INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **29**: 385–416 (2009)
Published online 6 January 2009 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/joc.1827

RMetS
Royal Meteorological Society

Multi-model ensemble estimates of climate change impacts on UK seasonal precipitation extremes

H. J. Fowler[a]* and M. Ekström[b]

[a] *Water Resource Systems Research Laboratory, School of Civil Engineering and Geosciences, Newcastle University, Newcastle upon Tyne, UK*
[b] *School of Geography, Archaeology and Earth Resources, University of Exeter, Exeter, UK*

ABSTRACT: Thirteen regional climate model (RCM) integrations from the Prediction of Regional Scenarios and Uncertainties for Defining European Climate change risks and Effects (PRUDENCE) ensemble are used together with extreme value analysis to assess changes to seasonal precipitation extremes in nine UK rainfall regions by 2070–2100 under the SRES A2 emissions scenario. Model weights are based on similarities between observed and modelled UK extreme precipitation calculated for a combination of (1) spatial characteristics: the semi-variogram parameters sill and range, and (2) the discrepancy in the regional median seasonal maxima. These weights are used to combine individual RCM bootstrap samples to provide multi-model ensemble estimates of percent change in the return value magnitudes of regional extremes. The contribution of global climate model (GCM) and RCM combinations to model structural uncertainty is also investigated. The multi-model ensembles project increases across the UK in winter, spring and autumn extreme precipitation; although there is uncertainty in the absolute magnitude of increases, these range from 5 to 30% depending upon region and season. In summer, model predictions span the zero change line, although there is low confidence due to poor model performance. RCM performance is shown to be highly variable; extremes are well simulated in winter and very poorly simulated in summer. The ensemble distributions are wider (projections are more uncertain) for shorter duration extremes (e.g. 1 day) and higher return periods (e.g. 25 year). There are rather limited differences in the weighted and unweighted multi-model ensembles, perhaps a consequence of the lack of model independence between ensemble members. The largest contribution to uncertainty in the multi-model ensembles comes from the lateral boundary conditions used by RCMs included in the ensemble. Therefore, the uncertainty bounds shown here are conservative despite the relatively large number of RCMs contributing to the multi-model ensemble distribution. Copyright © 2009 Royal Meteorological Society

KEY WORDS precipitation; extremes; seasonal maxima; regional climate models; ensemble; probabilities; climate change; UK

Received 31 October 2008; Accepted 4 November 2008

1. Introduction

Under enhanced greenhouse conditions, it is likely that projected changes to climatic and hydrologic extremes will have the largest impacts on human society (Tebaldi *et al*., 2006). The high latitudes of the Northern Hemisphere are currently experiencing a trend towards increased precipitation and enhanced variability (e.g. Meehl *et al*., 2005; Alexander *et al*., 2006; Trenberth *et al*., 2007), particularly in winter; observational analyses in Europe suggest significant positive trends in precipitation intensities over the past decade (e.g. Brunetti *et al*., 2000; Frei and Schär, 2001; Fowler and Kilsby, 2003a,b). Global climate model (GCM) studies consistently suggest that increases in the frequency and intensity of heavy precipitation are likely under enhanced greenhouse conditions (Tebaldi *et al*., 2006; Meehl *et al*., 2007), even in regions that may experience a reduction in mean precipitation (Frei *et al*., 2006).

The assessment of global and regional climate model simulations and future projections of extremes is now relatively common, yet few studies have examined the uncertainties introduced by using multiple climate model outputs (Fowler *et al*., 2007a), perhaps due to a scarcity of suitable datasets. For example, Frei *et al*. (2006) and Beniston *et al*. (2007) compared projections of change in extreme precipitation for multiple regional climate models (RCMs) across Europe, but did not combine them into a multi-model ensemble estimate. However, a variety of applications in other fields have demonstrated that combining models through a multi-model ensemble generally increases the skill, reliability and consistency of predictions (Tebaldi and Knutti, 2007). Indeed, Kendon *et al*. (2008) suggest that three models is the minimum useful ensemble size when considering changes in precipitation extremes due to the large influence of natural variability found amongst individual ensemble members.

In addition to natural variability, other sources of uncertainties must be considered within the climate modelling process for GCMs, e.g. grid resolution, process parameterization, model structure and emissions scenario (e.g. Giorgi and Francisco, 2000; Covey *et al*.,

* Correspondence to: H. J. Fowler, Water Resource Systems Research Laboratory, School of Civil Engineering and Geosciences, Cassie Building, Newcastle University, *Newcastle upon Tyne* NE1 7RU, UK. E-mail: h.j.fowler@ncl.ac.uk

2003). When using RCM data the sources of uncertainty increase, as outputs are influenced by RCM resolution, numerical scheme, physical parameterizations and the forcing lateral boundary conditions (Rummukainen *et al*., 2001; Déqué *et al*., 2007; Elía *et al*., 2008). In a recent application by the authors, Fowler *et al*. (2007b) applied equal weighting to combine the results of six RCM integrations from the Prediction of Regional scenarios and Uncertainties for Defining European Climate change risks and Effects (PRUDENCE) (Christensen *et al*., 2007) ensemble. This first attempt to produce probabilistic estimates of change in annual extreme precipitation was tested on nine climatically homogenous rainfall regions in the UK, generating larger uncertainty ranges than studies using integrations from a single RCM (e.g. Ekström *et al*., 2005) and demonstrating the large influence of GCM lateral boundary conditions on model projections; although all models projected increases in extreme precipitation, ECHAM4/OPYC-driven RCMs projected increases 20% higher on average than HadAM3H-driven RCMs.

In this paper we extend the approach taken by Fowler *et al*. (2007b) to derive regional multi-model ensemble estimates for future change in UK precipitation extremes by: (1) performing the analysis on a seasonal instead of an annual temporal resolution; (2) developing a model-specific weighting scheme. Recent approaches to the construction of probability density functions (PDFs) of change, 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. To date, most weighting methods have been based on model skill in simulating areal average climate characteristics. However, Blenkinsop and Fowler (2007) noted that RCMs can skilfully simulate areal average precipitation yet show poor skill in simulating its spatial distribution; the same holds true for extremes (Fowler *et al*., 2007b). Skilful simulation of areal averages may hide compensating biases; therefore models that produce errors with a less well-defined spatial structure may artificially score better at representing physical precipitation processes.

In this paper we therefore develop a set of region-specific weights for each RCM depending on its similarity in spatial characteristics and overall magnitude of extreme precipitation statistics to its observed counterpart. The weights are a combination of a discrepancy measure (region specific) and a measure that reflects the overall similarity in spatial dependence structure between observed and RCM extreme precipitation fields (UK-wide). Thus, each model weight is conditioned on both its closeness to observed extreme values in a particular region and how well it captures the observed spatial variability in precipitation extremes across the UK.

To describe UK-wide spatial characteristics, it is necessary to use a measure that is not sensitive to phase errors, i.e. errors that arise due to discrepancies in spatial location (or temporal lag) of higher and lower values between fields of comparison. For example, if a particular RCM models correctly the magnitude of extremes but at a slightly shifted position to that shown in the observed field, the RCM will be given a poor score when using a measure such as root-mean-square error. For this application, where it is unreasonable to expect RCMs to simulate extreme precipitation magnitudes with high precision on a grid cell resolution, a flexible measure is required to identify which RCMs have poor representations of regional variability in comparison with observations. For such a task the semi-variogram is appropriate as it quantifies the spatial dependence structure of a variable in space.

