, and several individual members of QUMP are competitive with the better CMIP5 models. Figure 1 also demonstrates that all contemporary models still contain significant systematic biases, since the best performing models (or model variants) invariably give larger RMSE values than those obtained from the differences between the two observational datasets (green dots in Figure 1).

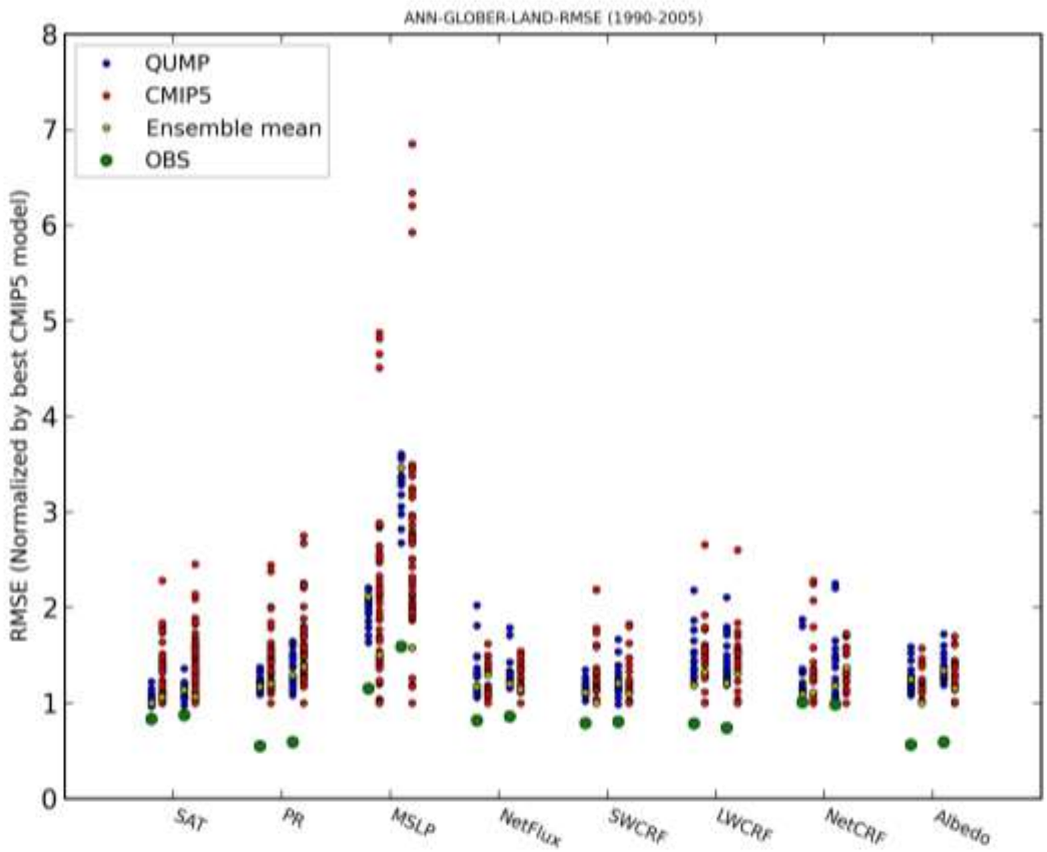

Figure 2. *As Figure 1 but for land regions only, and for a reduced set of variables.*

The fact that the precipitation simulations in QUMP stand up well against CMIP5 is noteworthy, considering that the AR5 found a general improvement since CMIP3 in the simulation of global precipitation patterns (see section 2a). We return to this point in the discussion of Fig. 4 below.

Figure 2 provides corresponding normalised RMSE values for some of the variables of Figure 1, calculated for land regions only. The ensemble of scores for SAT and precipitation in QUMP is highly competitive with that for CMIP5 models, although none of the QUMP variants outperforms the best CMIP5 model, in contrast to the results in Figure 1 for land and sea regions combined (see Figure 4 and related discussion below). The results for other variables in Figure 2 are similar to those of Figure 1, demonstrating that the QUMP simulations of annual mean baseline climatology are competitive with those of the CMIP5 ensemble in populated regions.

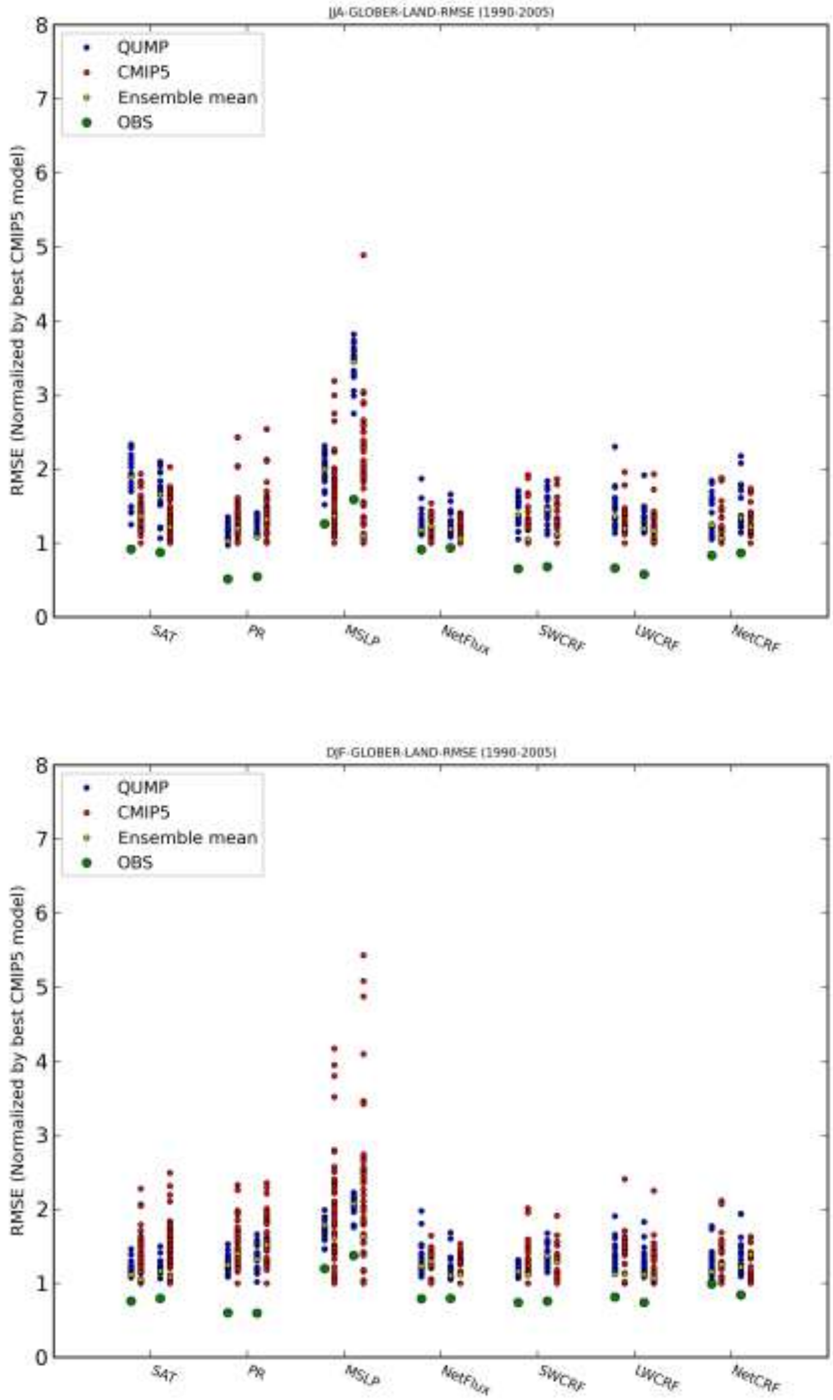

Figure 3. *As Figure2 but for Northern Hemisphere land only, for summer (top panel) and winter (bottom panel).*

Corresponding results for land climate in northern hemisphere summer (June to August, JJA) and winter (December to February, DJF) are shown in Figure 3. These seasonal comparisons of QUMP and CMIP5 climatology generally support the favourable results found for annual mean climatology in Figure 2. There are only a few variables (some of the radiative fluxes and JJA surface air temperature) where the worst simulation belongs to the QUMP ensemble rather than CMIP5. In JJA, errors in surface air temperature are typically somewhat larger in QUMP than in CMIP5. This arises from a warm bias in central regions of the Eurasian and north American continental landmasses (Murphy et al, 2009). Analysis of a more recent ensemble of variants of the earth system configuration of HadCM3 (Murphy et al, 2014) shows that warm biases in summer SAT over continental regions in the northern hemisphere are largest in model variants which simulate the lowest values of soil moisture content. It may be possible to use this relationship to refine the weights assigned to model variants in a future update of UKCP09, although the presence of large uncertainties in observed soil moisture may limit the effectiveness of such a constraint. For mean sea level pressure in JJA, the performance of QUMP relative to CMIP5 appears acceptable when verifying observations from the HadSLP2 dataset (Allan and Ansell, 2006) are used (left-hand set of dots in the MSLP column of Figure 3), but not when ERA-Interim reanalysis data (Dee et al, 2011) is used (right-hand set of dots). The different scores reflect large differences between the verification datasets in mountainous regions (especially over the Tibetan plateau). The RMSE scores of Figure 3 suggest that several CMIP5 models score better against either observational dataset than the two datasets score against each other, however this is more likely to reflect limitations of the verification data than a true absence of systematic biases in any of the models. These results should therefore be interpreted with caution. Over ocean regions (not shown), RMSE scores are much more consistent between the two observational datasets, and QUMP performs well with respect to CMIP5.

Overall, Figs. 1-3 suggest that QUMP performs well against CMIP5, with the exception of summer surface air temperature over northern hemisphere continents. However, it is important to note that the QUMP simulations used flux-adjustments[3] in order to reduce systematic biases in regional sea surface temperature (SST) and salinity values, and also to reduce the risk of under-sampling the effects of parametric model uncertainties. Some CMIP3 models also used flux adjustments, but no CMIP5 models do. While the use of flux adjustments can be justified in a climate prediction context (as a means of supporting uncertainty quantification and reducing the effects of historical SST biases on simulated future changes), their use does not improve the fundamental quality of a climate model from a process perspective. It is therefore important to understand to what extent their use may favourably influence the scores of model performance metrics, compared to the scores for models which do not use flux adjustments.

In Figure 4 we compare RMSE values for the key variables of surface air temperature, precipitation and MSLP, found in flux-adjusted and non-flux adjusted configurations of the version of HadCM3 using standard parameter settings (noting that the non-flux adjusted version also omitted aerosol-cloud interactions). Values for CMIP5 coupled ocean-atmosphere models are also shown. RMSE values are calculated over both land and ocean, and are normalised in this case using the spatial standard deviation of the observed field, so that the values represent typical regional errors relative to the observed contrasts between different regions. The black crosses in Figure 4 show a combined skill metric for each model, obtained by averaging the normalised RMSE value for the three variables. The flux-adjusted HadCM3 simulation gives a better combined score than all CMIP5 models in JJA (lower panel), and for all but one model in DJF (top panel). The HadCM3 simulation without flux adjustments gives a combined score within the body of CMIP5 results, though somewhat below the average performance level.

[3] Flux-adjustment is an empirical technique to limit the development of systematic biases in sea surface temperature and salinity in coupled models. The regionally and seasonally varying adjustments consist of additional surface fluxes of heat and moisture designed to counter the combined effects of climatological errors in surface exchanges and ocean transports. Flux adjustments were used in QUMP (and other coupled ocean-atmosphere ensembles run for UKCP09) to avoid the risk that small structural deficiencies in the simulation of planetary energy balance might restrict the sampling of parametric model uncertainties to an artificially small region of the model parameter space (Collins et al, 2011) .

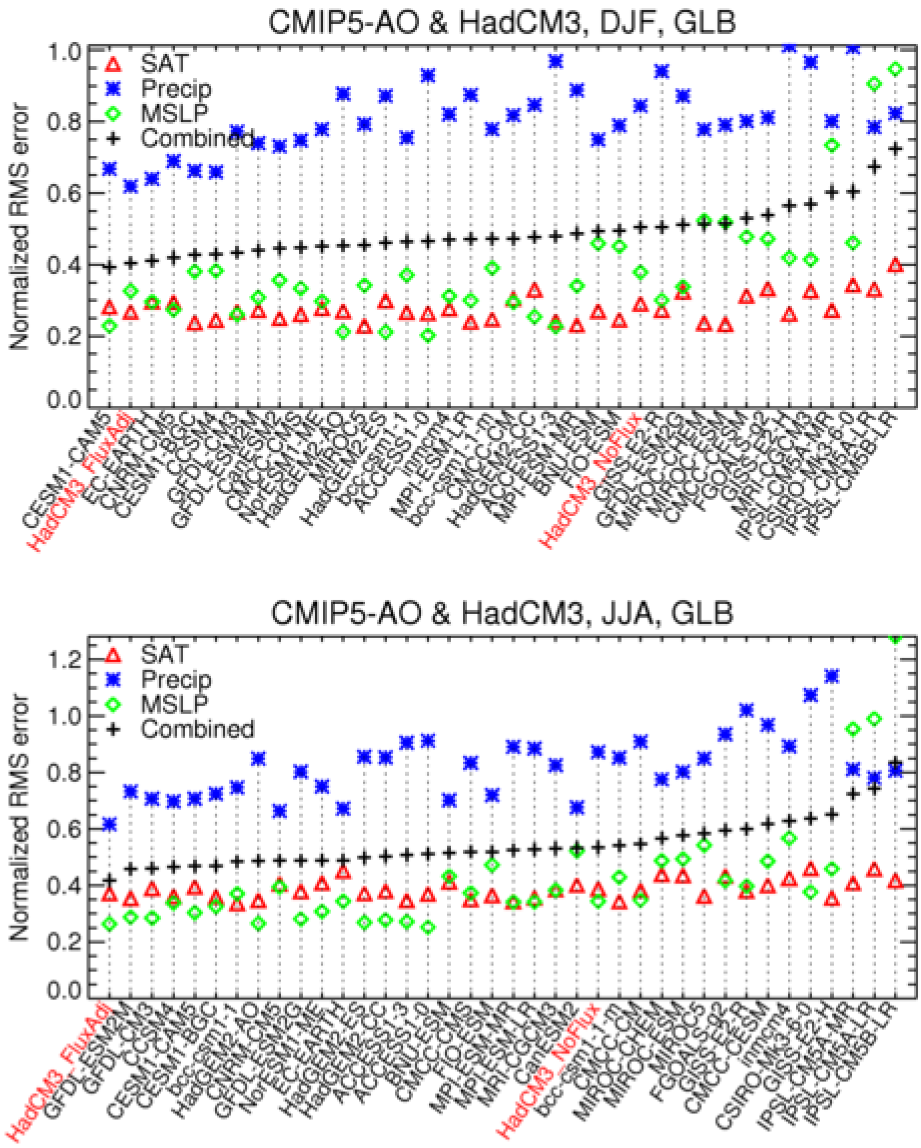

Figure 4. *Normalised root mean square errors of 1.5m air temperature, precipitation, and MSLP over the whole globe for CMIP3 and CMIP5 models. The normalisation is different to that used in Figures. 1-3; here the mean square error is divided by the spatial gridpoint-wise standard deviation of the observed climatological field. Models are ranked according to a combined score, obtained by averaging normalised RMSE values for each of the three variables. Top panel shows DJF, bottom panel shows JJA.*

Over land (Figure 5), the impact of flux adjustments is much smaller: both HadCM3 simulations give scores towards the upper end of the CMIP5 performance range. In Fig. 4, the main impact of using flux adjustments is to improve the precipitation score. This arises mainly from improvements over the tropical oceans. For example, McSweeney et al (2012) note improvements in the tropical Pacific, resulting from partial correction of a cool bias in the region of the equatorial cold tongue in SST. Such improvements are also consistent with previous work demonstrating reduction of precipitation biases resulting from the correction of SST errors (e.g. Ashfaq et al, 2011).

