Estimating change in extreme European precipitation using a multi-model ensemble

Fowler, H.J.[1*], Ekström, M.[2], Blenkinsop, S.[1], and Smith, A.P.[1]

[1] Water Resource Systems Research Laboratory, School of Civil Engineering and Geosciences, Newcastle University, Newcastle upon Tyne, NE1 7RU

[2] Centre for Air Transport and the Environment, Manchester Metropolitan University Faculty of Science and Engineering, Chester Street, Manchester, M1 5GD

***corresponding author**

Dr. H.J. Fowler

WRSRL

School of Civil Engineering and Geosciences

Cassie Building

University of Newcastle

Newcastle upon Tyne

NE1 7RU

UK

h.j.fowler@ncl.ac.uk

Submitted to Journal of Geophysical Research - Atmospheres

Submission date: 5 March 2007; revised 7 June 2007

Abstract

Using the results from multi-model ensembles enables the assessment of model uncertainty in present and future estimates of extremes and the production of probabilities for regional or local scale change. Six regional climate model (RCM) integrations from the PRUDENCE ensemble are used together with extreme value analysis to assess changes to precipitation extremes over Europe by 2070-2100 under the SRES A2 emissions scenario, investigating the contribution of the formulations of global (GCM) and regional climate models to scenario uncertainty. RCM ability to simulate precipitation extremes is evaluated for a UK case study. RCMs are shown to underestimate 1 day return values but reasonably simulate longer duration (5 or 10 day) extremes. A multi-model approach by which probabilities can be produced for regional or local scale change in extremes is then developed.

A key result is that all RCMs project increases in the magnitude of short and long duration extreme precipitation for most of Europe. Individual model projections vary considerably but are independent of changes in mean precipitation. The magnitude of change is strongly influenced by the driving GCM but moderated by the RCM, which also influences spatial pattern. Therefore, when designing future ensemble experiments (a) the number of GCMs should at least equal the number of RCMs; (b) if spatial pattern is important then integrations from different RCMs should be incorporated. For impact studies, both the resolution and number of models in the ensemble will influence projections of change. The use of a multi-model approach therefore provides more robust estimates.

Index Terms:

1630 Impacts of global change

1637 Regional climate change

1807 Climate impacts

1817 Extreme events

1854 Precipitation

Keywords: precipitation, extremes, Regional Climate Model, ensemble, probabilities, climate change

Introduction

Global analyses of precipitation intensities in observed data (e.g. Frich *et al*., 2002; Alexander *et al*., 2006) indicate that high latitudes of the northern Hemisphere are currently experiencing a trend towards increased rainfall and enhanced variability (e.g. Easterling *et al*., 2000; Meehl *et al*., 2005), particularly in winter. This is supported by regional studies in Europe (e.g. Fowler and Kilsby, 2003a, b; Brunetti *et al*., 2000; Frei and Schär, 2001) which show significant positive trends in intensity over the past decade. Such a trend is likely to continue into the future as modeling studies with global climate models (GCMs) (e.g. Giorgi *et al*., 2001; Palmer and Räisänen, 2002; Tebaldi *et al.*, 2006) consistently suggest that under enhanced greenhouse conditions there will be increases in the frequency and intensity of heavy precipitation.

Coarse resolution global climate models are unable to simulate realistic extreme events, particularly in areas of complex topography (Mearns *et al*., 2001). Regional detail on extremes can however be obtained by using simple interpolation, statistical downscaling or high-resolution dynamical modeling using Regional Climate Models (RCMs) (Haylock *et al*., 2006). Dynamical modeling confers advantages over other methods as it still represents physical processes but at a higher resolution. Therefore, much recent attention has been focused on the simulation of extremes by RCMs (e.g. Christensen and Christensen 2003, 2004; Pal *et al.* 2004; Räisänen *et al.*, 2004; Ekström *et al.*, 2005; Frei *et al*., 2006).

Estimates of future precipitation are, however, subject to several sources of uncertainty (Allen and Ingram, 2002; Covey, *et al*., 2003, Collins and Allen, 2002; Jenkins and Lowe, 2003). Two major sources are related to model structure and parameterization scheme and are likely to be reduced by further research. In addition, major uncertainties result from the emission rates of greenhouse gases which are determined by society through policies. When using RCM data the sources of uncertainty increase, as outputs are influenced by RCM resolution, numerical scheme, physical parameterizations and the forcing boundary conditions (Rummukainen *et al*., 2001; Déqué *et al*., 2007). Recently, results from multi-model ensembles have become available in projects such as PRUDENCE (Prediction of Regional scenarios and Uncertainties for Defining European Climate change risks and Effects; Christensen *et al*., 2007), ENSEMBLES and NARCCAP (North American Regional Climate Change Assessment Program), enabling assessment of model uncertainty in present and future estimates of extremes. Although ENSEMBLES used only four GCMs to drive nine RCMs, NARCCAP will use a better balance of GCMs and RCMs, including some of the same GCMs as PRUDENCE (Mearns *et al*., 2006).

Multi-model ensembles have also allowed the generation of probability density functions (pdfs) of the impacts of global warming. Despite the global emphasis taken by most probabilistic climate change assessments, there are now examples in the literature of the production of pdfs for regional-scale (e.g. Tebaldi *et al.*, 2004, 2005; Greene *et al.*, 2006; Ekström *et al.*, 2007) and even point-scale (Furrer *et al.*, 2007) changes. Although most studies concentrate on mean changes, Palmer and Räisänen (2002) used 19 global climate models to quantify the increases in the probability of extreme precipitation for different regions of the world under global warming using equal weighting of the results from different models. However, more recent approaches, e.g. Tebaldi *et al.* (2004, 2005) or Lopez *et al.* (2006), suggest that non-uniform weighting may be more appropriate as models have unequal skill in the simulation of contemporary climate.

There are several different methods available to define extreme events. Many studies have focused on 'soft' extremes (Klein Tank and Können, 2003), typically 90th or 95th percentiles, principally because the detection probability of trends decreases for even moderately rare events (Frei and Schär, 2001). For the European continent, there have been a number of recent studies summarized by Frei *et al.* (2006). Other studies have used extreme value analysis to examine more rare events. A series of publications have assessed different versions of Hadley Centre RCMs for the United Kingdom (e.g. Jones and Reid, 2001; Huntingford *et al.*, 2003; Fowler *et al.*, 2005; Buonomo *et al.*, 2007). Most recently, for HadRM3H, Ekström *et al.* (2005) found a 10% increase in 1 day precipitation intensities for return values from 10 to 50 years across the UK, estimating more spatially variable changes for 5 and 10 day events. The most extensive European study, by Frei *et al.* (2006), used six RCMs driven by the HadAM3H atmosphere-only GCM and found that in winter, precipitation extremes tend to increase north of about 45°N while there are insignificant changes or decreases to the south. In summer, the models produce divergent estimates of change, with RCM structure and parameterization contributing significantly to scenario uncertainty.

Here, we use PRUDENCE integrations from four models: three RCMs and an atmosphere-only GCM with a similar spatial resolution, with lateral boundary conditions taken from two different GCMs. Regional Frequency Analysis (Hosking and Wallis, 1997) is used to fit the Generalized Extreme Value (GEV) distribution to annual maxima using the method of L-moments, to define extremes with return values from 5 to 25 years. We then compare return values for the current (1961-1990) and future (2070–2100) climate, under the SRES A2 emissions scenario, to analyze changes to precipitation extremes over Europe. This work builds on Frei *et al.* (2006) by investigating the contribution of the formulation of the driving GCM to scenario uncertainty in extremes. This was found to be greater than that from the RCM for mean climate response, particularly temperature, by Déqué *et al.* (2007).