Specifically, the semi-variogram is defined as the variance between two values as a function of the separation distance (lag) between them and is traditionally used to estimate variables at previously unsampled locations, with numerous examples of applications to precipitation fields (e.g. Pardo-Igúzquiza, 1998; Goovaerts, 2000; Kyriakidis *et al*., 2001; Pardo-Igúzquiza *et al*., 2006). However, the semi-variogram can also be used to investigate particular spatial properties of precipitation (e.g. Holawe and Dutter, 1999; Germann and Joss, 2001). Here, the semi-variogram is calculated for extreme precipitation fields to give a summary of UK-wide characteristics in terms of spatial variability and range of dependence structure. This information is contained in the parameters (partial sill and range) of the semi-variogram model (a two-dimensional function fitted to each of the experimental semi-variograms) and these are then used in combination with the regional discrepancy measure to provide model-specific weights. Thus, the overall aim of this study is to create regional multi-model ensemble distributions for each season showing projected change to the 5- and 25-year return values of long- and short-duration precipitation extremes using the weights developed for each region and RCM.

The paper is divided into the following sections: Section 2 presents the RCM and observed data used in the study; Section 3 introduces the statistical methods used for analysis; Section 4 presents an evaluation of the RCMs' ability to simulate extreme seasonal precipitation over the UK and their future projections; Section 5 explores how probabilistic estimates of change in seasonal extremes may be developed for homogeneous rainfall regions in the UK using simple weighting methods and compares these to unweighted multi-model estimates; and Section 6 provides a discussion of the results and concludes the study.

2. Model output and observations

2.1. Models

The PRUDENCE ensemble contains daily outputs for a range of climatic variables for control (1961–1990; CTRL) and future (2071–2100; SCEN) RCM integrations. We use 13 integrations, 12 from the PRUDENCE ensemble and an additional RCM driven by HadAM3H, METNO, under the IPCC SRES A2 emissions scenario (Nakićenović *et al*., 2000).

DOI: 10.1002/joc

Nine of these RCM integrations 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. A variable resolution global atmospheric model, Arpége, at approximately the same resolution as the RCMs, is nested directly into HadCM3. Two RCM integrations, HIRHAM and RCAO, are driven by lateral boundary conditions from the ECHAM4/OPYC3-coupled ocean–atmosphere GCM as well as by HadAM3H.

HadCM3 (Gordon *et al*., 2000; Johns *et al*., 2003) is a coupled ocean–atmosphere GCM at ∼300 km resolution and provides boundary conditions for HadAM3H/P. HadAM3H (Pope *et al*., 2000) and HadAM3P (Jones *et al*., 2005) have a resolution of ∼150 km in the mid-latitudes and can be considered as essentially the same model for Europe (Moberg and Jones, 2004). For 1961–1990, the HadAM3H/P models used an observed ocean state; for 2071–2100, the ocean state was constructed by adding anomalies from a transient simulation of HadCM3 for the SRES A2 emissions scenario to observations. The ECHAM4/OPYC3-coupled ocean–atmosphere GCM (Roeckner *et al*., 1996, 1999) at ∼300 km resolution was developed by the Max Planck Institute for Meteorology (MPI) and the German Climate Computing Centre (DKRZ). The HadAM3H/P and ECHAM4/OPYC3 global mean temperature responses are similar for the IPCC SRES A2 emissions scenario (3.1 and 3.56 °C, respectively; Tim Osborn, personal communication); mid-range in the global mean temperature responses for all GCMs presented by the IPCC Third Assessment Report (2001).

Details for all RCMs considered in this study are listed in Table I. All operate with grid spacing of ∼0.5° longitude by ∼0.5° latitude (∼50-km spatial resolution) over a European domain and outputs were re-gridded to a regular 0.5° × 0.5° grid using an inverse distance weighted interpolation algorithm to allow direct comparison between models. Suffixes E and H denote RCMs driven by ECHAM4/OPYC3 and HadAM3H/P/HadCM3 GCMs, respectively. More details on the experimental design of the PRUDENCE integrations can be found in Jacob *et al*. (2007).

2.2. Observations

An observational precipitation dataset at a comparable scale to the RCM outputs was produced by taking a daily average across the 5-km boxes contained within each 0.5° × 0.5° grid cell for each day of 1961–1990 for the UK Meteorological Office (UKMO) dataset (Perry and Hollis, 2005a,b).

3. Methods

3.1. Extreme value analysis and uncertainty estimates

The statistical analysis of extreme seasonal precipitation is based on daily precipitation totals and uses the same methods as detailed in Fowler *et al*. (2007b). The multi-model ensembles are generated using a combination of nonparametric bootstrapping and regional frequency analysis (RFA) for each of the nine UK rainfall regions (Figure 1) delineated by Wigley *et al*. (1984). The homogeneity of these regions for extreme precipitation was tested by Fowler and Kilsby (2003a). For each RCM integration seasonal maximum (SM) series of 1-, 2-, 5- and 10-day precipitation totals are extracted for each grid cell. These SM series are standardized by their median (Rmed – equivalent to the 2-year return value; following Fowler *et al*., 2005) to remove grid cell-specific factors from the regional analysis and to allow the regional pooling of standardized data for each region.

To represent regional uncertainty in extreme precipitation, the intra-regional variability is estimated using a nonparametric bootstrap re-sampling method (Efron, 1979). A set of 10 000 bootstrap samples are drawn from each pooled standardized regional SM dataset, to produce separate samples for each region and event duration. The literature on bootstrapping suggests that when re-sampling maxima or extreme order statistics (e.g. Zelterman, 1993; Shao and Tu, 1995), the size of a bootstrap sample should be smaller than that of the original sample. However, as the regionally pooled standardized SM distributions have relatively light tails, even though the generalized extreme value (GEV) distribution provides a good fit as proved by a quantile–quantile plot (not shown), a bootstrap sample of the same size as the original sample gives representative statistics for each region. Distributional agreement between the bootstrap samples and the original samples is tested using the two-sample Kolmogorov–Smirnov (KS) test which determines if two independent random samples are drawn from the same underlying continuous population. The null hypothesis in the KS test states that the two samples have similar empirical distribution functions and that there is no significant difference between them. For the boot strap samples of the same size as the original sample, the null hypothesis could not be rejected at $\alpha = 0.05$, whereas using the semi-parametric approach of Zelterman (1993) the KS test suggested that the null hypothesis is rejected for all bootstrap samples. Therefore, bootstrap samples of the same size as the original sample population are used.

A GEV distribution is then fitted to each bootstrap sample using L-moment ratios (Hosking and Wallis, 1997), and return values of precipitation intensities with average recurrence of 5 and 25 years are estimated for each region and event duration. The estimates are then re-scaled using the regionally averaged Rmed from the appropriate original SM dataset (see Appendix for more details on the RFA methodology). This technique is used to estimate regional distributions of return values for UKMO, CTRL and SCEN.

Multi-model estimates of change are then generated using the non-parametric bootstrap samples for each RCM and region for 1-, 2-, 5- and 10-day extremes following the methodology outlined in Fowler *et al*. (2007b). A random number generator is used to sample

DOI: 10.1002/joc

Table I. The 11 regional climate models (RCM) and 13 integrations used in this study.