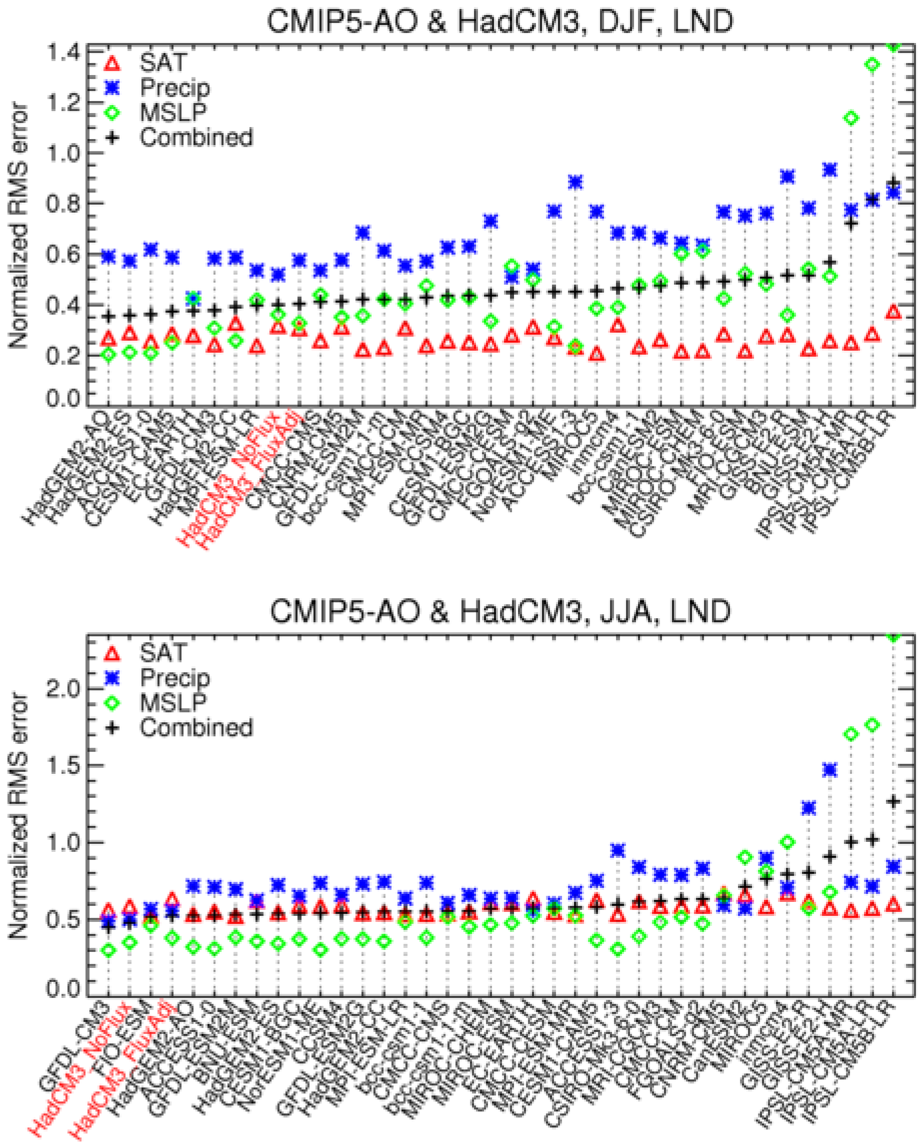

Figure 5. *As Figure 4, for global l land regions only.*

Overall, Figs. 4 and 5 show that the use of flux adjustments does play a role in improving the performance of HadCM3 in simulating long term mean climate, particularly for precipitation over the oceans, however the performance of HadCM3 remains reasonably competitive with CMIP5 models even when flux adjustments are not used, especially over land. This suggests that the representation of processes determining the climatological average characteristics in this core UKCP09 ensemble remains consistent with the latest performance benchmark set by CMIP5, despite the improvements in CMIP5 performance relative to CMIP3 models summarised in section 2a. This is because HadCM3 is one of the better CMIP3 models (e.g. Reichler and Kim, 2008), and was therefore an appropriate choice as the basis for the systematic approach to uncertainty quantification, achieved via its use in perturbed parameter ensembles for UKCP09.

So far we have focussed on assessing the performance of QUMP in simulating the mean climate, consistent with the status of the probabilistic component of UKCP09 as projections of future change in long term (30-year) averages. However, the occurrence of high-impact climate extremes is also strongly influenced by climate variability at finer time resolutions (potentially ranging from sub-seasonal to decadal time scales). Ultimately, provision of the best possible information on variability within a changing climate is dependent on developing new (post-CMIP5) models capable of improved simulations of a wide range of key phenomena, such as the remote influences on European climate of Atlantic SSTs (Sutton and Dong, 2012), ENSO events and stratosphere-troposphere interactions (Ineson and Scaife, 2009), blocking (Scaife et al, 2011), and extremes driven by local convective events (Kendon et al, 2012).

Here, we provide a basic assessment of the simulated historic North Atlantic Oscillation (NAO) in QUMP and CMIP5 models, given that the NAO is a key driver of interannual to decadal variability in both winter and summer, and that some CMIP5 models may offer improvements in the simulation of detailed processes influencing the NAO, for example due to enhancements in horizontal and vertical resolution. In particular, improved resolution of the stratosphere in some CMIP5 models (see section 2a) may improve simulation of the influences of both ENSO (Ineson and Scaife, 2009) and solar cycles (Ineson et al, 2011) on the winter NAO.

The NAO index is defined as the difference in mean sea level pressure between Azores and Iceland. Positive values of NAO indicate westerlies and warmer winters, negative values in winter often mean cold weather. Figs. 6 and 7 show that the QUMP ensemble compares very well with CMIP5 models in its simulation of the NAO variability in winter and summer. In both winter and summer, the two ensembles tend to overestimate observed variability on 1-30 year time scales to some degree, while variability on longer time scales is underestimated in nearly all ensemble members in summer, though not in winter.

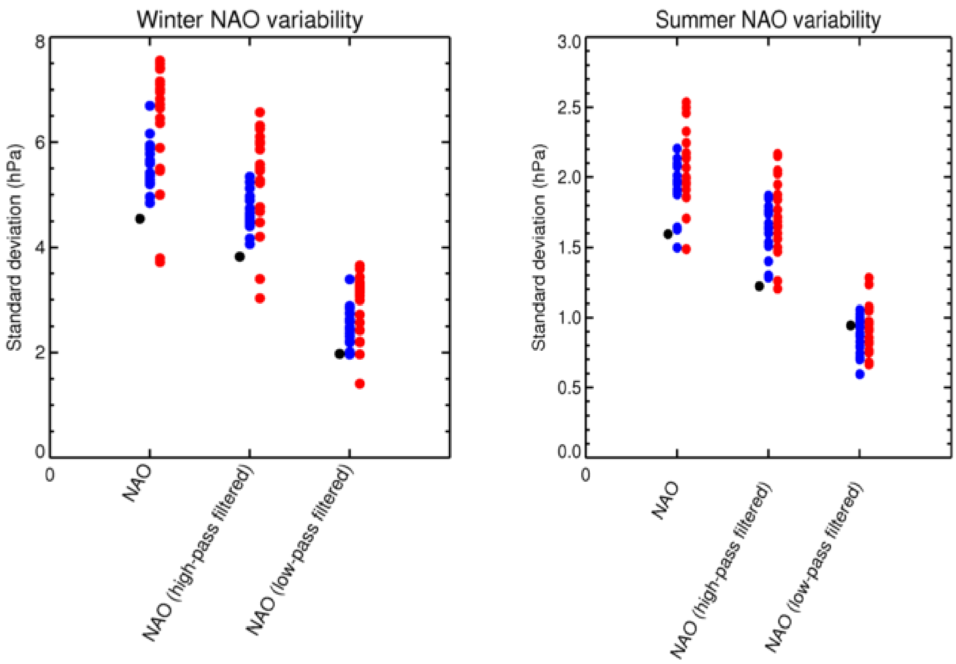

Figure 6. *Comparison of the standard deviation of the North Atlantic Oscillation index in winter (left panel) and summer (right panel) from 1861-2010 for QUMP (blue) and CMIP5 (red) simulations, and observations (black). Comparisons are also made for filtered frequencies higher and lower than the 10-year time scale.*

While the results of Figs. 6 and 7 demonstrate that the *magnitude* of NAO variability in the core modelling behind UKCP09 is quite realistic, and competitive with CMIP5 models, it does not follow that either set of simulations captures correctly all the driving mechanisms. For example, the QUMP simulations, and many of the CMIP5 models, will not capture potential stratospheric influences, due to limited vertical resolution.

3. Future projections in UKCP09 and CMIP5

In section 3a we assess projections of global mean temperature from the UKCP09 methodology, based on the latest understanding provided in IPCC AR5. Section 3b then extends the comparison to the UK, focusing mainly on surface temperature and precipitation changes, and showing some quantitative comparisons between UKCP09 and CMIP5.

3a. Global temperature

Globally averaged temperature changes over the 21st century are a key driver of regional changes projected over the UK for many of the UKCP09 variables, noting that uncertainties in spatial patterns of change and the effects of climate variability are also important, particularly for variables strongly influenced by regional aspects of the water cycle and atmospheric circulation. For a given emissions scenario, the spread of future global temperature changes in UKCP09 is mainly determined by the ranges of global climate feedbacks arising from physical processes, and, after the 2050s, the strength of carbon cycle feedbacks. In addition to simulations from complex three-dimensional climate models, our understanding of how feedback processes can be constrained by appealing to observations is an important determinant of assessments of their credible ranges at any given time. In particular, IPCC AR5 assessed results from a number of recent studies aimed at inferring the future sensitivity of the climate system from observed changes during the 20th century. Below, we briefly assess current understanding regarding (i) physical climate feedbacks, (ii) observational constraints on physical feedbacks offered by 20th century climate change, and (iii) carbon cycle and other earth system feedbacks, before comparing projections of global mean temperature between QUMP, CMIP3 and CMIP5 simulations.

(i) Physical climate feedbacks

The equilibrium climate sensitivity (ECS) and transient climate response (TCR) are two standard benchmarks of the physical response of the climate system to increases in greenhouse gases. ECS is defined as the equilibrium change in annual mean global surface temperature to a doubling of atmospheric CO_2 concentration, while TCR is defined as the change in annual mean global surface temperature at the time of CO_2 doubling, following a 1% per year increase in CO_2 over a period of 70 years. Both metrics are influenced by physical feedback processes affecting the surface and atmospheric response to the imposed changes in greenhouse gases, while the TCR is also influenced by the rate of uptake of heat by the oceans, and is the metric most relevant to understanding rates of 21st century warming projected in model simulations driven by non-intervention emissions scenarios (such as those used in UKCP09).

Vial et al (2013) show that the range in equilibrium climate sensitivity explored by eleven CMIP5 models is similar to CMIP3. This is supported by results from the larger sample of 23 CMIP5 models given by Flato et al (2013) in IPCC AR5. As in CMIP3, Vial et al find that the main contribution to the uncertainty in their CMIP5 sample comes from a spread in tropical cloud feedbacks, primarily arising from shortwave cloud properties in regions of shallow convection. The other major physical feedback components, arising from changes in water vapour, tropospheric lapse rate, sea-ice and snow albedo, continue to provide a robust net positive feedback in CMIP5 models, contributing relatively modest uncertainty to ECS compared to cloud feedbacks.

The 90% credible range for ECS in UKCP09 is 2.4-4.3ºC. This result was derived using the methodology described by Sexton et al (2012) and Harris et al (2013), based on climate model results from HadCM3 perturbed parameter ensemble simulations combined with CMIP3 models, and formally constrained by a multivariate set of observational metrics of model performance derived from recent multiyear mean climate, and also from historical trends in several large scale temperature indices (Braganza et al., 2003). The IPCC AR5 assesses a likely range of 1.5-4.5ºC, obtained from an expert assessment of several lines of evidence (Collins et al, 2013). The UKCP09 PDF of climate sensitivity suggests only a 5% probability of an outcome smaller than 2.4ºC, significantly above the AR5 estimate of 1.5ºC. Possible reasons for this difference are discussed in sub-section (ii) below.

The Transient Climate Response is assessed to have a likely range of 1.0-2.5ºC by IPCC AR5, again based on several lines of evidence including results from CMIP5 models. The UKCP09 methodology gives a corresponding 90% range of 1.6-2.4ºC (Harris et al, 2013). As for ECS, the upper ends of the two ranges are similar, while the lower end is larger in UKCP09.