The RCMs are also evaluated with respect to their ability to simulate precipitation extremes. To complement previous evaluations focusing on the Alps (Frei *et al.*, 2006) and southern Germany (Beniston *et al.*, 2007) using some of the same models, we focus on the UK; approximately 20 x 18 0.5° grid cells, updating a study by Fowler *et al.* (2005). We then consider how estimates from different models may be combined. We develop a method using non-parametric bootstrap re-sampling by which probabilities can be produced for regional or local scale change in extremes.

The paper is divided into the following sections: Section 2 presents the data and descriptions of the climate models used in the study; Section 3 introduces the statistical methods used for analysis; Section 4 presents an evaluation of the models' ability to simulate mean and extreme precipitation statistics using a UK case study; Section 5 presents future scenarios of precipitation extremes over Europe; Section 6 explores how probabilistic estimates of change in extremes may be developed for homogeneous rainfall regions in the UK; and Section 7 provides a discussion of the results and concludes the study.

1. Model Descriptions and Data

Within the FP5 PRUDENCE project (Christensen *et al.*, 2007), four Atmosphere-Ocean and Atmosphere-only GCMs were used to drive nine RCMs and one variable resolution global atmospheric model over a European domain for two time slice integrations: control (1961-1990; CTRL) and future (2071-2100; SCEN). Daily grid-point values for a range of climatic variables are available at http://prudence.dmi.dk. Here, we examine integrations from four models within the PRUDENCE ensemble, three RCMs and one variable resolution global atmospheric model, driven by two different GCMs for the IPCC SRES A2 emissions scenario (Nakićenović *et al.*, 2000). This subset of available model integrations was chosen to evaluate the relative contribution to RCM uncertainty in future projections by assessing:

- Same bounding GCM in combination with different RCMs
- Same RCM in combination with different bounding GCMs

Two of the RCM integrations analyzed in this study, HIRHAM and RCAO, were conducted by nesting into the atmosphere-only high-resolution GCM HadAM3H of the UK Hadley Centre. One RCM, HadRM3P, is nested into HadAM3P, a more recent version of the same atmosphere-only GCM. The latter version contains changes to the moisture parameterizations which affect biases seen in parts of the globe outside Europe; therefore HadRM3H and HadRM3P can be considered as essentially the same model for Europe (Haylock *et al.*, 2006). Additionally, the variable resolution global atmospheric model, ARPEGE, with a resolution of 50 km to 70 km over Europe (Hagemann *et al.*, 2004), is nested directly within HadCM3.

HadCM3 (Gordon *et al.*, 2000; Johns *et al.*, 2003) is a coupled ocean-atmosphere GCM at a resolution of approximately 300 km from which both HadAM3H and HadAM3P take their boundary conditions. HadAM3H (Pope *et al.*, 2000) and HadAM3P (Jones *et al.*, 2005) have a resolution of about 150 km in the mid-latitudes. For CTRL, they were forced by observed sea surface conditions from the same period. For SCEN, sea surface conditions were constructed by adding anomalies from a transient simulation of HadCM3 to observations. Atmosphere-only GCMs were favored over HadCM3 for driving RCMs in PRUDENCE, as their higher resolution provides an improved control climate, particularly with respect to the positioning of storm tracks in the Northern Hemisphere (Hudson and Jones, 2002). Furthermore, the representation of clouds and condensation are substantially improved (Stratton *et al.*, 2004).

Additionally, two RCM integrations analyzed here, HIRHAM and RCAO, are driven by lateral boundary and sea surface conditions from the ECHAM4/OPYC3 coupled ocean-atmosphere GCM (Roeckner *et al.*, 1996; 1999) developed in co-operation between the Max-Planck-Institute for Meteorology (MPI) and the German Climate Computing Centre (DKRZ) in Hamburg, Germany. These are included to sample the dependence of results on the driving GCM.

The HadAM3H/P and ECHAM4/OPYC3 global mean temperature responses are similar for the IPCC SRES A2 emissions scenario (3.1°C and 3.56°C respectively; Tim Osborn, personal communication); mid-range in the climate sensitivities presented by the IPCC (2001).

The RCMs considered in this study are listed below with further details in Table 1. All operate with grid spacing of about 0.5° longitude by 0.5° latitude (approximately 50 km spatial resolution) over a European domain. More details on the experimental design of the PRUDENCE integrations can be found in Jacob *et al.* (2007).

1. The HadRM3P model of the UK Hadley Centre (Jones *et al.*, 2004a) is an updated version of the HadRM3H model (Hudson and Jones, 2002). Precipitation extremes for the UK for HadRM3H are described by Fowler *et al.* (2005) and Ekström *et al.* (2005). The main changes in HadRM3P are related to the calculation of large-scale cloud and assumptions made about the radiative effects of convective clouds. Changes were also made to precipitation efficiency parameters to ensure reasonable vertical cloud profiles, cloud forcing and radiation fields (Déqué *et al.*, 2007);
2. The HIRHAM model of the Danish Meteorological Institute is an updated version of HIRHAM4 (Christensen *et al.*, 1996; 1998), incorporating high resolution physiographical datasets of surface topography and land use classification (Hagemann *et al.*, 2001; Christensen *et al.*, 2001). Regional simulations of extreme precipitation by HIRHAM are described in Christensen and Christensen (2003; 2004) and for the whole of Europe by May (2007);

3. The RCAO model of the Swedish Meteorological and Hydrological Institute consists of an atmospheric part RCA2 (Jones *et al*., 2004b) and an ocean model RCO (Meier *et al.*, 2003), described in Döscher *et al.* (2002). The simulation of extreme precipitation over part of northern Europe is described by Räisänen *et al.* (2004);
4. The global ARPEGE/IFS variable resolution model of the French Meteorological Service, an updated version of Déqué *et al.* (1998), is not strictly a RCM. Within PRUDENCE, however, it is used with maximum resolution over the Mediterranean Sea (Gibelin and Déqué, 2003) and so its resolution over Europe is approximately the same as the other RCMs. The simulation of extreme precipitation over France is described by Déqué (2007).

Daily precipitation data for CTRL and SCEN integrations were re-gridded onto the common 0.5° x 0.5° CRU grid to allow direct comparison between models. Suffixes E and H denote RCMs driven by ECHAM4/OPYC3 and HadAM3H/P/HadCM3 GCMs respectively.

For the UK evaluation, a daily observed 5 km precipitation grid produced by the UK Meteorological Office (Perry and Hollis, 2005a, b) was aggregated to the common CRU grid by taking a daily average across the 5 km boxes contained within each 0.5° x 0.5° grid cell for each day of 1961-1990. This daily data set is referred to as UKMO. No similar daily observational dataset is available for Europe, although one is under construction in the FP6 ENSEMBLES project (Malcolm Haylock, personal communication).

2. Statistical Analysis

The statistical analysis of extreme precipitation is based on daily precipitation totals. First, diagnostics of mean precipitation and wet day frequency are used to characterize the frequency distribution of precipitation. A threshold of 1 mm d^{-1} is used to discriminate wet days as lower daily totals may be sensitive to under-recording in observed series. Secondly, we analyze return values of precipitation intensities with average recurrence of 5 to 25 years. The return value for a return period of T years is defined as the precipitation intensity that is exceeded by an annual extreme with a probability of $1/T$. Return values for return periods in excess of 25 years were considered less reliable due to the short (30 year) length of climate model integrations and so results are not presented here. Return values are examined for 1, 2, 5 and 10 day precipitation totals. Here, results are presented only for illustrative low (5 year) and high (25 year) return periods using 1 and 10 day sums, representing short and long duration precipitation events respectively.

a. Return period estimation

For each RCM integration, annual maximum (AM) series are extracted for 1, 2, 5 and 10 day precipitation totals for each grid cell. The AM series are standardized by their median (Rmed; following Fowler *et al.*, 2005) and a GEV distribution is fitted using L-moments (Hosking and Wallis, 1997). Grid cell return period magnitudes are then derived for each model by multiplying the fitted GEV growth factor by its respective Rmed.