Model acronym	Institution/model origin and references	RCM	Global climate model
ARPH	French Meteorological Service; ARPEGE/IFS variable resolution global model. Model: Déqué *et al*. (1998)	Arpège	HadCM3
HADH	Hadley Centre, UK Meteorological Office, Exeter; Regional model at the Hadley Centre. Model: Jones *et al*. (2004a).	HadRM3P	HadAM3P
HIRH	Danish Meteorological Institute, Copenhagen; Dynamical core from HIRLAM, Parameterizations from ECHAM4. Model: Christensen *et al*. (1996, 1998). Physiographic datasets: Christensen *et al*. (2001); Hagemann *et al*. (2001)	HIRHAM	HadAM3H
HIRE			ECHAM4/OPYC
RCAOH	Swedish Meteorological and Hydrological Institute, Norrköping; Rossby Centre Atmosphere Ocean Model. Model: Döscher *et al*. (2002); Meier *et al*. (2003); Jones *et al*. (2004b); Räisänen *et al*. (2004)	RCAO	HadAM3H
RCAOE			ECHAM4/OPYC
CHRMH	Swiss Federal Institute of Technology (ETH), Zurich; Climate High-Resolution Model. Model: Lüthi *et al*. (1996); Vidale *et al*. (2003)	CHRM	HadAM3H
CLMH	GKSS, Institute for Coastal Research, Geesthatcht, Germany; Climate version of 'Lokalmodell' of German Weather Service. Model: Steppeler *et al*. (2003)	CLM	HadAM3H
REMOH	Max Planck Institute for Meteorology, Hamburg, Germany; Dynamical core from 'Europamodell' of German Weather Service, Parameterizations from ECHAM4. Model: Roeckner *et al*. (1996); Jacob (2001)	REMO	HadAM3H
PROMH	Universidad Complutense de Madrid, Spain; Climate version of PROMES model. Model: Castro *et al*. (1993); Arribas *et al*. (2003).	PROMES	HadAM3H
REGH	The Abdus Salam International Centre for Theoretical Physics, Italy (ICTP); Dynamical core from MM5, Parameterizations from CCM3. Model: Giorgi *et al*. (1993a,b, 1999); Pal *et al*. (2000)	RegCM	HadAM3H
RACH	The Royal Netherlands Meteorological Institute (KNMI), Netherlands; Dynamical core from HIRLAM, Parameterizations from ECMWF physics. Model: Tiedtke (1989, 1993); Lenderink *et al*. (2003)	RACMO2	HadAM3H
METH	Norwegian Meteorological Institute; Version of HIRHAM. Model: Christensen *et al*. (2001); Hanssen-Bauer *et al*. (2003)	MetNo	HadAM3H

The acronyms are adopted here to provide an easier understanding of the format of each model run. The first part of each acronym refers to the RCM and the second to the GCM data used to provide the boundary conditions. All RCM integrations are from PRUDENCE except for METH.

the return values estimated for CTRL and SCEN from the 10 000 non-parametric bootstrap samples separately for each region, return period and event duration. The percentage change in the return value between CTRL and SCEN is then calculated for each one of these samples. Assuming that the models have equal skill, the 10 000 estimates of percentage change in return values from each RCM are then pooled for each region, return period and event duration 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.

3.2. Deriving model weights

Here, we present the method that is used to quantify differences in the spatial patterns of extreme precipitation simulated by RCMs and noted by Fowler *et al*. (2007b). Two measures of similarity are calculated and then combined into a single measure (*RV*, equal weighting is applied for the two measures); the first measure is region specific (R) and the second is representative of the entire UK-wide return value field (V). The combined weight is expressed as a percentage, and is used to determine how many bootstrap samples to take from each RCM for the multi-model-weighted ensemble. Thus, models which have greater skill in simulating the observed regional magnitude and spatial distribution of UK precipitation extremes provide a larger contribution to the multi-model ensemble distribution for each region.

3.2.1. UK-wide spatial-similarity weights

To quantify spatial similarity between fields of UKMO and CTRL precipitation extremes we use two parameters of the semi-variogram model, which is the modelled experimental semi-variogram. The experimental

DOI: 10.1002/joc

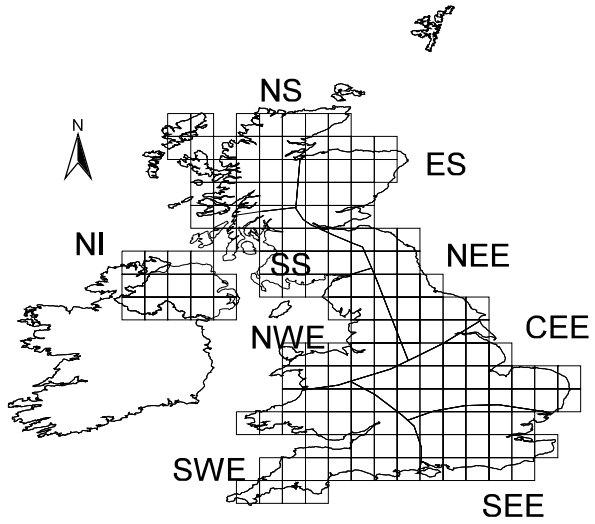

Figure 1. The re-gridded regular $0.5^\circ \times 0.5^\circ$ RCM grid and the nine coherent rainfall regions. The regions are North Scotland (NS), East Scotland (ES), South Scotland (SS), Northern Ireland (NI), Northwest England (NWE), Northeast England (NEE), central and eastern England (CEE), Southeast England (SEE) and Southwest England (SWE) (from Fowler *et al*., 2007b).

variogram is given by the mean of the halved squared increment as a function of distance for an intrinsic random function $Z(x)$ (Chilès and Delfiner, 1999):

$$2\gamma(h) = E\{[Z(x) - Z(x+h)]^2\} \quad (1)$$

where $2\gamma(h)$ is the expected squared difference between two data values at separation distance (or lag) h. In principle, the variogram shows how dissimilarity between $Z(x)$ and $Z(x+h)$ evolves with distance h, where $Z(x)$ represents an extreme value estimate at location x. The semi-variogram $[\gamma(h)]$ is one half of the variogram and is a measure of variability, which increases as samples become more dissimilar. The definition of the variogram assumes that the value Z is intrinsically stationary, which implies that the expectation of Z is constant through the spatial domain (here the UK) and that the semi-variance is dependent on the lag only and not on the position of x. In this application, both assumptions are likely to be violated as the expectation of Z is likely to vary over the domain as extremes vary in magnitude across the UK. Visual inspection of the seasonal UKMO experimental semi-variograms indicated a trend in the data at around 700–800 km, a feature that was also found in some, but not all, CTRL extreme value fields. However, these violations are not critical to the present study as the semi-variograms are simply used to summarize spatial characteristics in the extreme precipitation fields; they are not used for interpolating or simulating extreme rainfall. Since the return values are equally spaced over the entire domain, the resulting experimental semi-variogram is simply the average semi-variogram for the region and the semi-variance comprises, in addition to the sum of the semi-variogram, the squared difference between the means (systematic difference) (Germann and Joss, 2001). Furthermore, to ensure that distances in all directions contributed equally to the experimental semi-variogram, distances further than 330 km apart were excluded from the calculation as otherwise distances in the west–east direction would be underrepresented in relation to north–south distances (due to the shape of the UK).

The shape of the experimental semi-variogram can be described by theoretical models. Geo-statisticians use specific models that fulfil certain criteria, which become important when the models are used to interpolate or simulate the particular property in space (see Gringarten and Deutsch, 2001 for details). Commonly used models are the spherical, exponential, circular and Gaussian models. These models fulfil the required criteria and are defined by two parameters: the *sill* and the *range*. These parameters have a physical interpretation, where the sill is the semi-variance value that corresponds to zero correlation; the semi-variances below the sill indicate values that are positively correlated while semi-variances above the sill indicate negatively correlated values (Gringarten and Deutsch, 2001). The range is simply the distance associated with the sill, giving the distance at which the spatial auto-correlation between data point pairs, on average, ceases or becomes negatively correlated. Here, the sill and range are used in a crude way, simply indicating the variance and the average size of structures in the extreme value fields of UKMO and CTRL.