(ii) Observational constraints on global climate sensitivity from 20th century climate change

In Box12.2, Figure 1 in IPCC AR5 Chp.12 (Collins et al 2013), evidence for ECS values lower than 2ºC comes mainly from several recent studies that derive ECS distributions by applying some form of simple climate model to simulate surface temperature trends and ocean heat uptake during some or all of the 20th century, which are then compared with observed estimates (e.g. Aldrin et al, 2012; Lewis, 2013; Otto et al, 2013). In these approaches, the implementation of climate sensitivity in the simple climate model is typically a single input parameter, hence any value can be prescribed "from the top down" in order to test how well different simple model realisations match the observed historical changes. In contrast, ECS in a comprehensive climate model *emerges* out of simulated interactions between all the detailed physical and biogeochemical processes represented in the model. The simple model techniques are dependent on a number of choices and assumptions regarding historical radiative forcing components and their uncertainties, the efficacy with which different forcing components drive a surface temperature response, the estimated uncertainties attached to the constraining observations, the assumed characteristics or sampling of observed internal climate variability, which parts of the historical record are used to provide constraints, and the choice of statistical methodology used to sample possible ECS values and attach estimates of relative likelihood to them. For example, some studies of this type reported in IPCC AR5 report smaller probabilities for low ECS values, compared with the three studies highlighted above.

The lowest ECS values amongst CMIP3 or CMIP5 models is 2ºC (see Figure 8 below). That is, it is difficult to find a comprehensive climate model with a plausible representation of the mean climate that can sample the low end of the IPCC AR5 ECS range, for example by providing a negative cloud feedback strong enough to counter the net positive feedback from the water vapour, tropospheric lapse rate and albedo components. Nevertheless, an important task for future work is to explore whether the differences can be reconciled between ECS distributions derived from climate model ensembles, and those obtained from some of the estimates based on observed post-industrial climate change. The UKCP09 methodology combined both types of approach, being based on climate model ensembles constrained by observations of present day climate, but also using a simple model combined with pattern-scaling techniques to add further constraints based on 20th century surface temperature changes.

Harris et al (2013) showed that the range of historical aerosol forcing explored by UKCP09 was in line with the distribution provided by IPCC AR4, based on assessment of modelling and observational information available at the time. In the fifth assessment, a new distribution has been supplied (Myhre et al, 2013). While uncertainty in aerosol forcing remains substantial, the AR5 distribution shows a modest shift towards less negative values. Myhre et al attribute this mainly due to a less negative contribution from aerosol-cloud interactions.

However, it is important to recognise that understanding of past changes in aerosol forcing is still developing, as is the capability to represent aerosol species and processes in climate models. So the distribution given by AR5 is itself likely to be further updated in future. For example, Carslaw et al (2013) have recently shown that a major uncertainty for estimates of present day aerosol forcing arises from the background emissions of natural sources such as dimethyl sulphide, volcanic eruptions, sea salt and natural wildfires, alongside uncertainties relating to past anthropogenic emissions or aerosol processes. This uncertainty is not represented in CMIP5, UKCP09, or the IPCC AR5 assessment. One of their conclusions is that our limited ability to constrain the natural aerosol state will hamper "top-down" studies of the type discussed in this sub-section, that explore combinations of ECS, ocean heat uptake, and aerosol forcing to find those that best explain observed temperature changes.

(iii) Carbon cycle feedbacks and other Earth System processes

An estimate of the effects of uncertainties in terrestrial carbon cycle processes on future global mean temperature changes was represented in UKCP09. Model results available to date suggest that the net carbon cycle feedback is positive (albeit with a highly uncertain magnitude), and leads to an enhancement in global temperature change, especially during the second half of the 21st century. Carbon cycle feedbacks can be split into components that depend on the amount of carbon in the atmosphere (carbon-carbon feedbacks), and components that relate to temperature change (climate-carbon feedbacks) (Gregory et al 2009). Compared with the C^4MIP experiment (Friedlingstein et al, 2006) assessed in IPCC AR4, there has not been much change in the range of these two feedbacks in CMIP5 Earth system models that do not represent the nitrogen cycle.

The UKCP09 science report (Murphy et al, 2009) emphasised that the projections could only account for uncertainties in drivers of future change sufficiently well understood to be included in climate or earth system models available at the time of their construction. Here, we highlight two factors that are both likely long term developments in Earth system modelling, that could lead to increase in the upper tail of the spread of plausible future temperature increases:

Nitrogen cycle. Inclusion of the nitrogen cycle is likely to feature in the long term development of Earth System model capabilities. As a principal nutrient for plant growth, nitrogen can both limit future carbon uptake and stimulate it, depending on changes in nitrogen availability. Only two Earth System models in CMIP5 include the nitrogen cycle. These two models suggest that carbon-carbon feedbacks become less negative, as in nitrogen-limited soils there is a limit on plant growth which restricts how much vegetation can take up carbon, hence more carbon remains in the atmosphere. In contrast, the climate-carbon feedbacks become less positive, as increasing temperatures enhance soil organic matter that in turn makes more nitrogen available to plants, thus enhancing nitrogen uptake and carbon storage in vegetation. The carbon-carbon mechanism is considered to be robust, and wins out compared with the reduced temperature feedback in studies conducted to date (Bonan and Levis, 2010; Zaehle et al, 2010a). Therefore, the inclusion of the nitrogen cycle in Earth System models could lead to an increase in atmospheric CO_2 and greater temperature rise (Ciais et al 2013).

Permafrost thaw and the subsequent release of methane or CO_2 are not included in any CMIP5 models. IPCC AR5 (chapter 12 and chapter 6 respectively) assesses that permafrost extent is virtually certain to decrease in the future, but attributes low confidence to quantitative estimates of carbon lost as a result. Burke et al (2012) quantify relative sources of uncertainty, and find significant contributions from the chosen future emissions scenario, biases in initial permafrost carbon pools and detailed descriptions of soil processes within models. Burke et al (2013) further examined physical permafrost thaw across CMIP5 models, and found a wide spread of both initial permafrost extent and the degree of future change. They assessed future carbon loss from permafrost using a diagnostic of the present day permafrost store and future changes in soil temperatures from CMIP5 models, estimating that by 2100 there would be a likely 25-50ppm increase in carbon in the atmosphere. This would lead to further temperature increase of a few tenths of a degree Celsius by the end of 21st century.

We note also that carbon release from methane hydrates in the ocean, and from warm wetlands becoming wetter in the future, are further potential feedback mechanisms not included in current earth system models.

(iv) Comparison of global temperature projections

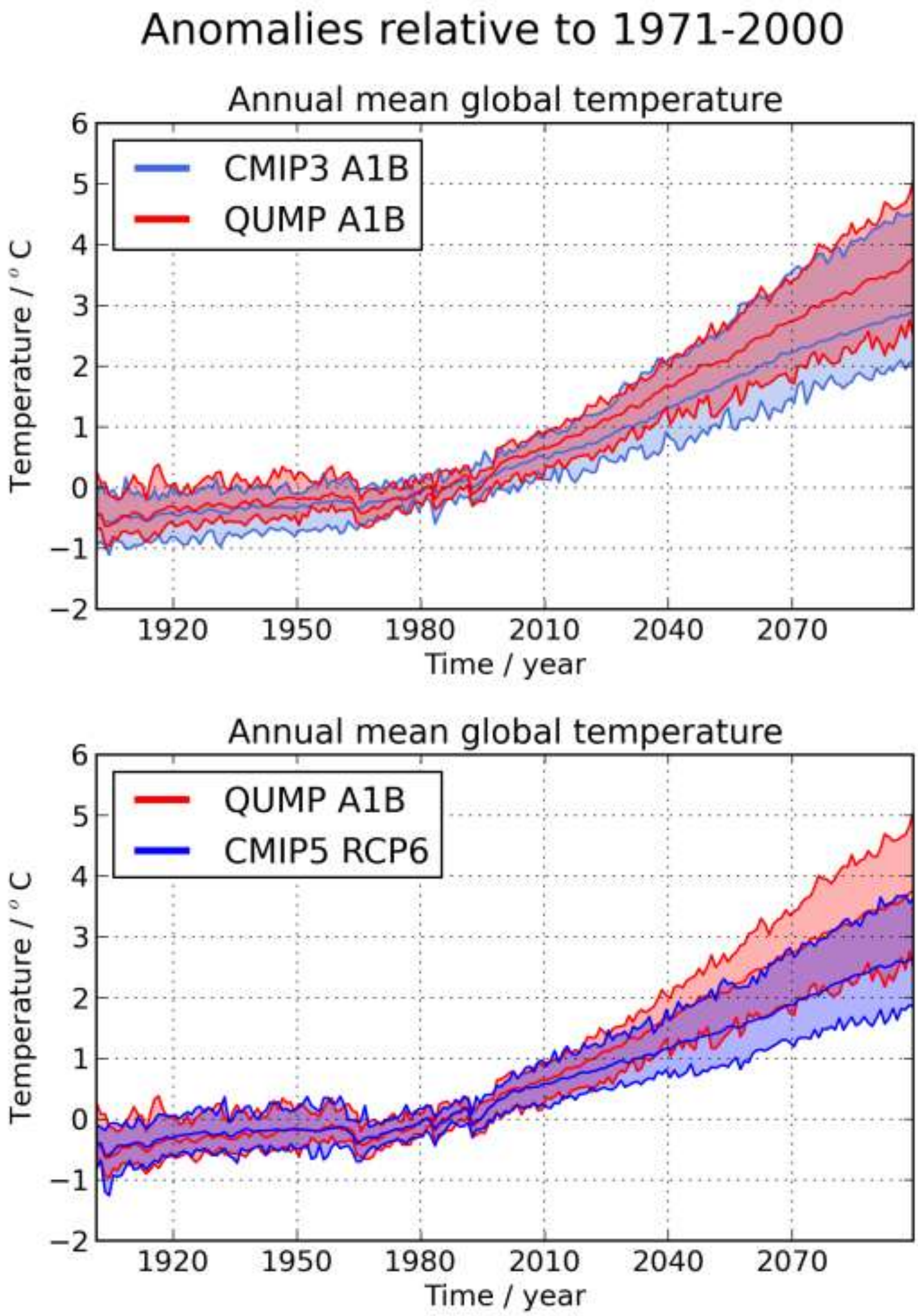

Figure 7. *Plumes of ensemble range of annual mean global mean temperature change relative to 1971-2000 for (top panel) 17 QUMP and 21 CMIP3 ensemble members under the A1B SRES emissions scenario; (bottom panel) 17 QUMP and 21 CMIP5 members for the A1B SRES and RCP6.0 scenarios respectively.*

In Figure 7, we compare projections of global temperature given by in CMIP3, CMIP5 and the 17-member QUMP ensemble of HadCM3 variants, one of the core experiments run for UKCP09 (see section 2b). The comparison is not straightforward, as the emission scenarios for UKCP09 and CMIP3 were based on the SRES family (Nakicenovic et al., 2000), whereas CMIP5 uses RCPs. The CMIP3 and QUMP simulations were forced by the A1B SRES scenario (medium scenario in UKCP09). The CMIP5 simulations in Figure 7 are those driven by RCP6, which is the closest in future radiative forcing to A1B. However, it is also the least well populated of the RCP scenarios in the CMIP5 archive. In the assessment of regional changes presented in section 3b below, we therefore often use RCP8.5 as well as or instead of RCP6.

Figure 7 shows that the range of future changes spanned by the QUMP ensemble is shifted to warmer values than the corresponding CMIP3 and CMIP5 ranges, more particularly in the latter case. This may be partly because specific forcing agents vary between the two emissions scenarios (for example CO_2 concentrations by the year 2100 are greater in A1B (703 ppm) than in RCP6 (667 ppm), although the total applied forcing by 2100 is similar (~6Wm^{-2} in both cases). Other potential reasons for differences between QUMP and CMIP5 would include different sampling of ECS or TCR values, and potentially different realisations of ocean heat uptake. For example, Figure 8 shows that the QUMP ensemble samples a wider of range of climate sensitivities than CMIP5, particularly at the upper end. However, this is not necessarily indicative of a problem with the QUMP ensemble, or HadCM3 - indeed the QUMP ensemble (and the larger ensemble of variants of the slab-ocean configuration (section 2b) from which it was derived) was explicitly designed to explore a plausible range of climate sensitivities (Collins et al, 2011), prior to the production of a posterior PDF of climate sensitivity. This PDF, also shown in Figure 8, accounts for results from CMIP3 models and the effects of observational constraints, as well as results from ensembles of perturbed HadCM3 variants (Sexton et al 2012). It agrees well with the CMIP5 range. The difference between the ranges implied by the PDF and the QUMP ensemble illustrates the impact of applying the full UKCP09 methodology to obtain projections, rather than relying solely on one of its core components.

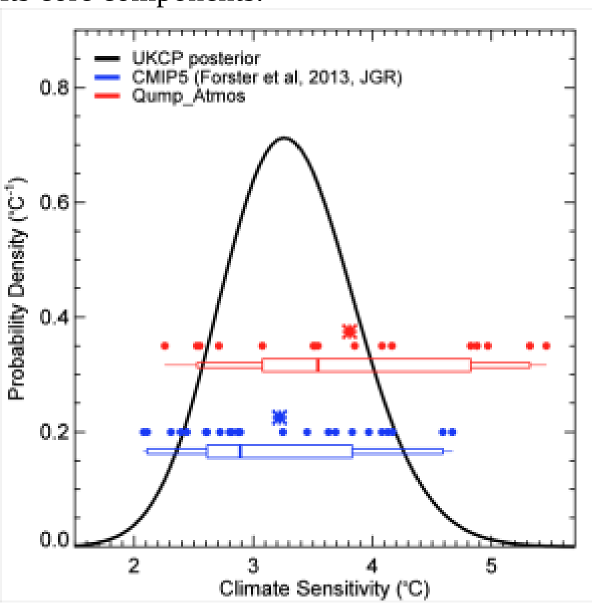

Figure 8. *Comparison of the posterior PDF of climate sensitivity used in UKCP09 (black curve) with climate sensitivities estimated for CMIP5 models (blue dots) (Forster et al, 2013) and the 17 QUMP variants (red dots). The stars show the mean and the box-whiskers show the median, full range and 5th, 25th, 50th, 75th, and 95th percentiles of the two climate model ensembles. The heights of the red and blue dots and box-whiskers are irrelevant and only the values on the x-axis need to be compared.*

By the same token, the probabilistic projections of global temperature rise given by the UKCP09 methodology (Harris et al, 2013) differ from those shown for the QUMP ensemble in Figure 7, because the methodology includes broader sampling of the HadCM3 parameter space, inclusion of CMIP3 model results, and the application of observational constraints. Harris et al find a 90% credible interval of 2.4-4.2ºC for changes for 2080-2100 relative to 1961-90 under CO_2 concentrations prescribed from the A1B scenario, lower than the range implied by the QUMP ensemble in isolation, and somewhat closer to (though still above) the CMIP5 range shown in Figure 7. It is not surprising that an unconstrained range of outcomes from an ensemble of opportunity such as CMIP5 is not identical to the probabilistic range obtained from the UKCP09 method.