For the UK case study, regional frequency analysis (RFA) is used to estimate return values of precipitation intensities for nine UK rainfall regions (Figure 1) developed by Wigley *et al*. (1984). The homogeneity of these regions for extreme precipitation was tested by Fowler and Kilsby (2003a). Within each region, standardized AM data for each grid cell is pooled and a GEV distribution fitted using regionally averaged L-moment ratios. The return values of precipitation intensities are then estimated by multiplying the fitted growth factor by the regional average Rmed. This technique is used to estimate regional return values for UKMO, CTRL and SCEN. The methodology is explained in more detail in Fowler and Kilsby (2003a).

We use the return value estimates for UKMO and CTRL to evaluate the RCMs with respect to their representation of precipitation extremes across the UK. We then consider the difference between CTRL and SCEN to give estimates of change in the return values of precipitation intensities across Europe.

b. Confidence intervals

For the UK case study regional confidence intervals on return values were estimated using a non-parametric bootstrap re-sampling method (Efron, 1979). If each dataset of AM is based on *n* data points then, as defined by Efron and Tibshirani (1993), bootstrapping samples the original dataset with replacement multiple times to produce multiple independent samples of size *n*. Thus, 10000 bootstrap samples, each of 30 values are drawn from each pooled standardized regional AM dataset for the RCMs and UKMO. For each bootstrap sample, a GEV distribution is fitted and the 5 and 25 year return values estimated by multiplying the growth factor by the regional average Rmed. This allows the construction of 5th and 95th percentiles for return value estimates for individual regions and RCMs. More detail on this method is given in Fowler *et al.* (2005).

c. Multi-model estimates

Multi-model estimates of change were generated using the non-parametric bootstrap samples. Using the 5 year return value as an example; for each RCM, region and aggregation, e.g. HADH for CEE at 1 day, a random number generator is used to sample the return values from the CTRL and SCEN regional pools. The percentage change in the return value between CTRL and SCEN is then calculated. This procedure is repeated 10000 times, giving 10000 estimates of change in the 5 year return value for each RCM and each region. A kernel density function is then fitted to visualize the estimated change.

Assuming that the models have equal skill, the 10000 estimates from each RCM are then pooled and the distributional properties examined using box plots for each UK region for the 5 and 25 year return values of 1 and 10 day precipitation extremes.

3. Evaluation in the UK

This section presents an evaluation of the simulation of UK precipitation extremes; comparing CTRL to UKMO. UKMO has the same grid resolution as the RCMs and therefore the resulting statistics are directly comparable (Osborn and Hulme, 1997).

Firstly, we compare mean precipitation and wet day frequency to characterize the frequency distribution of precipitation across the UK. Subsequently, we compare the mean and standard deviation of AM series and Rmed for each region. Finally, we compare the estimated return values for individual grid cells and regions.

a. Mean and frequency diagnostics

The CTRL integrations broadly simulate the observed annual cycle of precipitation over the UK (Figure 2a). However, model skill differs throughout the year and is most influenced by choice of RCM, except in autumn when precipitation tends to be greatest. Here, the driving GCM provides significant differences: CTRL integrations driven by ECHAM4/OPYC3 overestimate precipitation and those driven by Hadley GCMs underestimate precipitation. This perhaps suggests a seasonal disparity in model ability to simulate precipitation mechanisms. In contrast, most RCMs overestimate the wet day frequency (WD, >1 mm, Figure 2b) and the lack of clustering of RCMs according to driving GCM suggests that RCMs have a large influence over the precipitation occurrence process. Frei *et al*. (2003) identified similar problems with simulations of WD over the European Alps using ARPEGE and HIRHAM models driven by observed data. However, a selection of other RCM simulations including HadRM3H performed reasonably well. The simulations of WD for the UK would seem to confirm their speculation that the errors are not region specific but are inherent to specific model parameterizations.

Some models that provide good estimates of areal average precipitation show poor skill in the simulation of its spatial distribution (Figure 2c). Figure 2a shows that HIRHAME simulates the seasonal pattern of areal average precipitation well. However, this is due to the compensating large underestimates over the north and west and overestimates over central and eastern England (Figure 2c). In contrast, although ARPEGEH performs poorly in simulating areal average precipitation, the lack of a clear pattern to its spatial anomalies in Figure 2c suggests that it may be better at representing physical precipitation processes than models which produce errors with a well-defined spatial structure.

b. Annual maxima

The regional mean AM and standard deviation of AM within a region were calculated to identify percentage differences between pooled 1 and 10 day AM series for UKMO and CTRL (Tables 2 and 3).

At 1 day, all CTRL integrations underestimate the regional mean AM. At 10 days, differences between CTRL and UKMO are smaller and, for some regions, the CTRL integrations overestimate the regional mean AM (Table 2b). At both 1 and 10 days, the largest differences between CTRL and UKMO are simulated for northern Scotland (NS) and the smallest for CEE.

The standard deviation of AM within a region gives an indication of the ability of an RCM to reproduce the observed spatial variability in extremes. At 1 day, RCAOE and RCAOH produce large underestimates. At 10 days, RCAOH and RCAOE still underestimate the standard deviation of AM in most regions (Table 3b). At both 1 and 10 days, HADH provides the most similar or lowest mean model anomaly value (Tables 2b and 3b). The largest difference in standard deviation of AM between CTRL and UKMO is found for NS.

c. Rmed

Here we illustrate how Rmed distributions, used to rescale from the fitted GEV growth factor to the return value and equivalent to the 2 year return value, differ for each region. The mean and standard deviation of Rmed by model and region is compared for CTRL and UKMO in Figure 3. Color represents the CTRL and UKMO datasets whilst symbols represent regions, i.e. clustering of colors indicates similarity amongst models, whilst clustering of symbols indicates similarity within regions.

For 1 day AM, there is more similarity between Rmed values from the same RCM than Rmed values from the same region (Figure 3a). However, for 10 day AM, regional-clusters suggest that the mean and standard deviation of Rmed values from models within the same region are more similar (Figure 3b).

Differences between 1 and 10 day Rmed distributions for CTRL and UKMO are clearly identified in Figure 3. The UKMO estimates are placed further to the right than the CTRL markers in both plots. This shows that the CTRL integrations underestimate mean Rmed, although less so for the 10 day totals. Furthermore, with some exceptions, mainly for HADH, the CTRL markers show less vertical spread compared to UKMO markers, particularly at 1 day. This suggests that, in general, the CTRL integrations underestimate the spatial variability of Rmed.

Box plots of Rmed distributions for CTRL and UKMO were plotted to illustrate the two regions showing the smallest (CEE) and largest (NS) differences (Figure 4). The larger spatial variability of precipitation in NS is illustrated by the wider range of Rmed when compared to CEE (Figure 4). There are larger inter-model and CTRL-UKMO differences for NS than CEE, where the range of RCAOE/H is much smaller than the range of observed Rmed values.

d. Return values

Figure 5 shows the estimated 1 day, 5 year return value for UKMO and each of the CTRL integrations for individual grid cells (Figure 5a) and regions (Figure 5b). At the 1 day resolution, all CTRL integrations underestimate extreme precipitation amounts for both low (Figure 5a, b) and high return values (not shown). This was noted by Fowler *et al*. (2005) for HadRM3H and may be a result of the poor performance of RCMs in resolving convective precipitation processes. As precipitation is aggregated to the regional level, return value estimates are improved due to the effect of data pooling and use of the regional average Rmed (Figure 5b).

There is considerable variability in model performance over time and space. The RCAO models, in particular, show little spatial variation in 1 day, 5 year return values, with a UK range from 25 to 37 mm (Figure 5a). In comparison, the range estimated from UKMO is 32 to 87 mm. Better estimates are made for longer duration precipitation. For the 10 day, 5 year return value all RCMs show a reduction in simulation error when compared to the 1 day estimate, but a lack of spatial variability is still evident for the RCAO models (Figure 5c). Results for higher return values are comparable (not shown). The improvement in the simulation of longer duration precipitation extremes may reflect the models' ability to better capture large scale atmospheric processes or be due to the reduced influence of model parameterization when using temporal smoothing. Despite the dry bias in mean precipitation across the UK simulated by HADH, it produces the best estimates of return values for CTRL.