Fitted semi-variogram models can be split into two groups: those that reach a plateau and those that do not. For those that reach a plateau, that plateau is the sill and the distance at which it reaches the plateau is the range. For those that do not have a plateau, the models reach their sill asymptotically, and the range is an arbitrary 95% of the distance at which the sill is reached (Isaaks and Srivastava, 1989). When fitting the semi-variogram model to the experimental semi-variogram, the curve may not always have a zero intercept with the y-axis. If this is the case, the semi-variogram exhibits a *nugget effect* (c_0). The nugget is interpreted as non-spatial variation due to measurement error and variations in the data that relate to shorter ranges than the minimum sampled data spacing. The sill minus the nugget is sometimes known as the *partial sill* or *structural variance*.

To generate weights for each RCM, experimental semi-variograms were calculated for the gridded pattern of return values for UKMO and all CTRLs using only grid cells for which there is observational data. Prior to calculating the experimental semi-variograms, the precipitation field's grid coordinates were transformed to coordinates on a easting and northing grid reference system (metres) so that Euclidean distances between data pairs are calculated. Omni-directional semi-variograms (using combinations of data pairs in all compass directions) were calculated for the UKMO and CTRL extreme precipitation fields using a lag distance of ~60 km; the search radius for the experimental semi-variogram was cut off at 330 km. Four separate semi-variogram models

DOI: 10.1002/joc

(spherical, exponential, Gaussian and circular) were fitted using weighted least squares in geoR (Ribero and Diggle, 2001). The automated fitting procedure used a combination of over 100 different initial conditions (essentially first guesses of nugget, sill and range) to ensure a good fit. Additionally, visual inspection of all experimental semi-variograms and their respective best fit model was performed to ensure that the models captured the shape of the experimental semi-variogram. The sill and range for the best fitted model (based on the minimized weighted sum of squares) was retained for the calculation of individual RCM weights.

To identify which RCM extreme return value fields are more similar to their observed counterparts, we create a measure defined by the sill and range parameters of the fitted semi-variogram models. First the Euclidian distance is calculated between UKMO and CTRL using their respective sill and range parameters in a normalized parameter space. Parameters are normalized in order to avoid uneven influence by either parameter on the final semi-variogram weight. The distance is transformed to an inverse distance (Webster and Oliver, 2001) and then scaled to unity, providing model weights expressed as fractions. The RCM-specific weights, V, indicate which CTRLs are more similar to UKMO extreme precipitation patterns in terms of their scale of dependence and variance. The inverse distance weights imply that the contribution from models that are 'far away' in parameter space is relatively small.

3.2.2. Region-specific magnitude-discrepancy weights and combined weights

Although the model semi-variogram is able to capture spatial structure dependence, it does not reflect the magnitude of values in the field to which the semi-variogram is applied. Therefore, RCMs that have a similar spatial structure will be given large weights even though they may have significantly lower or higher return values when compared with the observed precipitation fields. To compensate for this potential drawback, the model-specific UK-wide weight, V, is combined with a region-specific weight, R, which provides information with respect to the discrepancy in return values between CTRL and UKMO.

As RCMs are not expected to capture fine scale (grid cell scale) patterns in observed return values, a discrepancy measure should be applied on a meaningful spatial scale. Here we consider the UK rainfall regions to be an appropriate scale and use as a discrepancy measure the absolute difference between the regionally averaged Rmed value for each RCM's SM and the observed SM. The region-specific weights are calculated and rescaled to sum to unity in a similar fashion to those of the UK-wide weights, i.e. by calculating inverse distances between each UKMO and CTRL regionally averaged Rmed and scaling to unity.

The final combined weight (RV) for each RCM is then given as the average of the UK-wide and region-specific weights:

$$RV_i = \frac{V_i + R_i}{2} \text{ for } i = 1, \ldots, 13 \text{ models} \qquad (2)$$

The combined weights are then used to determine the proportion of the 10 000 bootstrap samples for each RCM that contribute to the regional multi-model ensemble distribution for the weighted example.

4. Results

4.1. Evaluation of control climate

To evaluate RCM skill in simulating the median and range in magnitude of extreme regional precipitation in comparison with observations, we calculated the average and standard deviation of Rmed by model and region and compared these for CTRL and UKMO. Figure 2 illustrates how the distribution of Rmed, used to rescale from the fitted GEV growth factor to the return value, differs for each region. Colour represents the CTRL and UKMO datasets, while symbols represent regions. An example for 10-day precipitation extremes is shown in Figure 2. The results for 1 day are similar except for a larger difference in the standard deviation of 1-day summer Rmed between CTRL and UKMO and overall larger relative differences for mean 10-day Rmed.

Regional extremes are underestimated by the RCMs in all seasons. In winter and spring, the largest discrepancies in both the regional average and standard deviation of Rmed are found for regions that experience the heaviest precipitation, such as North Scotland (NS) and South Scotland (SS), and Northwest England (NWE) and Southwest England (SWE) (Figure 2(a) and (b)). Other regions show less discrepancy in the average Rmed but the standard deviation is generally underestimated. In summer, the major discrepancy between RCM simulations and observations lies in the underestimation of the regionally averaged Rmed (Figure 2(c)). This discrepancy is particularly large for regions with heavy summer rainfall, such as those mentioned above, but also for those regions where extremes are reasonably well captured by RCMs during the winter and spring seasons, e.g. Northeast England (NEE) and central and eastern England (CEE). This may be due to the predominance of local-scale convective processes which cannot be properly resolved at the scale of the RCM grid cell (~50 km). In autumn, the RCMs reproduce extremes well only in regions with low variability and low magnitude precipitation extremes [e.g. CEE and Southeast England (SEE)], failing to capture the high variability and high magnitude of observed precipitation extremes in northern and western regions. Due to the large scatter, it is hard to make generalizations with respect to individual model performance with the possible exception of CLMH which consistently clusters towards the right-hand side of the RCM markers, indicating a smaller discrepancy in the regionally averaged Rmed compared with other RCMs.