Given that the range of future projections of global temperature change in CMIP5 is somewhat different from that of either the QUMP ensemble, or the probabilistic range derived from the full UKCP09 methodology, we would expect to see some divergence in envelopes of regional change, because global mean warming is a key driver of regional uncertainties (e.g. Harris et al, 2006), particularly for temperature-related variables. To overcome this complication in comparing raw data from QUMP A1B and CMIP5 RCP6 and RCP8.5 simulations in section 3b below, we instead analyse the normalised response of temperature and other variables, defined as the change in the relevant variable per unit change in global temperature. This allows us to focus specifically on potential differences in the regional patterns of response.

3b. Normalised responses

Figures 9-11 show various distributions of future change in decadal-averaged temperature and precipitation for UK sub-regions, normalised by globally averaged temperature change as described above. The UKCP09 PDFs represent the combined evidence available from HadCM3 perturbed parameter ensembles, CMIP3 simulations and observational constraints (as in Figure 8 above), whereas the percentiles shown for the various climate model ensembles are obtained simply by ranking the relevant simulations (shown individually as dots in Figure 9) in order of response, without any attempt to weight or exclude any of the results according to model performance metrics (e.g. of the type discussed in section 2a). As explained in section 3a, we would not necessarily expect the percentiles derived from any particular multimodel ensemble to overlap precisely with the UKCP09 results, however any major disparity between the CMIP5 results and UKCP09 (e.g. the occurrence of several CMIP5 simulations outside the 90% credible interval of UKCP09, or of a CMIP5 median response outside the 25-75% interval of UKCP09 results), might suggest a need to update the current UKCP09-based advice in the light of CMIP5 results. Results from several different emissions or forcing scenarios are shown, in order to demonstrate potential dependencies of the comparisons on which subset of models are included in a particular CMIP experiment, or on the impacts of different specifications of forcing agents. For example, comparing the 1% per year CO_2 increase experiments of QUMP, CMIP3 and CMIP5 provides a clean comparison of how the different ensembles respond to greenhouse gas forcing, while comparisons between the 1%, A1B and RCP experiments for a given type of ensemble gives an indication of the effects of including different forcing agents (particularly aerosols), and (in the case of CMIP3 and CMIP5) of including different subsets of the models contributing to the relevant multimodel archive.

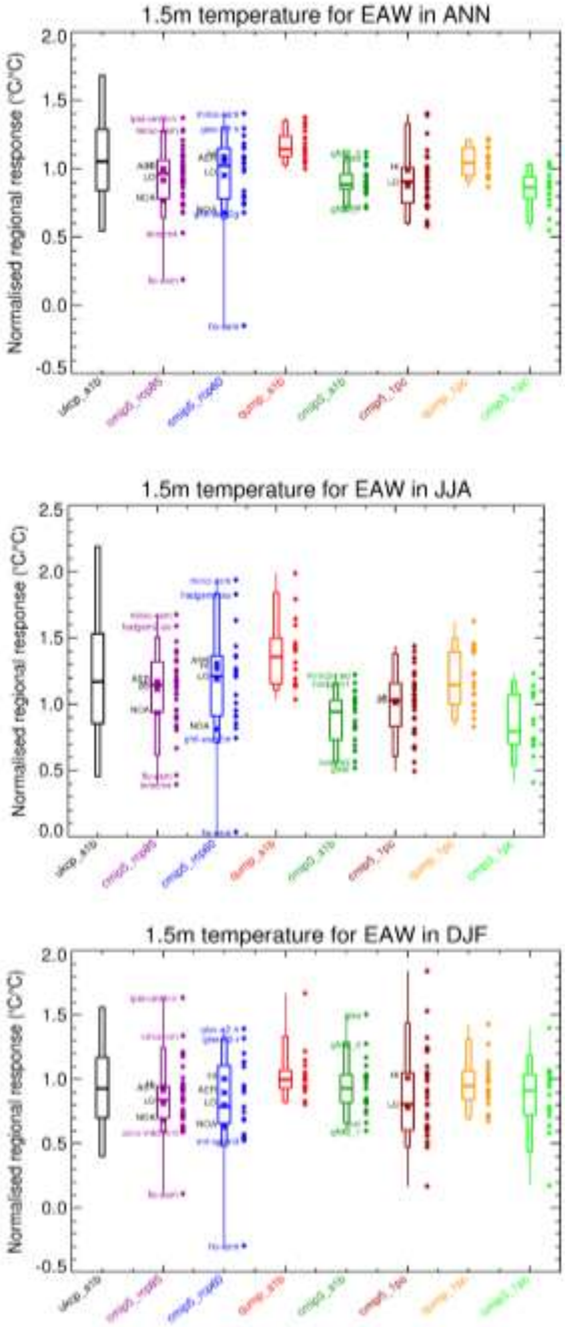

Figure 9. *Comparison of projected normalised future responses of England-Wales*[4] *temperature ($^{o}C/^{o}C$) for annual (top panel), summer (middle panel) and winter (bottom panel), for the 21st century. Results are provided for UKCP09, CMIP5 RCP8.5 (36 members) and RCP6 (21 members), and for the QUMP and CMIP3 simulations for A1B (QUMP_A1B and CMIP3_A1B). Boxes and lines show the 5th, 25th, 50th (thick solid), 75th, and 95th percentiles of the relevant distribution. Percentiles for the distributions based on specific multimodel ensembles are obtained simply by ranking the modelled outcomes and then interpolating to the required percentile. The UKCP09 percentiles are derived from probabilistic projections obtained by applying the statistical methodology of Sexton et al (2012) and Harris et al (2013). The whiskers show the full range of outcomes. Whiskers are not plotted for UKCP09, as the UKCP09 guidance recommended not using sampled outcomes beyond the 1% tails of the distribution. For a given multimodel archive (CMIP3*

[4] England and Wales are covered by a small number of grid boxes in current climate models, given a typical resolution of ~120km. As resolution also varies across the climate models, all model output was first regridded to the HadCM3 grid, using only land points from these models to produce values for HadCM3 land points.

or CMIP5), different scenarios are typically based on different subsets of the contributing international climate models. Results for QUMP, CMIP5, and CMIP3 simulations of the response to a 1% per annum increase in CO_2 (tagged _1pc) are also shown. Differences between 1% runs and RCP or SRES results indicate potential effects from changing aerosol emissions and other anthropogenic forcing agents, as well as sampling different multimodel subsets. "HI" and "LO" mark the average of the subset of multimodel members that are high-top or low-top respectively. "AER" and "NOA" mark the subset of multimodel ensemble members that respectively include or neglect aerosol-cloud interactions.

For surface air temperature in England and Wales (Figure 9), CMIP5 models give 5-95% ranges in normalised temperature response that invariably lie within the corresponding UKCP09 range, for annual mean, summer and winter changes (top, middle and lower panels). This indicates that CMIP5 results are consistent to first order with UKCP09 for this variable. However, this does not necessarily imply that updating UKCP09 to include CMIP5 would have no impact on the probabilistic projections. This is because the contribution of international models in the UKCP09 methodology is to provide sampling of the structural component of model uncertainties. In summer, for example, the QUMP_A1B model variants show a stronger normalised response than CMIP3, whilst UKCP09 covers both of these ranges. This demonstrates the benefit of including information from CMIP3 models in UKCP09. The potential impact of including CMIP5 information in a future update is best gauged by comparing the CMIP3 and CMIP5 distributions in Figure 9. In winter, the CMIP5 experiments show median responses slightly below the CMIP3 results, although the differences are relatively small. In summer, the CMIP5 ranges are shifted somewhat higher than their CMIP3 counterparts, and are closer to the ranges explored by the QUMP experiments. In Figure 9, we also note that for the annual and summer means, the 1% per year runs all show normalised responses shifted to somewhat lower values than their RCP and SRES counterparts for QUMP, CMIP3 and CMIP5. This suggests that aerosol forcing might be playing a significant role (through assumed future removal in SRES or RCP scenarios) in determining future UK changes. The CMIP5 models vary considerably in the levels of sophistication applied to their representations of aerosols. For example, the mean response amongst models which do not account for aerosol-cloud interactions in their RCP simulations (NOA in Figure 9) lies well below the mean response amongst models which do include these effects (AER).

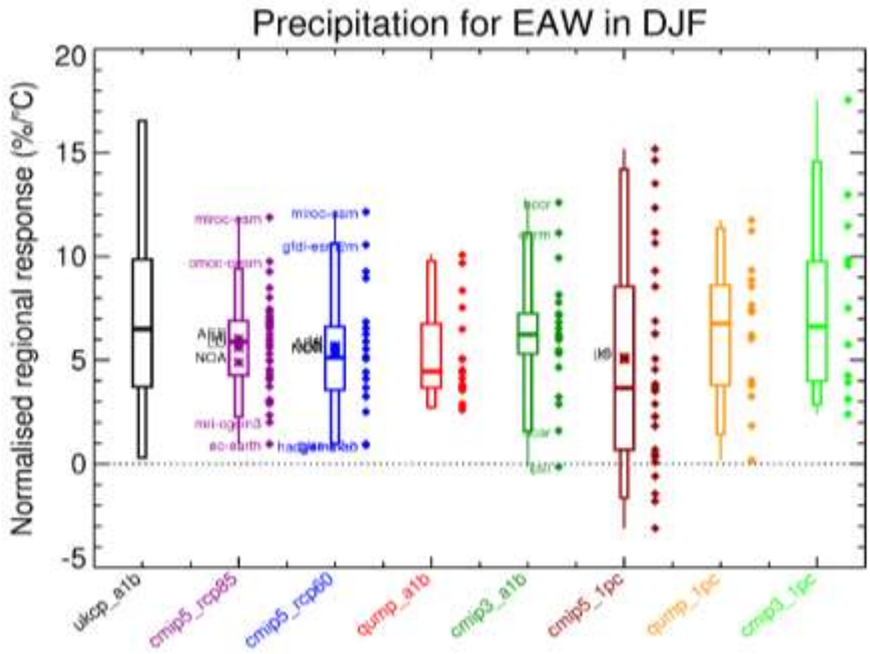

Figure 10. *As bottom panel in Figure 9 but for the normalised response in the percentage change in winter precipitation (%/°C) over England and Wales.*

For winter precipitation there is also good overlap between UKCP09 and CMIP5 (Figure 10). Almost all simulations continue to show a positive normalised response, consistent with the advice in UKCP09 of a long term trend towards wetter winters. For 1% experiments, CMIP5 does show a shift towards smaller values compared to CMIP3, but this difference is less clear when the CMIP5 RCP experiments are compared with the CMIP3 A1B simulations. We sub-sampled the CMIP5 results to assess whether the development of some high-top models with a fully resolved stratosphere in CMIP5 (mean responses labelled HI in Figure 10, cf mean responses labelled LO for low-top models), as well as the development of models with aerosol-cloud interactions (see above), might be responsible for any changes in the envelope of CMIP5 responses compared to earlier ensembles. Winter precipitation, in particular, has been reported to be sensitive to the introduction of a high-top. Both Scaife et al (2012) and Karpechko and Manzini (2012) have shown that extending the height of the atmosphere and increasing vertical resolution led to a equatorward shift of the storm track, offsetting the drying signal in winter over the Iberian peninsula and the Mediterranean. However, in the models considered in these studies (two by Scaife et al, and one by Karpechko and Manzini), the UK is near to the transition region from an enhanced wetter response to a drier response resulting from increased stratospheric resolution. In Scaife et al (2012) the two high-top models analysed have a slightly wetter UK response, whereas Karpechko and Manzini (2012) find a drier response. In Figure 10, we find no evidence of a systematic shift in winter precipitation response between the subsets of HI and LO CMIP5 models. Caution is required in interpreting this result: CMIP5 does not provide a clean test of the impact of enhanced stratospheric resolution, because few modelling centres ran parallel HI and LO simulations, so the subsets of HI and LO models will feature other differences potentially unrelated to vertical resolution. Also, the absence of a clear impact on winter precipitation over a small region such as the UK does not necessarily imply that impacts of stratosphere-troposphere interaction on circulation and precipitation will not be present over the broader Atlantic/European sector. Nevertheless, the results of Figure 10 are a salutary reminder that impacts of model changes reported in a small number of specific model studies need to be checked using a larger multi-model ensemble, in order to test the robustness of potential impacts in specific regions such as the UK.

For summer precipitation (Figure 11), the envelopes of normalised responses from CMIP5 favour future drying to a significantly lesser degree than the UKCP09 PDFs. Indeed, for the CMIP5 RCP6 scenario, over North England, 50% of the ensemble members have a positive percentage change in precipitation per degree of global warming. Interestingly, UKCP09 and CMIP5 1% pa runs show a similar percentage reduction in summer precipitation per degree of global warming, and the CMIP5 median is actually slightly drier than its CMIP3 counterpart. This experiment does not include changes in aerosol emissions, suggesting that future removal of aerosols in the SRES and RCP scenarios may play a role in partially offsetting a greenhouse-gas driven drying signal. In the RCP simulations, the average future response from CMIP5 models simulating aerosol-cloud interactions is shifted positive relative to that in models lacking aerosol-cloud interactions. This adds support to the above interpretation, although more detailed analysis would be needed to confirm it. Division of CMIP5 models between high-top and low-top also suggests that higher-top runs might lead to a greater likelihood of wetter summers in the future, although this only holds for RCP6. However, there is no indication in the recent literature that including stratosphere-troposphere dynamical interactions is necessarily likely to lead to a significant impact on summer climate over Europe (the focus has mainly been on impacts in winter, as described earlier), so it is possible that such differences may arise from sampling differences between the ensemble subsets unrelated to enhanced stratospheric resolution.