Whilst simulated return values for CTRL cannot be assessed relative to observations for the whole of Europe due to the lack of a gridded daily precipitation series, some useful indicators may be inferred from a comparison of the RCMs. Models agree that the largest 1 day, 5 year return values occur over the Alps, western Scandinavia, northwest Spain and the north Mediterranean coast (Figure 6). As for the UK case-study, the RCAO simulations show less spatial variability in 1 day return values than the other RCMs. In general, the models are in closer agreement on the distribution of high and low return values and the range of variability across Europe for longer duration precipitation events (5 or 10 days), with the main differences found over central and Eastern Europe (not shown).

4. Projected change in precipitation extremes

Increases in short duration extreme precipitation are projected over most of Europe (Figure 7), with projected changes for higher return levels generally larger but displaying more inter-model variability. GCM boundary conditions are important in generating differences in projected changes since large-scale circulation patterns within RCMs depend on the lateral boundary conditions inherited from their driving GCMs. These not only influence the simulation of mean precipitation changes over northern Europe (Beniston *et al.,* 2007) but also the extremes. For some regions, this results in different directions of change, e.g. moderate decreases in the magnitude of extreme precipitation events (<40%) over southern Iberia are projected by Hadley-driven models whereas ECHAM-driven models project increases.

Increases also predominate for longer duration extremes, e.g. 10 day precipitation intensities show modest increases in 5 year return values over most of Europe (Figure 8a). For 25 year return values, larger areas are projected to experience decreases (Figure 8b). However, the dominant pattern suggests larger increases over northern Europe, with smaller increases or potentially decreases in extremes over southern Europe. The uncertainty in the spatial pattern of change is strongly influenced by driving GCM, with ECHAM-driven models projecting much larger increases in return values than Hadley-driven models. Overall, there is a more coherent inter-model signal in projections for longer duration precipitation extremes, perhaps reflecting the better simulation of these types of events by RCMs.

Changes in extremes are not directly related to changes in mean precipitation. Models show much greater consistency in projections of mean precipitation change, with decreases over southern Europe and increases over the north (Blenkinsop and Fowler, 2007). Over northern Europe, increases in extremes are likely to be related to proportionately more precipitation in areas of existing storm tracks and associated dynamical moisture convergence resulting simply from the greater moisture holding capacity of warmer air together with a slight poleward shift of the mid-latitude storm tracks (Meehl *et al.*, 2005; Tebaldi *et al.*, 2006). However, over parts of southern Europe, increases in extremes are associated with decreases in mean precipitation. This inconsistency may be a result of an increased number of dry days together with more intense convective extremes, despite lower mean precipitation.

5. Estimating changes using a multi-model approach : UK example

Probability distributions of change in extreme precipitation were generated using the 10000 return value estimates generated for each model, aggregation period (1, 2, 5 or 10 days) and UK homogenous rainfall region by the non-parametric bootstrapping exercise. The methodology used is detailed in Section 3c.

Figures 9 and 10 show estimates of the percentage change in the 1 day 5 year and 10 day 25 year return values respectively for each RCM under the SRES A2 2071-2100 emissions scenario. Model projections of change vary considerably; estimates range from -20 to +60 % across the UK with a greater spread for higher return periods, and intra-regional differences are almost as large. What is striking is that few models predict decreases in any region; HADH proves the outlier by projecting decreases or no change. A distinct split between projections from Hadley- and ECHAM-driven models is seen at 1 day, except in southeast England (SEE) and central east England

(CEE). However, at 10 days, this is less distinct. ECHAM-driven models project large increases in extreme precipitation in all UK regions; generally larger than increases projected by Hadley-driven models, which also show large intra-ensemble differences.

Assuming that the models have equal skill, the 10000 estimates from each RCM were pooled to produce probability distributions of change for each UK region. Figure 11 shows box plots of the estimated percentage change in 1 and 10 day, 5 and 25 year return values. Data pooling produces large uncertainties in projections, particularly for 1 day extremes. Less spatial variability in change is estimated for 10 day extremes; all regions show median increases from 10-20%, with greater uncertainty surrounding estimates for the southern UK. At 1 day, larger changes are projected for northern regions (~20% increase), with 10-20% increases estimated for southern regions. At higher return levels there is greater uncertainty, as would be expected. However, it is likely that the choice of models in the pool heavily influences the results.

Estimates of change were then pooled by driving-GCM to examine uncertainty resulting from GCM boundary conditions and the structure and parameterization of RCMs. Figures 12 and 13 show three box plots of probability distributions of change estimated for 1 and 10 day, 5 and 25 return values. Firstly, all Hadley-driven model results were pooled to give an estimate of change using four ensemble members, HADH, ARPEGEH, HIRHAMH and RCAOH. Secondly, a strict comparison was made of pooled results from a 4 x 4 ensemble: RCAO and HIRHAM driven by HadAM3H and ECHAM4/OPYC3. The second plot (Hadley_sub) shows results for this pooled subset of Hadley-driven models: HIRHAMH and RCAOH. The third plot in each row shows pooled results for the same RCMs driven by ECHAM4/OPYC3: HIRHAME and RCAOE.

Figures 12 and 13 clearly illustrate the large effect of the driving GCM on the magnitude of estimated changes. For 1 day 5 year return values (Figure 12a), median increases of between 0 to 15% are estimated for all Hadley-driven RCMs but median increases for the Hadley subset are much lower for southern UK regions, from 0 to 5%. Median increases projected by the ECHAM-driven RCMs are significantly larger; from 25 to 45%. At higher return levels (Figure 12b), greater increases of up to 60% are projected. Projected increases are largest in east and south Scotland (ES and SS) and smallest in southern regions. For short duration precipitation extremes, the size of the RCM ensemble heavily influences change projections, with the spread significantly increasing with the inclusion of additional RCMs.

For 10 day extremes (Figure 13), differences between ECHAM- and Hadley-driven projections and the range of uncertainty are both smaller. Changes estimated from the subset do not differ significantly from estimates using all Hadley-driven models, suggesting increases of 10 to 15% in the 5 year return value (Figure 13a). Thus, ensemble size is much less important. ECHAM-driven models suggest greater increases, from 25 to 35%. For the 25 year return value (Figure 13b), increases are projected to be of similar magnitude.

6. Discussion and Conclusions

A selection of three RCMs and a variable resolution atmosphere-only GCM, with forcing from two different GCMs, were compared for the simulation of European precipitation extremes. A UK case study demonstrated each model's ability to reproduce observed climate statistics for the 1961-1990 control integration (CTRL) for nine rainfall regions. The study also examined some of the uncertainties associated with model structure and parameterization. Results showed that the driving GCM has a strong influence on the magnitude of change in extremes; similar results were obtained for mean change in precipitation and temperature over Europe by Déqué *et al.* (2007). However, RCM structure influences the spatial pattern of change in extreme precipitation and moderates the median magnitude of change; shown by the, sometimes large, differences between projections from Hadley-driven models.

All models were found to reproduce the form of the annual precipitation cycle over the UK. However, the lack of expected spatial patterns in mean precipitation suggests that models may not adequately capture the physical processes responsible for precipitation. Comparisons of regional means and standard deviations of the CTRL Rmed values with observed equivalents indicated much less intra-regional variability in modeled than observed precipitation, particularly so for the 1 day totals. Scatter-plots of regional Rmed mean and standard deviation showed that whilst 1 day Rmed clustered according to model (i.e. Rmed from same model but different regions showed similar values), the 10 day Rmed clustered according to region. This suggests that 1 day rainfall may reflect too much dependence on model specific behavior rather than regional climate characteristics, an effect that is reduced when averaging RCM data in the temporal or spatial domain.