DOI: 10.1002/joc

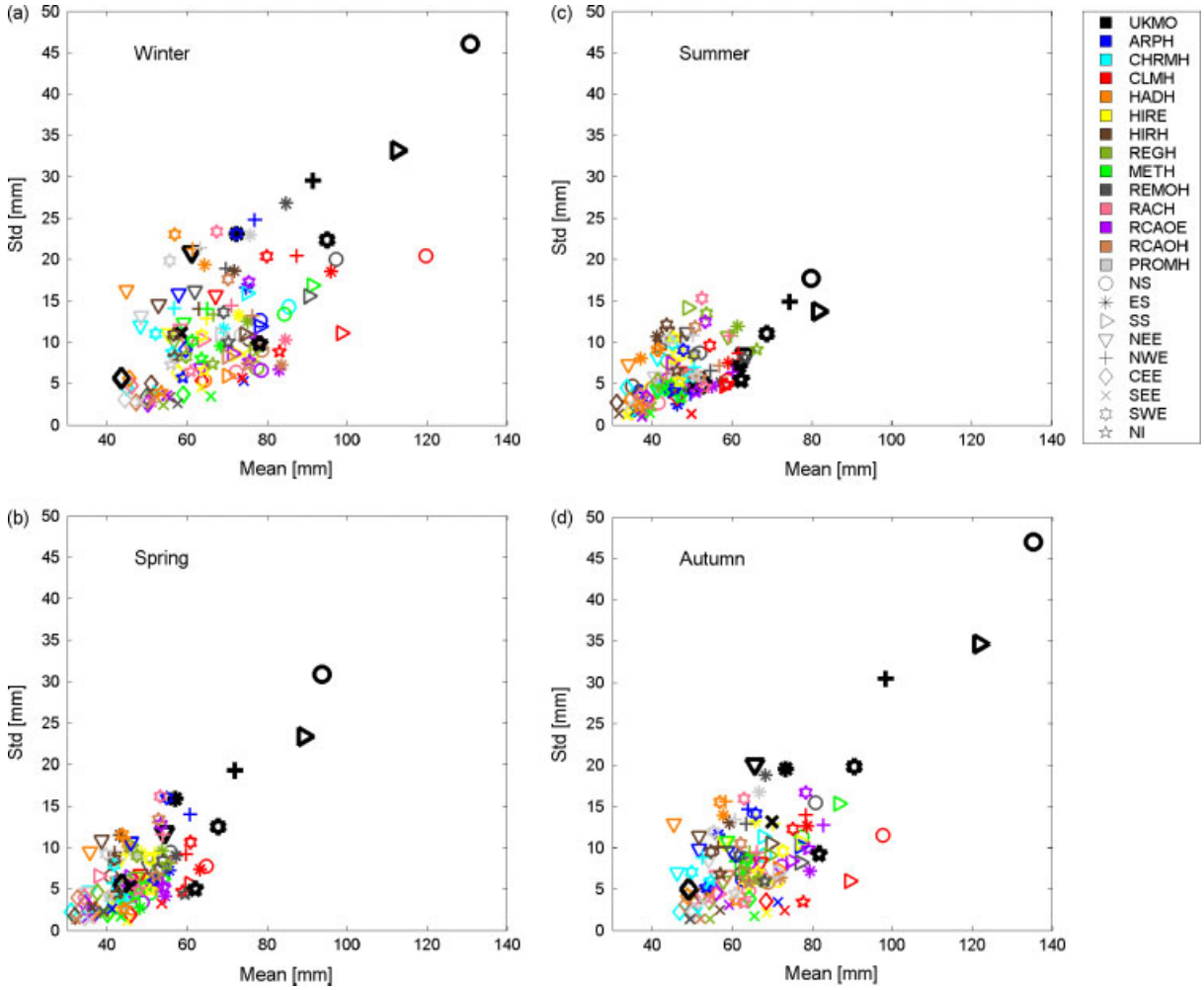

Figure 2. Scatter plots of the mean (regional average) and standard deviation of Rmed values, comparing the UKMO and CTRL integrations for 10-day extreme precipitation for (a) winter, (b) spring, (c) summer and (d) autumn. The CTRL datasets are denoted by different colours, UKMO is shown in bold black, while symbols represent regions. This figure is available in colour online at www.interscience.wiley.com/ijoc

Figure 3 shows the seasonal 1- and 10-day, 5-year return values estimated separately for each individual grid cell for UKMO. The spatial patterns of 1- and 10-day return values are very similar; differing only in magnitude. The highest magnitude extremes are found along the western side of the UK, particularly in the Scottish highlands, NWE, central and western Wales and the southwest moors during winter and autumn. In spring and summer, overall magnitudes are smaller and the highest magnitude extremes are found mainly in the Scottish highlands, NWE and central Wales.

In Figures 4 and 5, the percentage difference between UKMO and each CTRL is given for individual grid cells for the 5-year return value for 1- and 10-day totals, respectively. In winter, most CTRLs show a marked discrepancy from observations in CEE, NS and SS. In CEE, the CTRLs overestimate the observed extremes, particularly CLMH (1 day), HIRE (1 day) and RCAOE (10 days), while in NS the CTRLs underestimate the observed extremes, particularly CHRMH (1 and 10 days), HIRH (1 and 10 days), REGH (10 days) and METH (10 days). Although the spatial patterns of discrepancies are very similar between the 1- and 10-day winter plots, there are marked differences for certain regions and RCMs. For example, CHRMH and CLMH show larger overestimates in CEE for 1-day extremes than 10-day extremes.

In spring, discrepancy patterns are largely similar to those of winter, albeit with lower positive discrepancies in CEE and SS (Figures 4(b) and 5(b)). The overall features of the 1- and 10-day patterns are similar, with some regional differences for individual RCMs, e.g. CHRMH, HIRH and HADH show larger underestimates in NS, SS and NWE for the 10-day extremes. In summer, the discrepancy pattern is predominantly negative, meaning that RCMs underestimate summer precipitation extremes in comparison with observations; in particular for HADH and HIRH, where the negative discrepancy increases (becomes more negative) towards the south and west of the UK. However, there are exceptions; REGH shows a marked positive discrepancy in CEE. In autumn, the largest discrepancies between CTRLs and UKMO are found in NS and CEE. RCMs, particularly CHRMH and HIRH, tend to underestimate autumn precipitation extremes in NS and these negative discrepancies become

DOI: 10.1002/joc

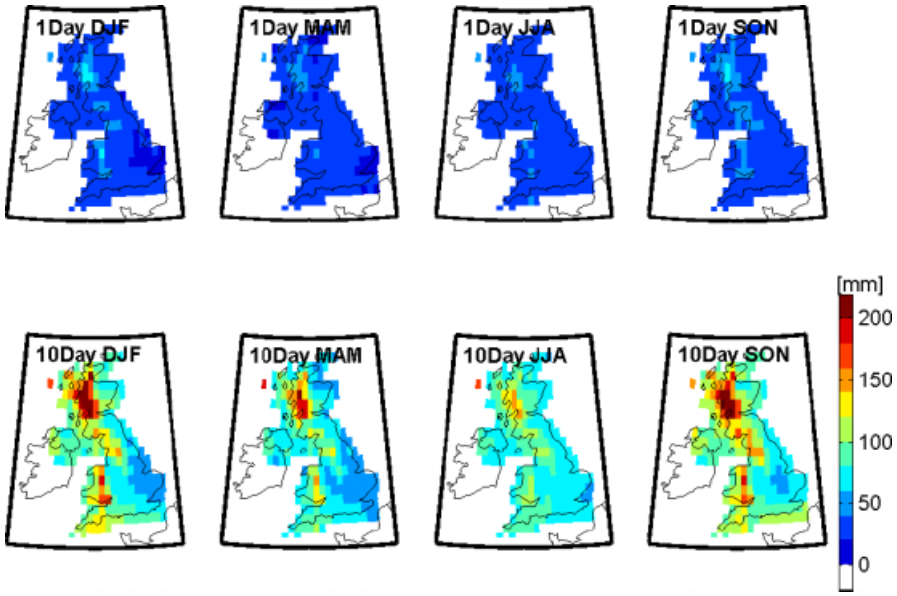

Figure 3. Return value estimates for the UKMO 1- and 10-day totals with 5-year return period for each meteorological season. Season and accumulation period is stated in each separate figure, where winter is denoted by DJF (December–February), spring by MAM (March–May), summer by JJA (June–August) and autumn by SON (September–November). This figure is available in colour online at www.interscience.wiley.com/ijoc