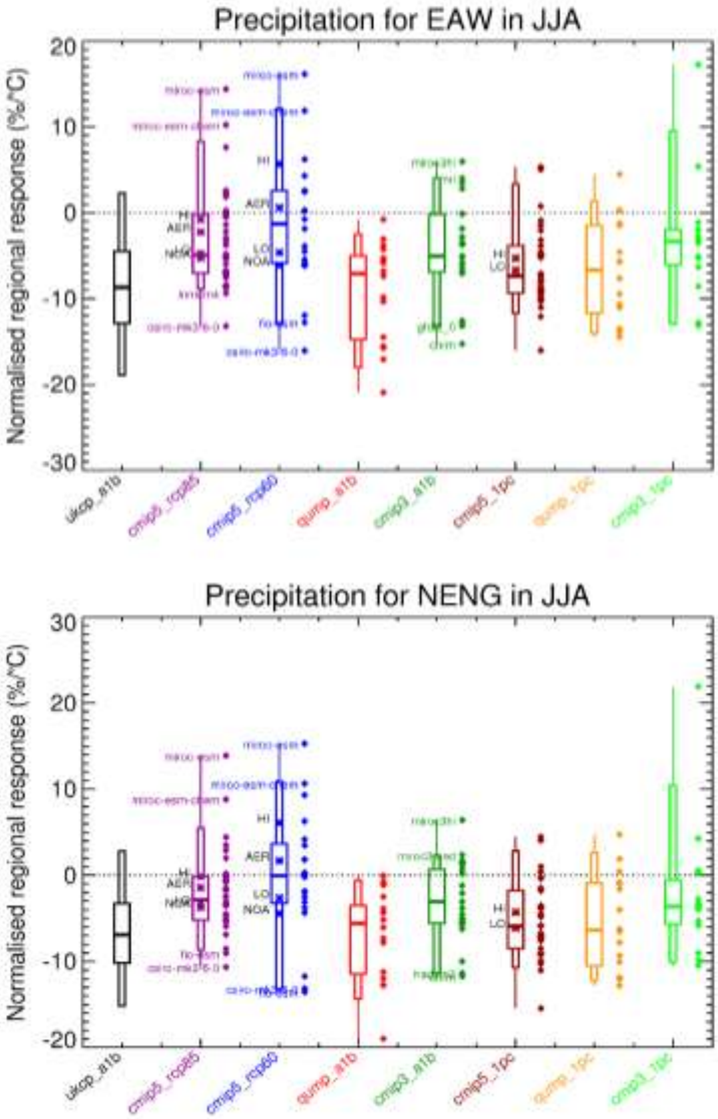

Figure 11. *As middle panel in Figure 9 but for the normalised response in the percentage change in winter precipitation (%/^{o}C) over England and Wales (top panel) and North England (bottom panel).*

Several additional factors contribute to the difficulty in drawing firm conclusions from apparent differences in ranges of responses such as those shown in Figure 11. First, as described in the Introduction, and emphasised above, the CMIP5 multimodel ensemble is an ad hoc sample of models from which it is hard to find robust statistical relationships that might isolate specific drivers of future change, due to the presence of multiple structural differences in the construction of individual ensemble members. Also, although the CMIP5 models have all undergone extensive testing to provide plausible simulations of climate, it is challenging to apply formal observational constraints in order to weight different members (Knutti et al, 2010), since multimodel ensembles lack the systematic construction of the perturbed parameter ensembles built for UKCP09, which facilitated the use of a statistical framework (Rougier, 2007) capable of supporting formal estimates of the relative likelihood of different ensemble members.

Nevertheless, results from multimodel ensembles do show clear differences in model performance, e.g. as seen in Figures 1-4. One alternative strategy to deriving some complex weighting system for multimodel ensemble members is to screen out a few truly poor performers. McSweeney et al (2014) describe such an approach, seeking to identify subjectively a subset of CMIP5 models suitable to derive plausible regional climate scenarios for Europe, Africa, and South East Asia. They excluded three models, including the two MIROC-ESM models, which gave particularly poor simulations of the summer circulation over the North Atlantic-Europe region and the Indian summer monsoon circulation. Interestingly, these are precisely the two models that produce the largest wet summer response over the UK in CMIP5 (see Figure 10).

Another issue relates to the degree of independence between different models. Some modelling centres contributed more than one configuration of their model to the CMIP5 archive (the two MIROC earth system models referred to above being one example), which would typically have shared the majority of their structural components. Screening the CMIP5 ensemble to remove near-duplicate results (not attempted here) could potentially affect some of the results shown in Figures 9-11. Taking all these cautionary notes into consideration, it is reasonable to conclude that more CMIP5 members (relative to CMIP3 and UKCP09) suggest that typical summers might become wetter than or remain similar to the 1961-90 average, however the extent of any change to the likelihood of wetter summers cannot currently be quantified. For CCRA2, and future use of UKCP09, we recommend that CMIP5 evidence should be considered alongside UKCP09 in assessing summer rainfall changes, and that use of UKCP09 in isolation would now over-estimate the risk of significant reductions in summer rainfall implied by evidence available from the latest climate models.

We assessed similar plots to Figures 9-11 for temperature and precipitation, two key variables for users, for all UK regions in all seasons. Based on a criterion that the CMIP5 median should lie within the 25-75 percentile range of UKCP09, we found in most cases a spread of CMIP5 results that had sufficient overlap to indicate that continued use of UKCP09 advice is justified, pending development of updated scenarios. We also assessed normalised future changes in downward shortwave radiation (SW_Down), given the potential impacts of aerosol removal found in CMIP5 results for summer temperature and precipitation. The MIROC-ESM and MIROC-ESM-CHEM models give very large increases compared to other models, and were excluded on the basis of the concerns over credibility cited above. Amongst the remaining CMIP5 models, those that include aerosol-cloud interactions give larger future increases in SW-Down than those that do not, by typically 4-6$Wm^{-2}K^{-1}$. Overall, the CMIP5 ensemble shows a wide spread of changes broadly compatible with UKCP09. Over SE England the CMIP5 distribution is shifted to higher values by ~2$Wm^{-2}K^{-1}$ compared with UKCP09, though this is modest compared with the 90% credible interval from UKCP09 of ~4-15 $Wm^{-2}K^{-1}$.

Another variable of major interest to users is surface wind speed. Probabilistic projections for this variable were provided as an additional UKCP09 product (Sexton and Murphy, 2010), and showed a broad range of future outcomes with little evidence of clear signals for increases or reduction, reflecting large uncertainties in simulated future changes in the regional circulation, especially at the UK scale. A like-for-like comparison with CMIP5 was not possible in the present report, as the UKCP09 PDFs were provided at a 25km scale, and included downscaling information from the regional climate model simulations run for UKCP09. Corresponding versions on the coarser HadCM3 grid used for the present comparisons were not made. However, inspection of wind speed changes from CMIP5 models also reveals no evidence of consistent signals of increase or reduction, essentially consistent with the UKCP09 information.

Although our main concern in this report relates to assessment of differences between the UKCP09 and CMIP5 projections over the UK, it is also important to evaluate this issue across the whole globe. This provides a broader view of the credibility of UKCP09 in the light of CMIP5, and also informs the extent to which the UKCP09 methodology (or a suitable future update to it) could continue to provide a basis for provision of information on international changes, for projects such as the compilation of worldwide impacts carried out for the UK's Department for Energy and Climate Change (DECC) in 2011 (see http://www.metoffice.gov.uk/climate-change/policy-relevant/obs-projections-impacts).

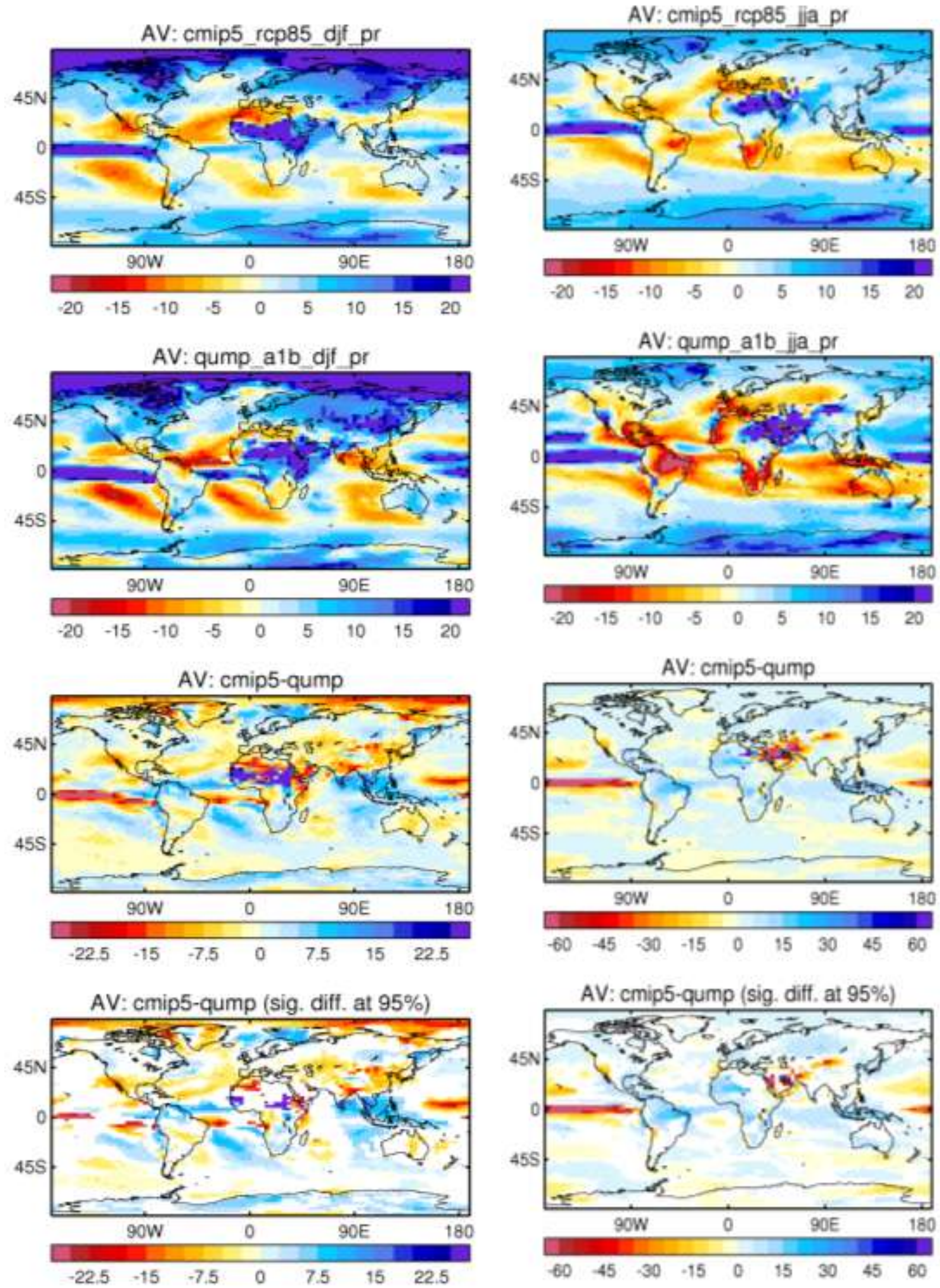

Figure 12. *Ensemble mean spatial pattern of the normalised response in percentage future change of DJF (left panels) and JJA (right panels) mean precipitation for CMIP5 and QUMP simulations (top two panels). The bottom panels show the difference between the CMIP5 and QUMP patterns, the lowest panel only displaying values where the difference is significant at the 5% level based on a two-tailed bootstrap test in which uncertainties in the difference between the two ensemble means are estimated by creating 2000 synthetic samples with replacement, for each ensemble.*

For DJF precipitation changes (see Figure 12 left panels), the ensemble mean spatial patterns for QUMP and CMIP5 (taken from 36 RCP8.5 simulations) look very similar, the only obvious difference in sign of response occurring over South East Asia. The bottom panel shows that there are actually regional differences in the *magnitude* of the normalised response over larger parts of

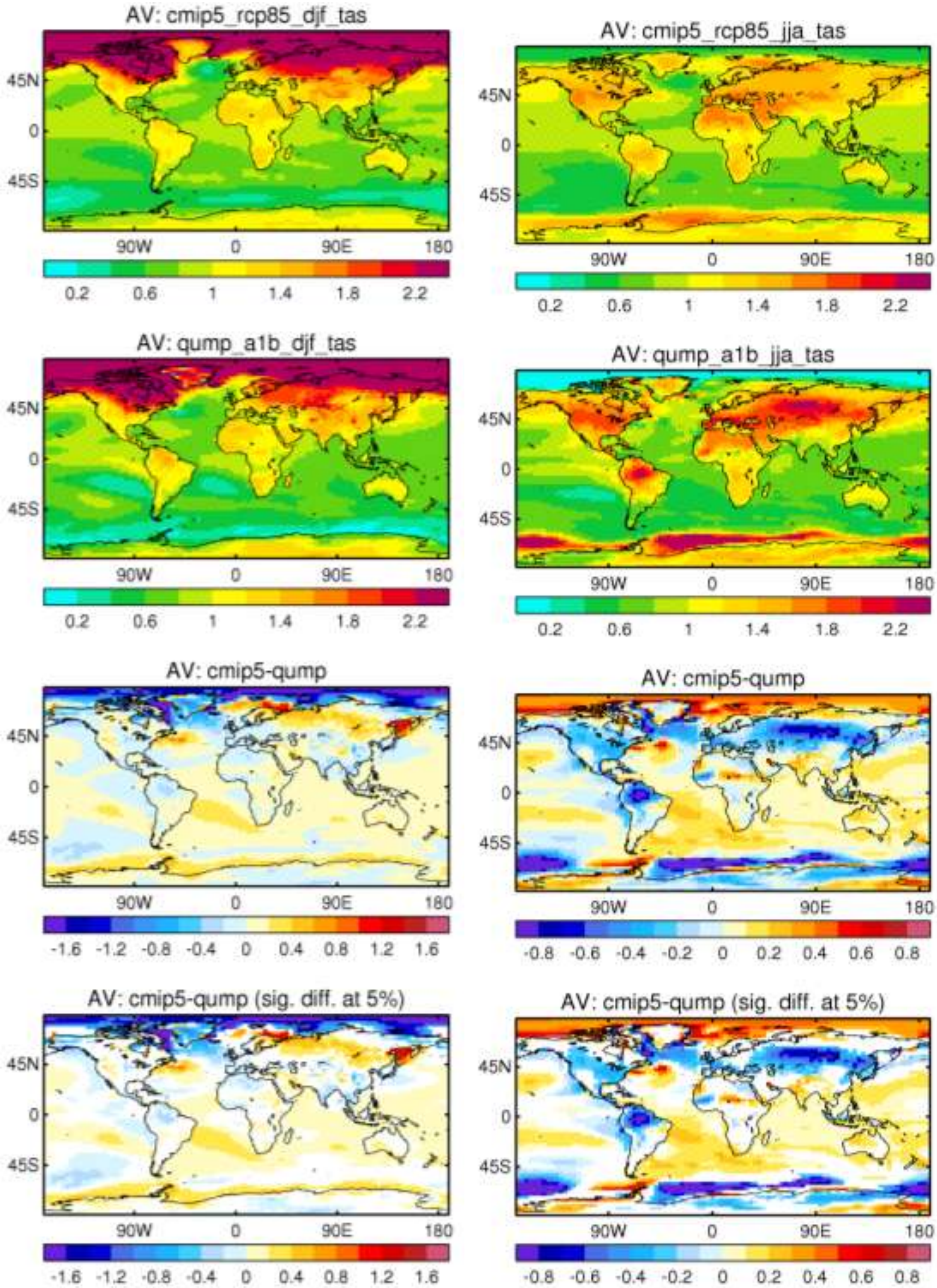

Figure 13. *As Figure 12 but for 1.5m air temperature.*

the globe. Such differences do not necessarily indicate a shortcoming in the QUMP (or indeed CMIP5) simulations, but do indicate regions where inclusion of CMIP5 information via a future update to the UKCP09 methodology might be appropriate (dependent also on differences between CMIP3 and CMIP5 results, not shown in Figure 12). However, the clearest, most spatially coherent difference in Figure 12 is a drier response over China and wetter response over South East Asia in CMIP5, compared to QUMP. This difference has been previously identified by McSweeney et al (2012) in comparison with CMIP3 models, where it was suggested that the use of flux adjustments in QUMP to mitigate SST biases may have caused the difference in the SE Asia region. In support of this suggestion, the normalised response in CMIP5's AMIPfuture[5] experiment are similar to CMIP5 over China but are more like QUMP over South East Asia. For JJA precipitation (right panels), regions of drying tend to show a stronger normalised response in QUMP, and the drying region over Europe extends further over into Eurasia. Otherwise the spatial variations in the sign of the normalised response look similar.