For extreme precipitation, all models, to varying degree, underestimated observed statistics and intra-regional spatial variability, hence giving a conservative measure of the magnitude of extremes. At 1 day, models provided poor estimates but simulations of longer duration extremes (5 or 10 days) were reasonable and show that RCMs are capable of representing the spatial patterns in extremes that are not resolved by GCMs. The nature of the deficiencies in model performance highlighted in this study has yet to be fully addressed; here, %-change was used to define probabilities for change in extreme precipitation rather than absolute values. In particular, there is a lack of rigorous comparative analyses of model skill in reproducing characteristics of the large-scale atmospheric circulation and its relationship with regional climate.

At the European scale, increases in both short and long duration extreme precipitation are projected, although there is uncertainty in the absolute magnitude. Coherent spatial patterns are rarely found for extreme precipitation projections as a result of the small-scale, local character of precipitation (Tebaldi *et al.,* 2006; Frich *et al.,* 2002). However, importantly for policy makers, reductions are projected over comparatively small areas, with general agreement amongst models for increases in longer duration events. This change is physically consistent with warmer air in the future climate being able to hold more moisture generated by increased evaporation from warmer oceans. When this moister air moves over land, more intense precipitation is produced (Meehl *et al.,* 2005). Change in extremes is driven by changes in mean precipitation in some areas but not in others, confirming the conclusions of Frei *et al.* (2006) who indicated that there is a component of change under global warming that is specific to extremes.

For the UK rainfall regions, results from all models were pooled per region and a non-parametric bootstrap re-sampling method, similar to Huntingford *et al*. (2003), was used to represent uncertainty resulting from natural climate variability. This combination allowed probability distributions of change in the 5 and 25 year return values to be estimated for the nine rainfall regions. The median change indicates an increase of 10 to 20% on 1961-1990 return values by 2071 under the A2 SRES emissions scenario. More uncertainty exists for change in short duration (1 or 2 days) precipitation extremes but consistently positive changes are predicted for longer duration events. If estimates of change for 1 day extreme precipitation are pooled by driving-GCM, Hadley-driven models estimate median increases of between 0 and 15% whereas ECHAM-driven models project larger increases of 25 to 45%. Estimated increases are largest in east and south Scotland and smallest in southern regions; similar to trends seen in observations (Fowler *et al*., 2003a, b). For 10 day extremes, the differences between the projections of change are smaller. Hadley-driven models suggest increases of 10 to 15%; ECHAM models suggest greater increases from 25 to 35%.

The large inter-model variability evident in the results suggests that introduction of further models into the analysis, particularly RCMs driven by different lateral boundary conditions, may well modify the estimated changes in extreme precipitation presented here. Furthermore, in this work all models were assumed to have equal skill. However, it is clear that models perform differently in the simulation of the magnitude and spatial distribution of extremes. Under global warming, the characteristics of precipitation are expected to change over both space and time. Models which cannot capture inter-regional differences in the present climate may not accurately predict change. Therefore, it is appropriate to assess the ability of models to simulate spatial as well as temporal climate characteristics. Ultimately, a multi-scale approach must be developed for weighting the results from different climate models. This should not only assess the simulation of synoptic-scale regional climate for the specific impact study, but also the simulation of continental-scale and global modes of variability such as the location of the storm track across Europe and the North Atlantic or the El Niño Southern Oscillations.

In summary, four lessons can be learnt for the design of future climate model ensemble experiments. Firstly, the number of driving GCMs should at least equal the number of RCMs, as the driving GCM seems to produce the largest uncertainties in response, particularly for precipitation. Indeed, the role of GCM uncertainty may be underestimated by PRUDENCE, as the global mean temperature response of HadCM3 and ECHAM4 are very similar for the IPCC SRES A2 emissions scenario (Déqué *et al.,* 2007). Secondly, if spatial variability or extremes are important then integrations from different RCMs must be incorporated into the analysis. Thirdly, the estimates of change are sensitive to the number of RCMs used. In the limited sensitivity analysis performed here, adding more ensemble members increased the uncertainty in estimates, particularly when used in combination with different driving-GCMs. Finally, an appropriate temporal and spatial resolution for RCM data must be chosen. Results based on 'smoothed' data, e.g. using 10 day totals or regions rather than grid cells, showed much less inter-model variability. Smoothing seems to reduce the influence of individual model characteristics, exaggerating precipitation patterns resulting from larger scale processes that are better resolved by RCMs in relation to precipitation resulting from processes operating at a higher temporal and spatial resolution.

The RCMs examined here indicate increases in extreme precipitation across the UK and most of Europe under global warming but considerable uncertainty as to the magnitude of change. The use of multi-model ensembles to assess the impacts of climate change offers considerable potential but also a significant challenge, for both resource planners and managers and for the research community in communicating the nature of these uncertainties.

Acknowledgments

Data have been provided through the PRUDENCE data archive, funded by the EU through contract EVk2-CT2001-00132. Data is available to download from http://prudence.dmi.dk/. The UKMO 5 km datasets were created with financial support from the Department of Environment, Food and Rural Affairs and are being promoted through the UK Climate Impacts Programme (UKCIP). They form part of the UKCIP02 national climate scenarios prepared for UKCIP by the Tyndall and Hadley Centres. This work was partially supported by the European Union FP6 Integrated Project AquaTerra (Project no. 505428) under the thematic priority sustainable development, global change and ecosystems and by a NERC Postdoctoral Fellowship award to Dr. Hayley Fowler (2006-2009) NE/D009588/1. Thanks to Lucy Manning and Aidan Burton (Newcastle University) for advice on the use of the R statistical software package and to the two anonymous reviewers for helpful comments that helped to improve the manuscript.

References

Alexander, L. V., X. Zhang, T. C. Peterson, J. Caesar, B. Gleason, A. M. G. Klein Tank, M. Haylock, D. Collins, B. Trewin, F. Rahimzadeh, A. Tagipour, P. Ambenje, K. Rupa Kumar, J. Revadekar and G. Griffiths (2006), Global observed changes in daily climate extremes of temperature and precipitation, *J. Geophys. Res., 111*, doi:10.1029/2005JD006290.

Allen, M. R. and W. J. Ingram (2002), Constraints on future changes in climate and the hydrologic cycle, *Nature*, *419*, 224-232.

Beniston, M., D. B. Stephenson, O. B. Christensen, C. A. T. Ferro, C. Frei, S. Goyette, K. Halsnaes, T. Holt, K. Jylhä, B. Koffi, J. Palutikof, R. Schöll, T. Semmler and K. Woth (2007), Future Extreme Events in European Climate: An Exploration of Regional Climate Model Projections. *Climatic Change, 81(Supplement 1),* 81-95.

Blenkinsop, S. and H. J. Fowler (2007), Changes in drought characteristics for Europe projected by the PRUDENCE regional climate models, *Int. J. Climatol., in press.*

Brunetti, M., L. Buffoni, M. Maugeri and T. Nanni (2000), Precipitation intensity trends in northern Italy*, Int. J. Climatol., 20,* 1017-1031.

Buonomo, E., R. Jones, C. Huntingford and J. Hannaford (2007), On the robustness of changes in extreme precipitation over Europe from two high resolution climate change simulations, *Q. J. R. Meteorol. Soc., 133,* 65–81.

Christensen, J. H. and O. B. Christensen (2003), Severe summertime flooding in Europe, *Nature, 421,* 805-806.

Christensen, J. H. and O. B. Christensen (2004), Intensification of extreme European summer precipitation in a warmer climate, *Global Planet. Change, 44,* 107-117.

Christensen, J. H., T. R. Carter, M. Rummukainen and G. Amanatidis (2007), Evaluating the performance and utility of regional climate models: the PRUDENCE project, *Climatic Change, 81(Supplement 1),* 1-6.

Christensen, O. B., J. H. Christensen, B. Machenhauer, and M. Botzet (1998), Very high-resolution regional climate simulations over Scandinavia: Present Climate, *J. Climate, 11,* 3204-3229.