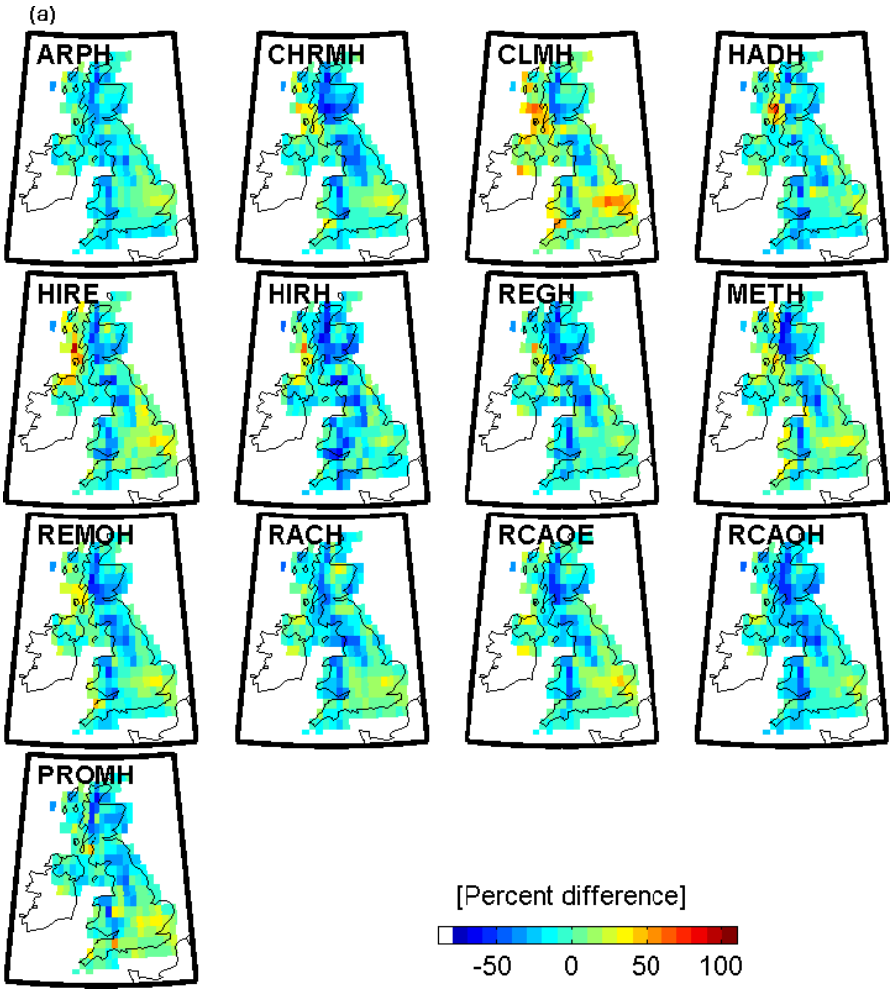

Figure 4. (a) Percent difference (discrepancy) between grid cells for UKMO and CTRLs for the 1-day 5-year return value during winter. (b) Percent difference (discrepancy) between grid cells for UKMO and CTRLs for the 1-day 5-year return value during spring. (c) Percent difference (discrepancy) between grid cells for UKMO and CTRLs for the 1-day 5-year return value during summer. (d) Percent difference (discrepancy) between grid cells for UKMO and CTRLs for the 1-day 5-year return value during autumn. This figure is available in colour online at www.interscience.wiley.com/ijoc

DOI: 10.1002/joc

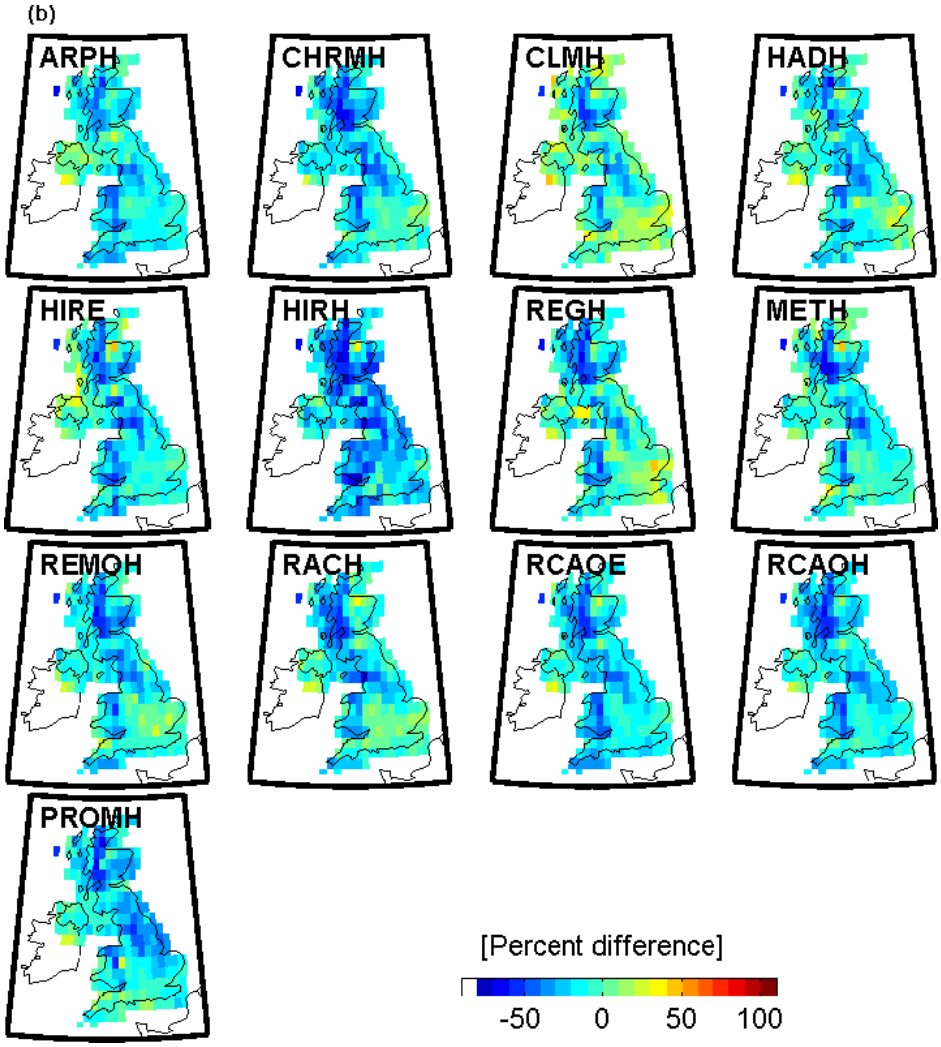

Figure 4. (*Continued*).

larger for most RCMs for 10-day extremes compared with 1-day extremes. In CEE and NEE, most RCMs show weak overestimation (somewhat larger in CLMH) for 1-day extremes but this discrepancy reduces for 10-day extremes, with the exception of RCAOE and to a lesser extent RCAOH.

4.2. Future projections for individual models

Probability distributions of change in extreme precipitation were created using the 10 000 return value estimates generated for each RCM, event duration (1, 2, 5 or 10 days) and region by the non-parametric bootstrapping exercise detailed in Section 3.1 and further elaborated in Fowler *et al.* (2007b). We show only results for 1- and 10-day totals for the 5-year return value due to space constraints.

In general, RCM projections of change vary considerably and uncertainty ranges are larger for 25-year than 5-year return values, as would perhaps be expected. Figure 6 shows estimates of the percent change in the 1-day 5-year and 10-day 5-year return values, respectively, for each RCM and each season under the SRES A2 2071–2100 emissions scenario for the NS region. Figures 7 and 8 show the same plots for the NWE and SEE regions, respectively. These illustrate the model uncertainty in predictions for contrasting regions of the UK, although discussion will be made of results for other regions.

RCM projections are most consistent for winter extremes. All RCMs project increases in the 1-day 5-year and 10-day 5-year return values; although there is an overlap of the probability distribution with the zero change line for some RCMs, mainly HADH, in southern regions, these are in the minority. ECHAM-driven RCMs, in particular RCAOE, project larger increases than Hadley-driven RCMs in winter. There is a greater uncertainty for 25-year return values; more RCMs project zero change or reductions. The results are similar for spring, with most RCMs projecting increases in extreme precipitation for both 1- and 10-day events. However, there is no distinction between projections from ECHAM- and Hadley-driven RCMs.

In summer, the projections are less consistent in terms of sign and, in general, span the zero change line for the 1-day 5-year event. At the 25-year return value, RCM predictions span a wide range but are skewed positive.