For DJF 1.5m temperature (Figure 13 left panels), the ensemble-mean normalised response looks broadly similar across CMIP5 and QUMP (top two panels). However, there are statistically significant differences at the 5% level, most notably CMIP5 has a stronger response over parts of northern Eurasia, yet a weaker response over north Atlantic between Greenland and UK. However, the difference over UK was also found in comparison with CMIP3, and this systematic difference between QUMP and CMIP3 was accounted for in the UKCP09 method (Murphy et al 2009; Harris et al 2013). For JJA 1.5m temperature (right panels), QUMP has a stronger ensemble-mean response over NH midlatitude land areas, possibly reflecting dry historical soil conditions (Murphy et al, 2014), making it easier under a warming climate for soil moisture to drop below the level where the cooling effect of evaporation becomes limited. However, uncertainties in observed soil moisture make it difficult to assess the relative credibility of different model simulations in this respect. Again, this difference also existed between QUMP and CMIP3 simulations, and was accounted for by combining both sources of modelling information in the UKCP09 method.

Finally, in Figure 14 we extend the international comparisons to consider uncertainties in worldwide regional changes, comparing the QUMP ensemble against the same set of 36 CMIP5 coupled ocean-atmosphere models used to construct Figures 12 and 13, for the same set of country regions assessed in the DECC worldwide impacts project referred to above. In Figure 14, the boxes, bars and whiskers denote the 5th, 25th, 50th, 75th and 95th percentiles of the relevant ensemble distributions. While there is substantial overlap between the two ensembles in many cases, some substantial regional differences are also apparent, demonstrating both the value of considering different sources of model information, and the need to evaluate further the extent to which inclusion of CMIP5 information might change probabilistic projections obtained by combining information from HadCM3 perturbed parameter ensembles with CMIP3 results (Harris et al, 2013). For example, the shift to a less dry summer precipitation response and a smaller warming found for the UK is also seen in the comparisons between CMIP5 and QUMP envelopes of change for other European countries. Again, the significance of such differences for potential future updates to the Harris et al. projections depends on differences between CMIP5 and CMIP3 results. These are not shown here; however Murphy et al (2014) find worldwide contrasts between the QUMP ensemble and CMIP3 models which are qualitatively similar in many cases to the CMIP5/QUMP contrasts shown in Figure 14.

[5] AMIPfuture is a 10-year AMIP run with prescribed SSTs made up from the standard AMIP SSTs plus a warming pattern estimated from the ensemble mean response of CMIP5 in 1%p.a. runs at four times CO_2 concentrations.

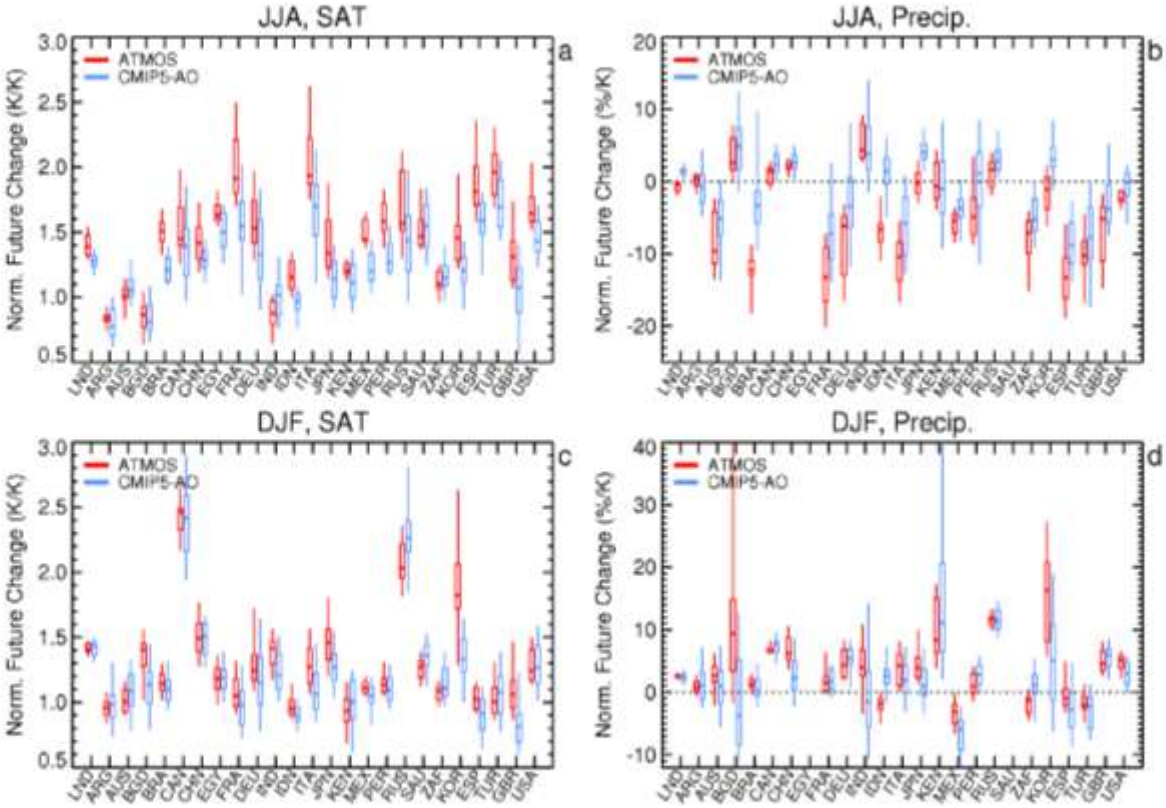

Figure 14. *Uncertainties in future regional responses, normalised by future global surface air temperature changes, for a selection of G20 and other nations. Results are shown for DJF (top panels) and JJA (bottom panels) and for surface air temperature (left panels) and precipitation (right panels). Results from the 17QUMP simulations driven by historical and A1B concentrations are shown in red, and from 36 CMIP5 atmosphere-ocean simulations forced by the RCP85 concentration pathway in blue. The whiskers, boxes and horizontal bars show the 5th, 25th, 50th, 75th and 95th percentiles of the relevant ensemble distribution of changes. Normalised future changes are estimated by applying linear regression to annual anomalies for the period 1990-2100 calculated with respect to a1980-1999 baseline. Precipitation changes are in units of percentage future change per degree of global SAT change. However values are omitted for Egypt and Saudi Arabia because baseline precipitation values simulated for 1980-99 are very small, leading to noisy distributions of percentage future change.*

4. Key drivers of UK climate

Here, we present a qualitative assessment of some key drivers of climate over UK (other than global mean temperature, covered earlier in section 3a) based on a survey of a wide range of expert opinion canvassed from Met Office Hadley Centre and (in the case of storm tracks) the Department of Meteorology at the University of Reading. The material in this section is not a comprehensive summary of all drivers of UK variability and change, but highlights a number of topics featured heavily in recent literature, focusing mainly on large-scale features of the climate system that affect variability over the UK and Europe. We include comments on both historical simulation and projected future changes, comparing some of the modelling behind UKCP09 against CMIP5 results.

Atlantic meridional overturning circulation (AMOC)

The AMOC provides a substantial northward transport of heat in the North Atlantic Ocean, which exerts a major influence on the atmosphere through transfer of heat at higher latitudes. Natural variability or externally-forced changes in the strength of the circulation affect surface temperatures, precipitation amounts and storm track characteristics over the North Atlantic and Europe, including the UK (Srokosz et al., 2012). Observations of the AMOC have improved dramatically in the last few years due to the availability of data from the RAPID observing array at 26.5N, which was deployed in 2004. Roberts et al (2013) show that the AMOC in HadCM3 and HadGEM2-ES is too shallow and too weak compared to observations, but some CMIP5 models do a much better job of capturing the magnitude and depth structure of the AMOC at 26.5N. Overflow transports across the Greenland-Scotland ridge are an important source of the dense waters that make up the deep return flow of the AMOC. Overflow transports in HadCM3 and HadGEM2-ES are too strong in the Denmark Strait and too weak in the Faroe Bank Channel. Again, overflow transports in some CMIP5 models (e.g. CCSM4) compare better with observations.

Although some CMIP5 models show some improvements in simulating the AMOC, there is little difference in projected future changes in its strength. Climate model projections used in UKCIP02, UKCP09, and CMIP3 suggested that the AMOC will weaken in response to increasing greenhouse gases. In the UKCP09 science report (Murphy et al., 2009), Fig A5.3 shows 10-30% reductions in the strength of AMOC, at the time of doubling CO_2 concentrations (70 years) in simulations driven by a 1% per annum increase. In CMIP5 simulations, the best estimate decrease by 2100 is about 20–30% for the RCP4.5 scenario and 36–44% for the RCP8.5 scenario, similar to CMIP3 results. On 26th September 2013, the Daily Telegraph reported that this would amount to a cooling of 1^{o}C by the end of the 21st century, and that “Scientists warn that the resulting cooling would mask the impacts of global warming on the country, but play havoc with the weather.” In fact, the lower ends (defined by the 5th percentile) of the CMIP5 responses for England and Wales annual temperature in the 2080s (relative to 1961-90) indicate a warming of about 1^{o}C and 2^{o}C for the RCP6 and RCP8.5 scenarios respectively, while the corresponding median responses amount to ~3^{o}C and ~4^{o}C. Therefore, the cooling from a reduction in AMOC strength only partly offsets the warming. This is also the case in the QUMP simulations driven by the A1B scenario.

Arctic sea ice and variability

Sea ice is a product of atmosphere-ocean interaction. There are a number of ways in which sea ice is influenced by and interacts with the atmosphere and ocean, and some of these feedbacks are still poorly quantified. Although it is also important to look at sea ice in terms of thickness, volume, and regional distribution, for this study we limit ourselves to the rate of reduction of total Arctic sea ice extent in September and January. Changes in the minimum Arctic sea ice extent, which occurs in September, have received a lot of attention recently. This is because the period 2007-2012 saw record low amounts, though there has been a recovery in 2013. In CMIP5, Massonnet et al (2012) show that the reduction in sea ice is better captured in those models that do not over-estimate the baseline sea ice extent. QUMP is in good agreement with the observed baseline values (not shown).

Figure 15 compares percentage change of Arctic sea ice extent in September in the A1B simulations from QUMP against CMIP5 simulations using RCP6.0, this being the RCP scenario closest in future forcing to A1B. QUMP has a slower downward trend than CMIP5 up to 2040 after which the means of the two ensembles agree very well. QUMP also has a narrower spread until 2030, possibly because the use of flux adjustments may have constrained historical baseline values of sea ice extent somewhat. Yet, both ensembles capture the observed year-to-year variations in their envelope. For January, the observed downward trend is slightly slower than those of CMIP5 and QUMP, with the observed values approaching the upper parts of the envelopes of both ensembles. From 2020, QUMP shows a more rapid decline in sea ice extent for January.

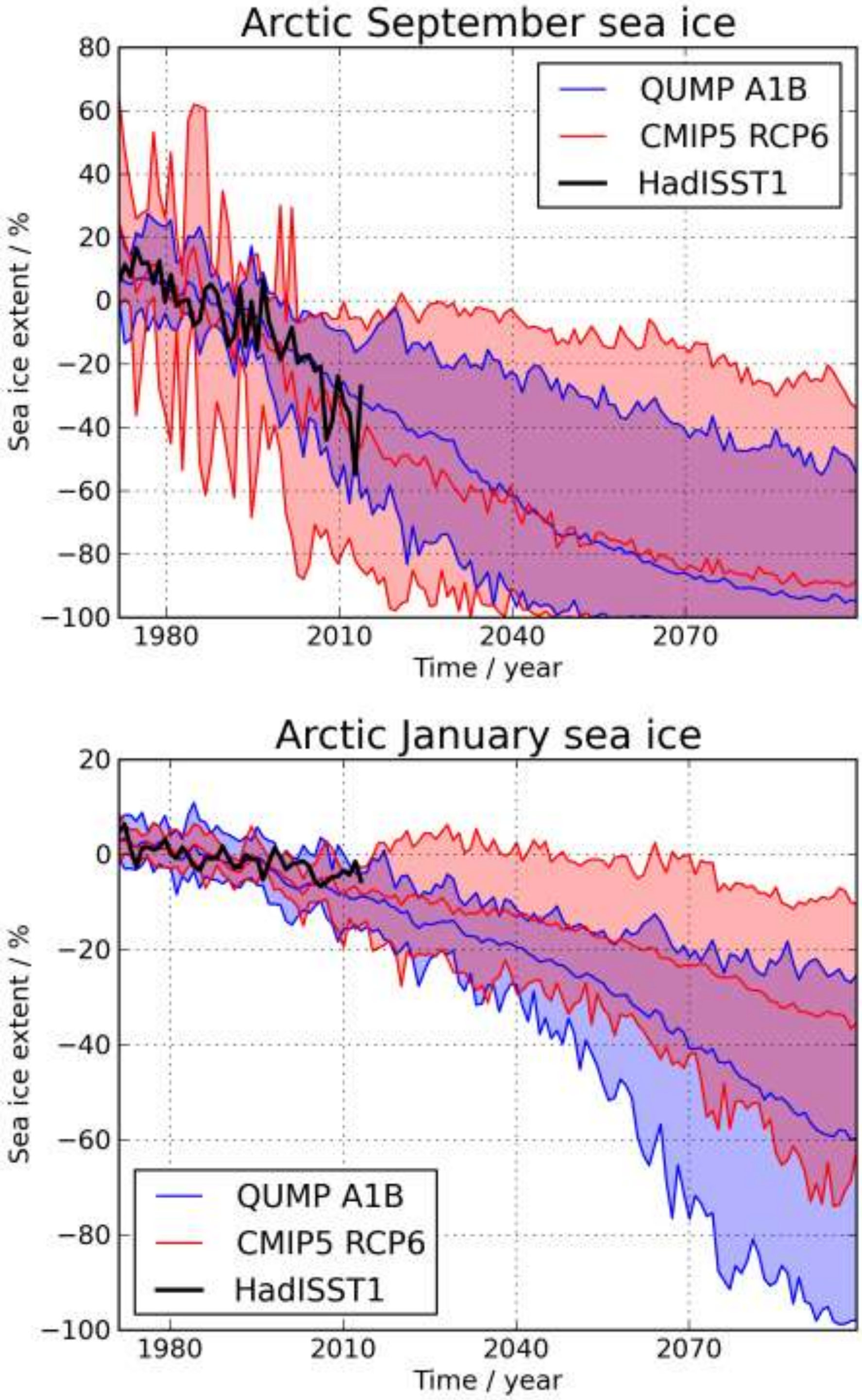

Figure 15. *Envelopes of percentage change in Arctic sea ice cover relative to 1971-2000 for September, the month of minimum sea ice extent (top panel), and January (bottom panel), for the QUMP ensemble (blue), and 17 members of the CMIP5 RCP6 ensemble (red). The middle blue and red lines show the averages of the two ensembles. Black line shows observed values from HadISST1.1.*