Christensen, J. H., O. B. Christensen, P. Lopez, E. Van Meijgaard and M. Botzet (1996), The HIRHAM4 Regional Atmospheric Climate Model, *Scientific Report 96-4*, 51 pp. Available from DMI, Lyngbyvej 100, Copenhagen Ø.

Christensen, J. H., O. B. Christensen and J. P. Schultz (2001), High resolution physiographic data set for HIRHAM4: An application to a 50 km horizontal resolution domain covering Europe, *DMI Technical Report 01-15*, Available from DMI, Lyngbyvej 100, Copenhagen Ø.

Collins, M. and M. R. Allen (2002), Assessing the relative roles of initial and boundary conditions in interannual to decadal climate predictability, *J. Climate, 15,* 3104-3109.

Covey, C., K. M. Achuta Rao, U. Cubasch, P. Jones, S. J. Lambert, M. E. Mann, T. J. Phillips and K. E. Taylor (2003), An overview of results from the Coupled Model Intercomparison Project, *Global Planet. Change*, *37,* 103-133.

Déqué M. (2007), Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values, *Global and Planetary Change, 57,* 16–26.

Déqué M., P. Marquet and R.G. Jones (1998), Simulation of climate change over Europe using a global variable resolution general circulation model, *Climate Dynam., 14*, 173-189.

Déqué, M., D. P. Rowell, D. Lüthi, F. Giorgi, J. H. Christensen, B. Rockel, D. Jacob, E. Kjellström, M. de Castro and B. van den Hurk (2007), An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections, *Climatic Change, 81(Supplement 1),* 53-70.

Döscher, R., U. Willén, C. Jones, A. Rutgersson, H. E. M. Meier, U. Hansson and L. P. Graham (2002), The development of the coupled regional ocean-atmosphere model RCAO, *Boreal Env. Res., 7,* 183-192.

Easterling, D. R., G. A. Meehl, C. Parmesan, S. A. Changnon, T. R. Karl and L. O. Mearns (2000), Climate extremes: Observations, modeling, and impacts, *Science*, *289*, 2068-2074.

Efron, B. (1979), Bootstrap methods: another look at the jack-knife, *Annals of Statistics*, 7, 1–26.

Efron, B. and R. J. Tibshirani (1993), *An Introduction to the Bootstrap*, 456pp, Chapman and Hall, New York.

Ekström, M., H. J. Fowler, C. G. Kilsby and P. D. Jones (2005), New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1. Future estimates and use in impact studies, *J. Hydrol., 300,* 234-251.

Ekström, M., B. Hingray, A. Mezghani and P. D. Jones (2007), Regional climate model data used within the SWURVE project 2: addressing uncertainty in regional climate model data for five European case study areas, *Hydrol. Earth Syst. Sci., 11(3),* 1085-1096.

Fowler, H. J. and C. G. Kilsby (2003a), A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000, *Int. J. Climatol., 23,* 1313-1334.

Fowler, H. J. and C. G. Kilsby (2003b), Implications of changes in seasonal and annual extreme rainfall, *Geophys. Res. Lett., 30,* 1720, doi:10.1029/2003GL017327.

Fowler, H. J., M. Ekström, C. G. Kilsby and P. D. Jones (2005), New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1. Assessment of control climate, *J. Hydrol., 300,* 212-233.

Frei, C. and C. Schär (2001), Detection probability of trends in rare events: Theory and application to heavy precipitation in the Alpine region, *J. Climate, 14,* 1568-1584.

Frei, C., J. H. Christensen, M. Déqué, D. Jacob, R. G. Jones and P. L. Vidale (2003), Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps, *J. Geophys. Res., 108(D3),* 4124, doi:10.1029/2002JD002287.

Frei, C., R. Schöll, S. Fukutome, J. Schmidli and P. L. Vidale (2006), Future change of precipitation extremes in Europe: An intercomparison of scenarios from regional climate models, *J. Geophys. Res., 111,* D06105, doi:10.1029/2005JD005965.

Frich, P., L. V. Alexander, P. Della-Marta, B. Gleason, M. Haylock, A. M. G. K. Tank and T. Peterson (2002), Observed coherent changes in climatic extremes during the second half of the twentieth century, *Climate Res., 19,* 193–212.

Furrer, R., S. R. Sain, D. Nychka and G. A. Meehl (2007), Multivariate Bayesian Analysis of Atmosphere-Ocean General Circulation Models, *Environmental and Ecological Statistics, in press.*

Giorgi, F., B. Hewitson, J. Christensen, M. Hulme, H. Von Storch, P. Whetton, R. Jones, L. Mearns and C. Fu (2001), Chapter 10. Regional Climate Information – Evaluation and Projections, in *Climate change 2001: The scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*., edited by J. T. Houghton et al., pp. 583-638, Cambridge University Press.

Gibelin, A. L. and M. Déqué (2003), Anthropogenic climate change over the Mediterranean region simulated by a global variable resolution model, *Climate Dynam.*, *20,* 327-339.

Gordon, C., C. Cooper, C. A. Senior, H. Banks, J. M. Gregory, T. C. Johns, J. F. B. Mitchell and R. A. Wood (2000), The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments, *Climate Dynam., 16,* 147–166

Greene, A.M., L. Goddard and U. Lall (2006), Probabilistic multi-model regional temperature change projections, *J. Climate, 19,* 4326-4343.

Hagemann, S., M. Botzet and B. Machenhauer (2001), The summer drying problem over south-eastern Europe: Sensitivity of the limited area model HIRHAM4 to improvements in physical parametrization and resolution, *Phys. Chem. Earth B, 26,* 391–396.

Hagemann, S., B. Machenhauer, R. Jones, O. B. Christensen, M. Déqué, D. Jacob and P. L. Vidale (2004), Evaluation of water and energy budgets in regional climate models applied over Europe, *Climate Dynam., 23*, 547–567.

Haylock, M. R., G. C. Cawley, C. Harpham, R. L. Wilby and C. M. Goodess (2006), Downscaling heavy precipitation over the United Kingdom: A comparison of dynamical and statistical methods and their future scenarios, *Int. J. Climatol*., *26,* 1397-1415.

Hosking, J. R. M. and J. R. Wallis (1988), The effect of intersite dependence on regional flood frequency analysis, *Water Resour. Res., 24,* 588-600.

Hosking, J. R. M. and J. R. Wallis (1997), *Regional frequency analysis: an approach based on L-moments*, 224 pp., Cambridge University Press, Cambridge.

Hudson, D. A. and R. G. Jones (2002), Regional climate model simulations of present-day and future climates of southern Africa, *Hadley Centre Technical Note No 39*, Met. Office, Exeter, UK.

Huntingford, C., R. G. Jones, C. Prudhomme, R. Lamb, H. H. C. Gash and D. A. Jones (2003), Regional climate-model predictions of extreme rainfall for a changing climate, *Q. J. Roy. Meteorol. Soc*., 129, 1607-1621.

Jacob, D., L. Bärring, O. B. Christensen, J. H. Christensen, M. de Castro, M. Déqué, F. Giorgi, S. Hagemann, M. Hirschi, R. Jones, E. Kjellström, G. Lenderink, B. Rockel, E. Sánchez, C. Schär, S. I. Seneviratne, S. Somot, A. van Ulden and B. van den Hurk (2007), An inter-comparison of regional climate models for Europe: model performance in present-day climate, *Climatic Change, 81(Supplement 1),* 31-52.

Jenkins, G. and J. Lowe (2003), Handling uncertainties in the UKCIP02 scenarios of climate change, *Hadley Centre Technical note 44*.

Johns, T. C., J. M. Gregory, W. J. Ingram, C. E. Johnson, A. Jones, J. A. Lowe, J. F. B. Mitchell, D. L. Roberts, D. M. H. Sexton, D. S. Stevenson, S. F. B. Tett and M. J. Woodage (2003), Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emission scenarios, *Climate Dynam*., *20,* 583-612.