DOI: 10.1002/joc

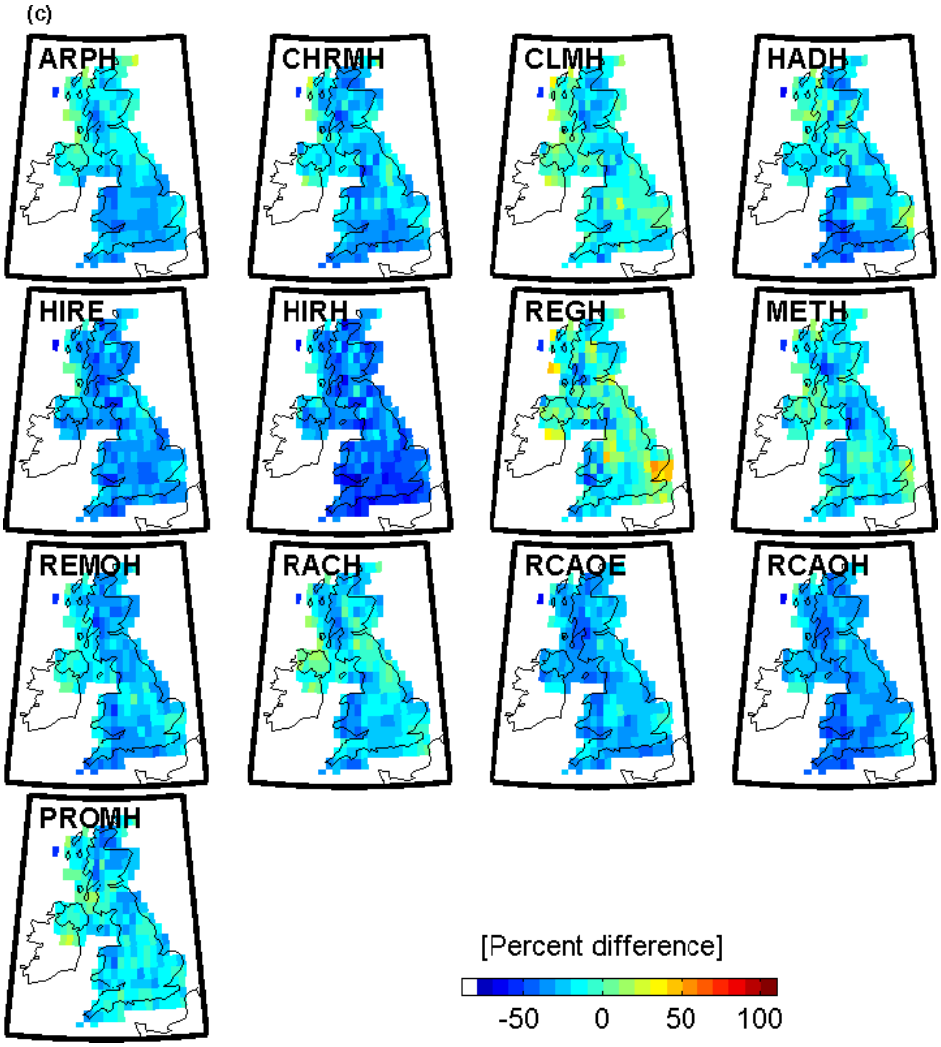

Figure 4. (*Continued*).

For the 10-day 5-year event, there is a split between northern regions (NS, SS, ES, NWE and NEE), where RCMs consistently predict a decrease in return value, and southern regions, where model predictions span the zero change line. In autumn, most RCMs predict increases in the 1-day and 10-day 5-year return values; at 1 day some RCMs predict very large increases, particularly the ECHAM-driven RCMs and particularly in England and Wales. For the 25-year return period, the predictions are more uncertain but most RCMs still predict increases (not shown).

5. Sensitivity of results to model weighting

There are obviously different ways by which we may combine RCMs. In many cases, simple weighted averages, using the historical relationship between forecasts and observations (e.g. Krishnamurti *et al*., 2000), have been found to perform better than simple averages where each model is weighted equally. Although it intuitively makes sense to trust, and thus weight, the better performing models more, it is difficult to objectively quantify model skill and therefore derive model weights (Tebaldi and Knutti, 2007). Here, we show results from pooling RCM estimates of change in extreme precipitation using equal weighting and for a simple weighting technique based on the skill of RCMs in simulating the magnitude and spatial distribution of observed extremes. Proportional weighting of large multi-model ensembles has previously been used to derive probabilistic estimates for future flows in the Eden catchment, UK (Fowler *et al*., 2007a; 2008) and for changes in temperature, rainfall and summer flow abstractions for the Thames catchment, UK (Manning *et al*., in press).

5.1. Future projections using simple weighting schemes

For the weighted estimates, we use the semi-variogram parameters of sill and range in combination with the regionally averaged Rmed values, following the methodology outlined in Section 3.2. Unlike the Rmed weights, the semi-variogram weights are the same for each region as the experimental semi-variograms are calculated using all UK grid cells to ensure a robust estimate of the semi-variogram. Thus, while the regionally averaged Rmed weights vary with region and season (as shown

DOI: 10.1002/joc

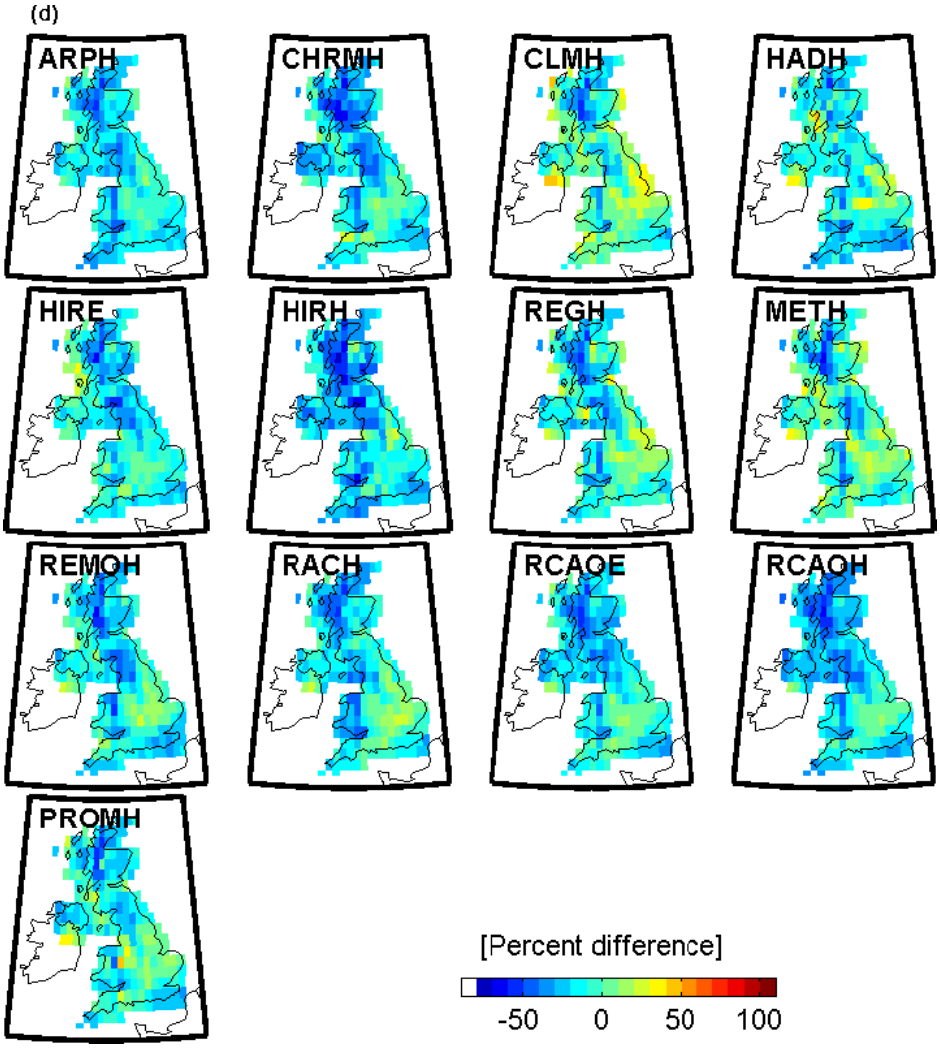

Figure 4. (*Continued*).