As well as being a key indicator of climate change, Arctic sea ice extent has recently received further attention relating to potential links between its year-to-year variability and changes in circulation over the North Atlantic and Europe (e.g. Overland and Wang, 2010), or related impacts such as the risk of cold Eurasian winters (Petoukhov and Semenov, 2010). For instance, Screen et al (2013) use two climate models to test the atmospheric response to recent observed changes in sea ice extent, and find that the only robust signal is a weak tendency towards negative NAO in

November and December, which is linked with colder weather over Europe. Other than that, there is little consistency across existing studies regarding the effect of sea ice extent changes on the NAO.

Storm tracks

Models are able to capture the general characteristics of storm tracks and extratropical cyclones, and there is some evidence of improvement since the AR4. Storm track biases in the North Atlantic have improved slightly, but models still produce a storm track that is too zonal. For example, Figure A3.6 of Murphy et al (2009) showed that the latitude of maximum storm track strength was typically displaced a few degrees to the south of observations in the QUMP ensemble, while the southerly displacement was somewhat larger in most CMIP3 models. Therefore, climate models tend to have more storms passing over the British Isles than observed. This error is apparent in both the CMIP3 and CMIP5 multi-model means, although the error is reduced in the CMIP5 models (Zappa et al 2013a). CMIP5 provided a greater wealth of data for storm track analysis, and Zappa et al (2013a) use this to show that cyclone intensities are typically underestimated. However, they also show that there is a large spread in the performance of individual climate models and that a few CMIP5 models show a good representation of both the location and intensity of extratropical storms, with the higher-resolution models being the best.

Projections for the end of the 21st century are similar in CMIP3 and CMIP5, and suggest a modest reduction in wintertime storm activity over Scandinavia and the Mediterranean and a slight increase over North Western Europe (Harvey et al 2012; Harvey et al 2013; Chang et al 2012, Zappa et al 2013b). The largest differences between CMIP3 and CMIP5 projections are associated with a reduction of storm activity over the ice-edge regions of the Arctic Ocean (Harvey et al., 2012).

Blocking

Blocking occurs when a high pressure system persists for several days or more, disrupting the prevailing mid-latitude westerly winds and storm track, causing a local reversal of zonal flow and in some locations a greater component of the wind in a north-south direction. Such systems are important for the UK as they are often associated with more extreme weather, such as very cold winters or heatwaves and droughts in summer. Climate models in the past have universally underestimated the occurrence of blocking, in particular in the Euro-Atlantic sector (Scaife et al, 2010).

Improvements in physical parameterization schemes and increases in horizontal (Matsueda, 2009) and vertical resolution (Anstey et al, 2012) can lead to improvements in the abilities of climate models to simulate blocking. Enhanced resolution has improved the simulation of blocking due to better representation of orography and atmospheric dynamics (Woollings et al, 2010), and to reduced ocean surface temperature errors in the extra tropics (Scaife et al, 2011). However, most of the CMIP5 models still significantly underestimate the frequency of winter Euro-Atlantic blocking (Anstey et al, 2012; Masato et al, 2012; Dunn-Sigouin and Son, 2013).

Murphy et al (2009) estimated blocking frequency using the Pelly and Hoskins (2003) variable-latitude index, and found that the QUMP model variants underestimated the frequency of blocks over the UK in winter by typically 30% (50 to 0% for individual QUMP members). Since CMIP3, analyses of blocking (Anstey et al 2013, Dunn-Sigouin and Son 2013, Masato et al 2012) have become more sophisticated and only Anstey et al provide a similar metric. Amongst CMIP5 models, Anstey et al (2013) conclude that all models underestimate observed the blocking frequency over Europe, and that deficits of ~50% are common, suggesting a similar distribution of biases to that found in the QUMP ensemble.

Most CMIP3 models simulated a reduction in future winter blocking frequency over the

European/North Atlantic sector in response to increases in greenhouse gases (Barnes et al., 2012). Such changes need to be interpreted with caution (Woollings, 2010), because they can arise either from a simple change in the multiyear mean background flow, or from changes in variability relative to the mean state associated with the dynamics of blocking.

The UKCP09 probability distributions accounted for modelling uncertainty represented in the set of climate models available at the time of their construction, and did not include uncertainty caused by structural errors common to all climate models. This has been a source of criticism of UKCP09, although the science report did discuss this issue (Murphy et al., 2009), classifying the unknown effects of such common errors as an unquantifiable component of uncertainty beyond the scope of the understanding and capabilities summarised in the probability distributions. The underestimation of blocking is a prevalent example: Future advice on 21st century blocking may change, if an ensemble of models can be developed which lacks the systematic biases seen in QUMP, CMIP3 and CMIP5 models. Until such an ensemble is available, however, it is difficult to anticipate to what extent the projections could change. We are not aware, for example, of any studies that attempt to quantify the impact of common structural errors in blocking on projections. We suggest that understanding the effects of common model errors on projections is a gap in our knowledge that would be challenging to overcome, but if progress could be made, this would be of benefit in enhancing the advice available to climate adaptation practitioners. Even if the effects of such common biases (or of missing earth system feedback processes) cannot be quantified, a greater focus on the provision of additional qualitative advice would be of benefit, to help users understand the degree of confidence that should be placed in the quantitative results.

5. Summary and recommendations

The purpose of this technical report was to evaluate the probabilistic projections of future climate change over land that from a core component of the current UKCP09 climate scenarios. Other components of UKCP09, such as the observed trends report, the marine and coastal scenarios, the weather generator tool and a subsidiary set of spatially coherent land projections, are not assessed here. Information on these is available from the UKCP09 website (see http://ukclimateprojections.metoffice.gov.uk/). In this report, the probabilistic land projections are assessed in the light of results from the recent ensemble of international climate models contributed to the CMIP5 archive and the fifth IPCC WG1 assessment (AR5). There were three parts to the evaluation of the probabilistic projections: assessment of historical model performance using long-term climatological averages of global scale fields and conclusions from the model evaluation chapter of AR5; comparison of future projections; review of model performance in simulating several large scale phenomena that influence climate variability over the UK, and how these might affect UK climate projections.

The AR5 concluded that generally models have improved since CMIP3, being enhanced through increases in resolution both horizontally and vertically (in the latter case, some CMIP5 models include a better representation of the whole stratosphere), and through the addition of processes, e.g. aerosol-cloud interactions. However the general conclusion depends on the variables under consideration. No instances are found where the CMIP5 ensemble as a whole performs significantly worse than the CMIP3 ensemble; for some variables the performance levels are similar, and for others the CMIP5 ensemble performs better. In terms of emergent properties of climate, there have been some improvements, mainly in the representation of certain aspects of variability and extremes. This is reflected in section 4, based on a consultation of experts at the Met Office and the University of Reading, who cite improvements in the model performance of the Atlantic Meridional Overturning Circulation, blocking, and storm tracks, which are all key drivers of climate variability over the UK, and are also potential drivers of future climate change. Despite these improvements, there are still deficiencies in the representation of these phenomena; for example, winter blocking frequency is still underestimated in CMIP5 models. More generally, the IPCC model evaluation chapter concludes that there are still some major challenges for climate modelling; one specific example that affects the UK is the timing of moist convection during days in summer.

The quantitative comparison of model performance between UKCP09 and CMIP5 (see Figures 1-5) is based on assessment of an ensemble of perturbed coupled ocean-atmosphere variants of the HadCM3 model (QUMP), a key underpinning component of the probabilistic projections, against CMIP5 coupled ocean-atmosphere models. The results show that the QUMP simulations (a key underpinning component of the UKCP09 modelling strategy) are still competitive in terms of their ability to simulate global climatological fields, with the exception of summer surface air temperature over the Northern Hemisphere continents. The relatively good performance of HadCM3 and its variants is mainly because HadCM3 is one of the better CMIP3 models (see Figures. 4-5). The use of flux adjustments, which act to reduce SST biases and related errors in the model, also improves the climatological performance of the QUMP simulations, particularly over the oceans.

We also compared quantitatively future projections from the UKCP09 probability distributions, CMIP3 and (in particular) CMIP5 multi-model simulations. However, the analysis is not straightforward. One issue is that the UKCP09 distributions are observationally constrained, whereas the CMIP3 or CMIP5 results are not. A second issue relates to limitations of the multimodel ensembles, in which the limited sample size exposes the analysis to possible effects of poor models, outliers, and pairs of interdependent models provided by the same climate modelling centre and sharing many structural modelling choices. Also the multimodel design is *ad hoc*, which makes it hard to determine underlying causes of any changes to the projections in CMIP5, relative to UKCP09. For these reasons, we would not expect their results to match precisely. We also note that UKCP09 contains information from CMIP3 simulations, which contributed an estimate of the structural component of modelling uncertainty. In assessing the potential impacts of including CMIP5 information in a future update to UKCP09, a key question therefore, is whether there are differences between the two multimodel ensembles, CMIP3 and CMIP5. Given these issues, the evidence presented in this report should be regarded as a survey aimed at identifying any major disparities between UKCP09 and CMIP5, rather than providing a detailed view of how the UK projections might change with inclusion of CMIP5 information.

Overall, the analysis reveals no first order shifts in CMIP5 future projections relative to CMIP3 and UKCP09 that would render the UKCP09 probabilistic projections out-of-date, notwithstanding the improvements in climate modelling noted above. Given the continued presence of common biases in CMIP5 models, there is no immediate prospect of resolving the limitation of UKCP09 that the projected distributions of future change do not account for the (currently) unknown implications of systematic model errors. Therefore, we suggest that more research needs to be done to ascertain which, if any, of the structural errors in the representation of processes already included in climate models, including those in CMIP5, could have an appreciable impact on future projections. We note, however, that this is likely to be a challenging task. The investment in model development, at the Met Office Hadley Centre and other worldwide modelling institutes, is an ongoing exercise aimed at addressing these deficiencies. However, it is difficult to place a timescale on the removal of such biases.

Some processes are still missing from current models, which further emphasises the status of results such as UKCP09 and CMIP5 as projections conditional on the understanding included in the set of models used to produce them. Interactive representation of the nitrogen cycle and permafrost-carbon feedbacks are two examples of missing processes discussed in this report. Representation of dynamical stratosphere-troposphere interactions in high-top models is one example of a process newly included in some CMIP5 models relative to CMIP3. However, this report shows that there also needs to be evidence from a larger ensemble of such models before the impact on future projections can robustly be assessed. It also shows an ongoing need for clear communication and advice to users when a major modelling advance arises.

Having said that, the CMIP5 projections for summer rainfall differ enough that this warrants an update to advice on how users should handle this variable. CMIP5 members (relative to CMIP3 and UKCP09) suggest that future summer climate might become drier to a lesser extent, while some members suggest little change relative to the historical 1961-90 average, and a few suggest typical summers might become wetter. However the extent of any change to the likelihood of wetter summers cannot currently be quantified, due to the issues of interpretation of the CMIP5 ensemble discussed above. It is possible that enhanced representation of aerosol-cloud interactions in some members might be a driver of the reduced probability of future drying in the CMIP5 ensemble, however more work is needed to investigate this further.

For CCRA2 and other adaptation advice, we therefore recommend that CMIP5 evidence should be considered alongside the UKCP09 probabilistic projections in assessing summer rainfall changes, and that use of UKCP09 in isolation would now over-estimate the risk of significant reductions in summer rainfall implied by evidence available from the latest climate models.

Acknowledgements

We would like to thank the following people for providing the wider perspective of how HadCM3 and CMIP3 compared with CMIP5. From the Hadley Centre: Adam Scaife, Chris Jones, Chris Roberts, Laura Jackson, Gareth Jones, Mark Webb, Ben Booth, Dave Rowell, Simon Brown, Neal Butchart, Mark McCarthy, Yoko Tsushima. From Reading University: Len Shaffrey, Ben Harvey, Tim Woollings. We would particularly like to thank Chris Roberts and Len Shaffrey, who contributed paragraphs of text.

References

Aldrin, M., M. Holden, P. Guttorp, R. B. Skeie, G. Myhre, and T. K. Berntsen, 2012: Bayesian estimation of climate sensitivity based on a simple climate model fitted to observations of hemispheric temperatures and global ocean heat content. Environmetrics, 23, 253–271. doi:10.1002/env.2140.

Allan, R., and T. Ansell, 2006: A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850-220. J. Climate, 19, 5816-5842.

Anstey, J. A., P. Davini, L. J. Gray, T. J. Woollings, N. Butchart, C. Cagnazzo, B. Christiansen, S. C. Hardiman, S. M. Osprey, and S. T. Yang, 2013: Multi-model analysis of Northern Hemisphere winter blocking: Model biases and the role of resolution. J. Geophys. Res., 118, 3956–3971. doi:10.1002/jgrd.50231.

Ashfaq, M., C. B. Skinner, and N. S. Diffenbaugh, 2011: Influence of SST biases on future climate change projections. Clim. Dyn., 36, 1303–1319. doi:10.1007/s00382-010-0875-2.