Jones, P.D. and P.A. Reid (2001), Assessing future changes in extreme precipitation over Britain using regional climate model integrations, *Int. J. Climatol.*, *21*, 1337-1356.

Jones, R. G., M. Noguer, D. C. Hassell, D. Hudson, S. S. Wilson, G. J. Jenkins and J. F. B. Mitchell (2004a), Generating high resolution climate change scenarios using PRECIS, 35pp, Tech. report available from Met. Office, Hadley Centre, Exeter, UK.

Jones, C. G., A. Ullerstig, U. Willén and U. Hansson (2004b), The Rossby Centre regional atmospheric climate model (RCA). Part I: Model climatology and performance characteristics for present climate over Europe, *Ambio*, *33,* 199-210.

Jones R. G., J. M. Murphy, D. C. Hassell and M. J. Woodage (2005), A high resolution atmospheric GCM for the generation of regional climate scenarios, *Hadley Centre Technical Note 63,* Met Office, Exeter.

Klein Tank, A. M. G. and G. P. Können (2003), Trends in indices of daily temperature and precipitation extremes in Europe, 1946-1999, *J. Climate*, *16,* 3665-3680.

Lopez, A., C. Tebaldi, M. New, D. Stainforth, M. Allen and J. Kettleborough (2006), Two approaches to quantifying uncertainty in global temperature changes, *J. Climate*, *19(19),* 4785-4796

May, W. (2007), The simulation of the variability and extremes of daily precipitation over Europe by the HIRHAM regional climate model, *Global and Planetary Change*, *57*, 59–82.

Mearns, L. O., and the NARCCAP Team (2006), Overview of the North American Regional Climate Change Assessment Program. *NOAA RISA-NCAR Meeting*, Tucson, AZ, March 2006.

Mearns, L. O., M. Hulme, T. R. Carter, R. Leemans, M. Lal and P. H. Whetton (2001), Climate scenario development. Chapter 13 in *Climate Change 2001: The Scientific Basis*, edited by J. Houghton et al., Intergovernmental Panel on Climate Change, Cambridge University Press.

Meehl G. A., J. M. Arblaster and C. Tebaldi (2005), Understanding future patterns of increased precipitation intensity in climate model simulations, *Geophys. Res. Lett.*, *32*, L18719, doi:10.1029/2005GL023680.

Meier, H. E. M., R. Döscher and T. Faxén (2003), A multiprocessor coupled ice-ocean model for the Baltic Sea. Application to the salt inflow, *J. Geophys. Res.*, *108(C8)*, 3273, doi:10.1029/2000JC000521.

Nakićenović, N., J. Alcamo, G. Davis, H. J. M. de Vries, J. Fenhann, S. Gaffin, K. Gregory, A. Grubler, T. Y. Jung, T. Kram, E. L. La Rovere, L. Michaelis, S. Mori, T. Morita, W. Papper, H. Pitcher, L. Price, K. Riahi, A. Roehrl, H-H. Rogner, A. Sankovski, M. Schlesinger, P. Shukla, S. Smith, R. Swart, S. van Rooijen, N. Victor, and Z. Dadi (2000), *Emissions Scenarios. A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*, 559 pp., Cambridge University Press, Cambridge.

Osborn, T. J., and M. Hulme (1997), Development of a relationship between station and grid-box rainday frequencies for climate model evaluation, *J. Climate, 10,* 1885–1908.

Pal, J. S., F. Giorgi and X. Bi (2004), Consistency of recent European summer precipitation trends and extremes with future regional climate projections, *Geophys. Res. Lett., 31*, L13202, doi:10.1029/2004GL019836.

Palmer, T.N. and J. Räisänen (2002), Quantifying the risk of extreme seasonal precipitation events in a changing climate, *Nature, 415,* 512.

Perry, M. and D. Hollis (2005a), The development of a new set of long-term climate averages for the UK, *Int. J. Climatol., 25,* 1023-1039.

Perry, M. and D. Hollis (2005b), The generation of monthly gridded datasets for a range of climatic variables over the UK, *Int. J. Climatol., 25,* 1041-1054.

Pope, V. D., M. L. Gallani, P. R. Rowntree and R. A. Stratton (2000), The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3, *Climate Dynam., 16,* 123-146.

Räisänen, J., U. Hansson, A. Ullerstig, R. Döscher, L. P. Graham, C. Jones, M. Meier, P. Samuelsson and U. Willén (2004), European climate in the late 21st century: regional simulations with two driving global models and two forcing scenarios, *Climate Dynam., 22,* 13-31.

Roeckner, E., K. Arpe, L. Bengtsson, M. Christoph, M. Claussen, L. Dümenil, M. Esch, M. Giorgetta, U. Schlese, and U. Schulzweida (1996), The atmospheric general circulation model ECHAM4: Model description and simulation of present-day climate. *Max Planck Institut für Meteorologie, Report No. 218,* pp. 90, Hamburg, Germany.

Roeckner, E., L. Bengtsson, J. Feichter, J. Lelieveld, and H. Rodhe (1999), Transient climate change simulations with a coupled Atmosphere–Ocean GCM including the tropospheric sulfur cycle, *J. Climate*, *12*, 3004–3032.

Rummukainen, M., J. Räisänen, B. Bringfelt, A. Ullerstig, A. Omstedt, U. Willen, U. Hansson and C. Jones (2001), A regional climate model for northern Europe: model description and results from the downscaling of two GCM control simulations, *Climate Dynam*., *17*, 339-359.

Stratton, R. A. (2004), Report on aspects of variability in high-resolution versions of HadAM3. *Hadley Centre Technical Note No. 53*, pp. 32, Hadley Centre for Climate Prediction and Research, U.K. Meteorological Office.

Tebaldi, C., L. O. Mearns, D. Nychka and R. L. Smith (2004), Regional probabilities of precipitation change: A Bayesian analysis of multimodel simulations, *Geophys. Res. Lett.*, *31*, L24213, doi:10.1029/2004GL021276.

Tebaldi, C., R. L. Smith, D. Nychka and L. O. Mearns (2005), Quantifying uncertainty in projections of regional climate change: a Bayesian approach to the analysis of multi-model ensembles, *J. Climate*, *18*, 1524-1540.

Tebaldi, C., K. Hayhoe, J. M. Arblaster, G. A. Meehl (2006), Going to the extremes; An intercomparison of model-simulated historical and future changes in extreme events, *Climatic Change, 79,* 185-211.

Wigley, T. M. L., J. M. Lough and P. D. Jones (1984), Spatial patterns of precipitation in England and Wales and a revised, homogeneous England and Wales precipitation series, *J. Climatol.*, *4,* 1-25.

Wilks D.S. (1995), *Statistical methods in the atmospheric sciences*, Academic Press.

List of Figures

Figure 1 The RCM 50 km grid (all re-gridded to the common CRU grid) and the nine coherent rainfall regions. The regions are: North Scotland (NS), East Scotland (ES), South Scotland (SS), Northern Ireland (NI), Northwest England (NWE), Northeast England (NEE), Central and Eastern England (CEE), Southeast England (SEE) and Southwest England (SWE).

Figure 2 Observed (UKMO) and RCM modeled (CTRL) (a) mean daily precipitation, (b) wet day frequency (> 1 mm), (c) spatial distribution of mean winter precipitation for the UK for the control period, 1961-1990. The shaded area in (a) represents the 95% confidence interval for the 30-year sample mean, and in (b) indicates the 95% confidence interval for the sample proportion assuming the Gaussian approximation to the binomial distribution as described in Wilks (1995).

Figure 3 Scatter plots of the mean and standard deviation of Rmed values, comparing the UKMO and CTRL integrations for (a) 1 day, and (b) 10 day. The CTRL datasets are denoted by different colors, UKMO is shown in bold black, whilst symbols represent regions.