Barnes, E.A., J. Slingo, and T. Woollings, 2012: A methodology for the comparison of blocking climatologies across indices, models and climate scenarios. Clim. Dyn, 38, 2467-2481.

Bonan, G. B. and S. Levis, 2010: Quantifying carbon-nitrogen feedbacks in the Community Land Model (CLM4). Geophys. Res. Lett., 37. doi:10.1029/2010GL042430.

Booth, B. B. B., D. Bernie, D. McNeall, E. Hawkins, J. Caesar, C. Boulton, P. Friedlingstein, and D. M. H. Sexton, 2013: Scenario and modelling uncertainty in global mean temperature change derived from emission-driven global climate models. Earth Sys. Dyn., 4, 95–108. doi:10.5194/esd-4-95-2013.

Braganza, K., D. J. Karoly, A. C. Hirst, M. E. Mann, P. Stott, R. J. Stouffer, and S. F. B. Tett, 2003: Simple indices of global climate variability and change: Part I - variability and correlation structure. Clim. Dyn., 20, 491–502.

Burke, E. J., I. P. Hartley, and C. D. Jones, 2012: Uncertainties in the global temperature change caused by carbon release from permafrost thawing. The Cryosphere, 6, 1063–1076. doi:10.5194/tc-6-1063-2012.

Burke, E. J., C. D. Jones, and C. Koven, 2013: Estimating the Permafrost-Carbon Climate Response in the CMIP5 Climate Models Using a Simplified Approach, J. Climate, 26, 4897-4909.

Ciais, P., C. Sabine, G. Bala, L. Bopp, V. Brovkin, J. Canadell, A. Chhabra, R. DeFries, J. Galloway, M. Heimann, C. Jones, C. Le Quere, R. Myneni, S. Piao and P. Thornton, 2013: Carbon and Other Biogeochemical Cycles. 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 [Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, in press.

Carslaw, K., L. A. Lee, C. L. Reddington, K. J. Pringle, A. Rap, P. M. Forster, G. W. Mann, D. V. Spracklen, M. T. Woodhouse, L. A. Regayre, and J. R. Pierce, 2013: Large contribution of natural aerosols to uncertainty in indirect forcing. Nature, 503, 67. doi:10.1038/nature12674.

Chang, E. K. M., Y. Guo, and X. Xia, 2012: CMIP5 multimodel ensemble projection of storm track change under global warming. J. Geophys. Res., 117. doi:10.1029/2012JD018578.

Collins, M., B. B. B. Booth, B. Bhaskaran, G. Harris, J. M. Murphy, D. M. H. Sexton, and M. J. Webb, 2011: Climate model errors, feedbacks and forcings. A comparison of perturbed physics and multi-model ensembles. Clim. Dyn., 36, 1737-1766. doi:DOI 10.1007/s00382-010-0808-0.

Collins, M., R. Knutti, J. M. Arblaster, J. Dufresne, T. Fichefet, P. Friedlingstein, X. Gao, W. J. Gutowski, T. Johns, G. Krinner, M. Shongwe, C. Tebaldi, A. J. Weaver, and M. Wehner, 2013: Long-term climate change: Projections, commitments and irreversibility. IPCC WG1 Fifth Assessment report, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, chapter 12.

Dee, D. P. et al, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc., 137, 656, 553-597. doi:10.1002/qj.828.

Dunn-Sigouin, E. and S. W. Son, 2013: Northern Hemisphere blocking frequency and duration in the CMIP5 models. J. Geophys. Res., 118, 1179–1188. doi:10.1002/jgrd.50143.

Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S. C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason, and M. Rummukainen, 2013: Evaluation of climate models. IPCC WGI Fifth Assessment Report, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, chapter 9.

Forster, P.M., T. Andrews, P. Good, J.M. Gregory, L.S. Jackson, and M. Zelinka, 2013: Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. J. Geophys. Res., 118, 1-12.

Friedlingstein, P., P. M. Cox, R. A. Betts, L. Bopp, W. von Bloh, V. Brovkin, P. Cadule, S. Doney, M. Eby, I. Fung, G. Bala, J. John, C. D. Jones, F. Joos, T. Kato, M. Kawamiya, W. Knorr, K. Lindsay, H. D. Matthews, T. Raddatz, P. Rayner, C. Reick, E. Roeckner, K. G. Schnitzler, R. Schnur, K. Strassmann, A. J. Weaver, C. Yoshikawa, and N. Zeng, 2006: Climate–carbon cycle feedback analysis, results from the C^4MIP model intercomparison. J. Climate, 19, 3337–3353. doi:10.1175/JCLI3800.1.

Gregory, J. M., C. D. Jones, P. Cadule, and P. Friedlingstein, 2009: Quantifying carbon-cycle feedbacks. J. Climate, 22, 5232-5250 doi:10.1175/2009JCLI2949.1

Harris, G. R., D.M.H. Sexton, B.B.B. Booth, M. Collins, J.M. Murphy, and M.J. Webb, 2006: Frequency distributions of transient regional climate change from perturbed physics ensembles of General Circulation Model simulations, Clim. Dyn., 27, 357-375.

Harris, G. R., D. M. H. Sexton, B. B. B. Booth, M. Collins, and J. M. Murphy, 2013: Probabilistic projections of transient climate change. Clim. Dyn., 40, 2937-2972. doi:10.1007/s00382-012-1647-y.

Harvey, B. J., L. C. Shaffrey, and T. J. Woollings, 2013: Equator-to-pole temperature differences and the extra-tropical storm track responses of the CMIP5 climate models. Clim. Dyn. doi:10.1007/s00382-013-1883-9.

Harvey, B. J., L. C. Shaffrey, T. J. Woollings, G. Zappa, and K. I. Hodges, 2012: How large are projected 21st century storm track changes? J. Geophys. Res., 39. doi:10.1029/2012GL052873.

Ineson, S., and A. A. Scaife, 2009: The role of the stratosphere in the European climate response to El Nino. Nature Geosci., 2, 32–36. doi:10.1038/ngeo381.

Ineson, S., A. A. Scaife, J. R. Knight, J. C. Manners, N. J. Dunstone, L. J. Gray, and J. D. Haigh, 2011: Solar forcing of winter climate variability in the Northern Hemisphere. Nature Geosci., 4, 753–757. doi:10.1038/NGEO1282.

Karpechko, A. Y., and E. Manzini, 2012: Stratospheric influence on tropospheric climate change in the Northern Hemisphere. J. Geophys. Res., 117. doi:10.1029/2011JD017036.

Kendon, E. J., N. M. Roberts, C. A. Senior, and M. J. Roberts, 2012: Realism of rainfall in a very high resolution regional climate model. J. Climate, 25, 5791–5806. doi:10.1175/JCLI-D-11-00562.1.

Knutti, R., G. Abramowitz, M. Collins, V. Eyring, P. J. Gleckler, B. Hewitson, and L. Mearns, 2010: Good practice guidance paper on assessing and combining multi model climate projections. IPCC Working Group I Technical Support Unit, University of Bern, 2010.

Lewis, N., 2013: An objective Bayesian improved approach for applying optimal fingerprint techniques to estimate climate sensitivity. J. Climate, 26, 7414–7429. doi:10.1175/JCLI-D-12-00473.1.

Masato, G., B. Hoskins, and T. J. Woollings, 2012: Winter and summer Northern Hemisphere blocking in CMIP5 models. J. Climate. doi:10.1175/JCLI-D-12-00466.1.

Massonnet, F., T. Fichefet, H. Goosse, C. M. Bitz, G. Philippon-Berthier, M. Holland, and P. Y. Barriat, 2012: Constraining projections of summer Arctic sea ice. The Cryosphere, 1383–1394.

Matsueda, M., R. Mizuta, and S. Kusunoki, 2009: Future change in wintertime atmospheric blocking simulated using a 20-km-mesh atmospheric global circulation model. J. Geophys. Res., 114, D12.

McSweeney, C. F., R. G. Jones, and B. B. B. Booth, 2012: Selecting ensemble members to provide regional climate change information. J. Climate, 25, 7100–7121. doi:10.1175/JCLI-D-11-00526.1.

McSweeney, C. F., R. G. Jones, R. W. Lee, and D. Rowell, 2015: Selecting CMIP5 GCMs for downscaling over multiple regions. Clim. Dyn. 44, 3237-3260.

Morice, C. P., J. J. Kennedy, N. A. Rayner, and P. D. Jones, 2012: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 dataset. J. Geophys. Res., 117, D08101. doi:10.1029/2011JD017187.

Murphy, J. M., B. B. B. Booth, C.A. Boulton, R. C. Clark, G. R. Harris, D. M. H. Sexton, and J. A. Lowe, 2014: Transient climate changes in a perturbed parameter ensemble of emissions driven earth system model simulations. Clim. Dyn. 43, 2855-2885.

Murphy, J. M., D. M. H. Sexton, G. J. Jenkins, B. B. B. Booth, C. C. Brown, R. T. Clark, M. Collins, G. R. Harris, E. J. Kendon, R. A. Betts, S. J. Brown, K. A. Humphrey, M. P. McCarthy, R. E. McDonald, A. Stephens, C. Wallace, R. Warren, R. Wilby, and R. A. Wood, 2009: UK climate

projections science report: Climate change projections. Met Office Hadley Centre, Exeter, UK.
Myhre, G. et al, 2013: Radiative forcing of the direct aerosol effect from AeroCom Phase II simulations. Atmos. Chem. Phys., 13, 1853–1877. doi:10.5194/acp-13-1853-2013.

Nakicenovic ́, N., J. Alcamo, G. Davis, B. de Vries, J. Fenhann, S. Gaffin, K. Gregory, A. Grübler, T. Y. Jung, T. Kram, E. L. La Rovere, L. Michaelis, S. Mori, T. Morita, W. Pepper, H. Pitcher, L. Price, K. Riahi, A. Roehrl, H.-H. Rogner, A. Sankovski, M. Schlesinger, P. Shukla, S. Smith, R. Swart, S. van Rooijen, V. Nadejda, and Z. Dadi, 2000: Emission scenarios. A special report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge University Press. 599pp.

Otto, A., F. E. L. Otto, O. Boucher, J. Church, G. Hegerl, P. M. Forster, N. P. Gillett, J. Gregory, G. C. Johnson, R. Knutti, N. Lewis, U. Lohmann, J. Marotzke, G. Myhre, D. Shindell, B. Stevens, and M. R. Allen, 2013: Energy budget constraints on climate response. Nature Geosci., 6, 415–416. doi:10.1038/ngeo1836.

Overland, J., and M. Wang, 2010: Large-scale atmospheric circulation changes are associated with the recent loss of Arctic sea ice. Tellus, 62A, 1-9.

Petoukhov, V., and V. Semenov, 2010: A link between reduced Barent-Kara sea ice and cold winter extremes over northern continents. J. Geophys. Res., 115, D21111.

Reichler, T. and J. Kim, 2008: How well do coupled models simulate today’s climate? Bull. Am. Meteorol. Soc., 89, 303–311. doi:10.1175/BAMS-89-3-303.

Roberts, C. D., F. K. Garry, and L. C. Jackson, 2013: A multimodel study of sea surface temperature and subsurface density fingerprints of the Atlantic Meridional Overturning Circulation. J. Climate, 26, 9155–9174. doi:10.1175/JCLI-D-12-00762.1.

Rougier, J., 2007: Probabilistic inference for future climate using an ensemble of climate model evaluations. Clim. Change, 81, 247–264.

Scaife, A. A., D. Copsey, C. Gordon, C. Harris, T. Hinton, S. Keeley, A. O’Neill, M. Roberts, and K. Williams, 2011: Improved Atlantic winter blocking in a climate model. Geophys. Res. Lett., 38. doi:10.1029/2011GL049573.

Scaife, A. A., T. Spangehl, D. R. Fereday, U. Cubasch, U. Langematz, H. Akiyoshi, S. Bekki, N. Butchart, M. P. Chipperfield, A. Gettelman, S. C. Hardiman, M. Michou, E. Rozanov, and T. G. Shepherd, 2012: Climate change projections and stratosphere-troposphere interaction. Clim. Dyn., 38, 2089–2097. doi:10.1007/s00382-011-1080-7.

Screen, J. A., I. Simmonds, C. Deser, and R. Tomas, 2013: The atmospheric response to three decades of observed Arctic sea ice loss. J. Climate, 26, 1230–1248. doi:10.1175/JCLI-D-12-00063.1.
Sexton, D. M. H., and J. M. Murphy, 2010: UKCP09: Probabilistic projections of wind speed. UKCP09 additional product report. Available from http://ukclimateprojections.metoffice.gov.uk/22530.

Sexton, D. M. H., J. M. Murphy, M. Collins, and M. J. Webb, 2012: Multivariate probabilistic projections using imperfect climate models Part I: Outline of methodology. Clim. Dyn., 38, 2513-2542. Nature Geoscience, 5, 788-792. doi:10.1007/s00382-011-1208-9.

Srokosz, M., M. Baringer, H. Bryden, S. Cunningham, T. Delworth, S. Lozier, J. Marotzke, and R. Sutton, 2012: Past, present and future changes in the Atlantic Meridional Overturning circulation. Bull. Amer. Met. Soc., Nov 2012, 1663-1676.

Sutton, R. T. and B. W. Dong, 2012: Atlantic ocean influence on a shift in European climate in the 1990s. doi:10.1038/NGEO1595.

Vial, J., J. Dufresne, and S. Bony, 2013: On the interpretation of intermodel spread in CMIP5 climate sensitvity estimates. Clim. Dyn. doi:10.1007/s00382-013-1725-9.

Woollings, T., 2010: Dynamical influences on European climate: an uncertain future. Phil. Trans. R. Soc. Lond. A, 368, 3733-3756.

Woollings, T., A. Charlton-Perez, S. Ineson, A. G. Marshall, and G. Masato, 2010: Associations between stratospheric variability and tropospheric blocking. J. Geophys. Res., 115, D06108.

Zaehle, S., P. Friedlingstein, and A. D. Friend, 2010: Terrestrial nitrogen feedbacks may accelerate future climate change. Geophys. Res. Lett., 37, L01401.

Zappa, G., L. C. Shaffrey, and K. I. Hodges, 2013a: North Atlantic cyclones and model biases in the CMIP5 models. J. Climate, 26, 5379–5396.

Zappa, G., L. C. Shaffrey, K. I. Hodges, P. Sansom, and D. Stephenson, 2013b: The response of North Atlantic cyclones to climate change in CMIP5 models. J. Climate, 26, 5846–5862.