Figure 4 Box plots of distributions of Rmed for CTRL integrations and observed (UKMO) for (a) CEE region, 1 day, (b) CEE region, 10 day, (c) NS region, 1 day, and (d) NS region, 10 day. The observed distribution is shown in bold.

Figure 5 Estimates of return value (in mm) for (a) 1 day, 5 year event for grid cells, (b) 1 day, 5 year event for regions, and (c) 10 day, 5 year event for grid cells.

Figure 6 Estimates of return value (in mm) for 1 day, 5 year event for all European grid cells.

Figure 7 Estimates of percentage change in return values for (a) 1 day, 5 year events and (b) 1 day, 25 year events for all European grid cells.

Figure 8 Estimates of percentage change in return values for (a) 10 day, 5 year events and (b) 10 day, 25 year events for all European grid cells.

Figure 9 Estimates of change in 1 day 5 year return value for the SRES A2 2071-2100 scenario for each of the nine UK homogenous rainfall regions for six RCMs. The x-axis shows the percentage change estimated and the y-axis shows the probability density. The uncertainty resulting from natural variability is shown by the width of the density function.

Figure 10 Estimates of change in 10 day 25 year return value for the SRES A2 2071-2100 scenario for each of the nine UK homogenous rainfall regions for six RCMs. The x- and y-axis labels are as for Figure 9. The uncertainty resulting from natural variability is shown by the width of the density function.

Figure 11 Estimates of change in 5 and 25 year return values for the SRES A2 2071-2100 scenario for each of the nine UK homogenous rainfall regions, pooling results from all RCMs and assuming equal weighting: (a) 1 day, and (b) 10 day. The box plot shows the smallest observation (lower bar), lower quartile (bottom of box), median (line through box), upper quartile (top of box), and largest observation (upper bar). Outliers, points which fall more than 1.5 times the inter-quartile range above the third quartile or below the first quartile, are indicated individually.

Figure 12 Estimates of change in 1 day (a) 5 year, and (b) 25 year return values for the SRES A2 2071-2100 scenario for each of the nine UK homogenous rainfall regions, pooling results from all Hadley-driven models (Hadley), a sub-set of Hadley-driven models (Hadley_sub), RCAOH and HIRHAMH, and the ECHAM-driven models (ECHAM), RCAOE and HIRHAME. Box plot details are the same as for Figure 11.

Figure 13 Estimates of change in 10 day (a) 5 year, and (b) 25 year return values for the SRES A2 2071-2100 scenario for each of the nine UK homogenous rainfall regions, pooling results from all Hadley-driven models (Hadley), a sub-set of Hadley-driven models (Hadley_sub), RCAOH and HIRHAMH, and the ECHAM-driven models (ECHAM), RCAOE and HIRHAME. Box plot details are the same as for Figure 11.

Table 1 The selection of PRUDENCE Regional Climate Models for which integrations are analyzed in this study. The first part of each model acronym refers to the RCM and the second to the GCM data used to provide the boundary conditions, either from Hadley Centre models (HadRM3H/P or HadCM3; suffix H) or ECHAM4/OPYC3 (suffix E).

Model Acronym	**Institution**	**RCM**	**GCM Driving Data**
HIRHAMH	Danish Meteorological Institute (DMI)	HIRHAM	HadAM3H
HIRHAME			ECHAM4/OPYC3
RCAOH	Swedish Meteorological and Hydrological Institute (SMHI)	RCAO	HadAM3H
RCAOE			ECHAM4/OPYC3
HADH	Hadley Centre – UK Meteorological Office	HadRM3P	HadAM3P
ARPEGEH	Météo-France, France	ARPEGE	HadCM3

Table 2 Percentage differences in mean and standard deviation of 1 day AM between UKMO and CTRL integrations for each UK region. UKMO values are given in mm.

Statistic	UKMO	ARPEGE_H	HAD_H	HIRHAM_E	HIRHAM_H	RCAO_E	RCAO_H
(a) Mean							
SEE	32.1	-20.1	-20.9	-19.5	-26.8	-21.3	-23.4
SWE	37.6	-25.1	-25.2	-22.5	-31.3	-22.6	-27.0
CEE	29.5	-17.2	-8.0	-16.0	-27.0	-15.1	-18.8
NEE	33.2	-18.9	-14.7	-25.4	-29.2	-24.2	-27.1
NWE	38.7	-26.3	-24.8	-24.8	-31.4	-28.0	-32.5
SS	40.9	-24.0	-12.5	-27.3	-37.0	-31.5	-33.9
ES	34.6	-18.1	-17.9	-27.8	-34.7	-25.2	-30.1
NS	44.4	-30.0	-14.4	-26.5	-38.2	-34.5	-40.1
NI	33.6	-15.9	-13.9	-14.7	-27.9	-19.8	-24.1
Mean anomaly		-21.7	-16.9	-22.7	-31.5	-24.7	-28.6
(b) Standard Deviation							
SEE	9.7	-5.5	-16.4	-27.5	-16.9	-42.7	-34.0
SWE	12.2	-42.7	-31.4	-33.6	-27.4	-48.6	-47.8
CEE	10.1	-33.4	-13.7	-32.9	-46.8	-42.9	-35.5
NEE	10.5	-39.0	-16.0	-41.8	-31.8	-55.7	-46.8
NWE	12.7	-47.8	-28.9	-38.3	-35.8	-54.7	-56.4
SS	11.7	-53.1	-13.8	-40.5	-30.5	-62.1	-55.5
ES	10.4	-33.2	-18.7	-41.1	-38.9	-41	-56.1
NS	17.3	-53.0	-9.3	-42.2	-48.0	-66.6	-63.3
NI	10.9	-45.0	-36.9	-41.4	-51.4	-59.6	-46.0
Mean anomaly		-39.2	-20.6	-37.7	-36.4	-52.7	-49.0

Table 3 Percentage differences in mean and standard deviation of 10 day AM between UKMO and CTRL integrations for each UK region. UKMO values are given in mm.

Statistic	UKMO	ARPEGE_H	HAD_H	HIRHAM_E	HIRHAM_H	RCAO_E	RCAO_H
(a) Mean							
SEE	85.5	-12.7	-26.7	-7.3	-10.8	-1.7	0.2
SWE	111.0	-15.7	-25.7	-12.1	-22.2	-5.5	-8.8
CEE	68.6	-7.2	-4.4	8.5	-6.8	16.0	10.4
NEE	86.1	-6.7	-17.3	-13.8	-21.8	0.4	-9.1
NWE	114.3	-11.1	-26.0	-16.5	-23.1	-10.0	-15.9
SS	135.0	-16.1	-18.0	-19.8	-30.0	-17.4	-21.8
ES	93.9	-9.7	-15.7	-12.8	-19.2	-3.6	-8.3
NS	154.0	-19.1	-16.5	-20.8	-34.4	-20.2	-26.1
NI	93.0	-8.7	-11.1	-0.3	-12.8	8.7	-0.8
Mean anomaly		-11.9	-17.9	-10.5	-20.1	-3.7	-8.9
(b) Standard Deviation							
SEE	23.5	-30.5	-34.6	-1.2	-16.2	-21.7	-26.2
SWE	29.3	-26.4	-22.8	-10.3	-28.1	-38.2	-44.0
CEE	17.1	-31.6	-4.3	7.6	-28.8	-2.8	6.6
NEE	26.2	-27.6	-1.5	-40.2	-40.5	-38.5	-47.5
NWE	37.3	-35.4	-35.4	-41.5	-32.1	-47.5	-58.8
SS	40.9	-49.0	-17.4	-32.7	-38.2	-57.8	-61.9
ES	30.8	-30.2	-17.7	-45.6	-36.2	-47.0	-49.6
NS	58.6	-46.1	1.7	-39.3	-47.8	-62.0	-60.9
NI	17.0	-7.6	1.2	3.5	-9.1	-10.2	-16.7
Mean anomaly		-31.6	-14.5	-22.2	-30.8	-36.2	-39.9