in Figure 2) the semi-variogram weights vary only with season. In Figure 9, the experimental semi-variograms and the respective semi-variogram models are shown for each season for the 1-day 5-year return values. The semi-variograms reflect the spatial characteristics as seen in Figure 3 and display seasonal characteristics that persist through event duration and return period. The largest spatial variability (as indicated by the sill value) is found during winter, followed by autumn, spring and summer; a similar order of intra-seasonal variability to that shown by the regional Rmed plots in Figure 2. Less variability is found in range, which tends to vary around ∼130 km. The CTRL semi-variogram models also show seasonal variability, albeit less so compared with the UKMO data, particularly for autumn and spring. Furthermore, the CTRL models appear to have a longer range compared with the UKMO models, indicating that the spatial dependence is greater in the CTRL extremes compared with the UKMO extremes; a result in agreement with the greater standard deviation shown for the UKMO regionally averaged Rmed values in Figure 2.

To visualize the impact of each of the two measures separately and in combination, weights for all three schemes are first derived for two regions with clear differences in precipitation characteristics, NS and SEE. Figure 10 shows the seasonal weights for each RCM for the 1- and 10-day, 5-year return values. The semi-variogram weights are not resolved on a regional basis but represent the overall performance of the RCM in capturing the spatial characteristics of the extreme precipitation field; thus weights vary only by season and event duration. The homogeneity of the semi-variogram weights shows that there are relatively small differences in the RCMs' abilities to replicate spatial structures in the extreme precipitation fields. Only on a few occasions is a RCM given a larger weight compared with the overall RCM ensemble; e.g. in winter and summer for the 1-day totals and only in winter for the 10-day totals. However, the model weights based on the regionally averaged Rmed values show large differences, not only between seasons but also between regions. In NS model weights are very similar, indicating that the RCMs perform equally well (or poorly in the case of NS). In SEE, on the other hand, some RCMs are given very large weights, particularly in winter and autumn for 1-day return values and in summer and spring for 10-day return values.

DOI: 10.1002/joc

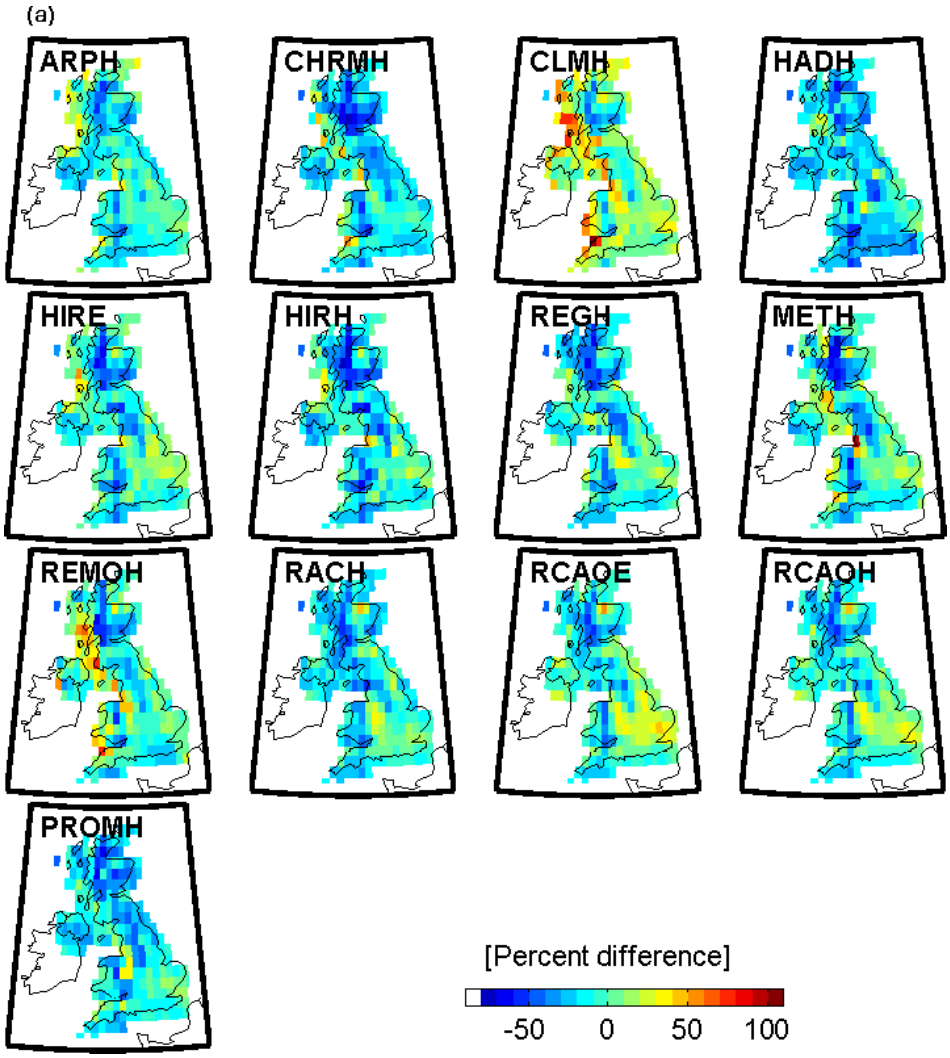

Figure 5. (a) Percent difference (discrepancy) between grid cells for UKMO and CTRLs for the 10-day 5-year return value during winter. (b) Percent difference (discrepancy) between grid cells for UKMO and CTRLs for the 10-day 5-year return value during spring. (c) Percent difference (discrepancy) between grid cells for UKMO and CTRLs for the 10-day 5-year return value during summer. (d) Percent difference (discrepancy) between grid cells for UKMO and CTRLs for the 10-day 5-year return value during autumn. This figure is available in colour online at www.interscience.wiley.com/ijoc

It is clear that some RCMs are given very large weights due to a high similarity to observations in their regionally averaged Rmed. However, given that the semi-variogram weights show clear differences between UKMO and each CTRL in terms of spatial characteristics, we suggest that using the model weights based on regionally averaged Rmed alone would bias the distribution towards RCMs that perform well in capturing the magnitude of extremes but may show deficits when considering spatial characteristics. Thus, here we use a combination of the two weights, also displayed in Figure 10.

Tables II–V give the individual RCM weights for each region and season for the 1-day, 5- and 25-year return values and the 10-day, 5- and 25-year return values. Some general conclusions can be drawn from Tables II–V; e.g. per season and region, weights are generally larger in autumn for the 1-day totals and during winter for the 10-day totals. When considering all regions together, the seasons with largest variability in weights are autumn (1 day) and summer (10 days). The regions that have most variable model weights are CEE, SEE and NEE for the 1-day totals and ES and SEE for the 10-day totals. Regions with least variability in model weights are SS and SWE for all event durations and return values. The RCM that is repeatedly given a high weight within different regions is CLMH. However, REMOH is given the highest individual weights, with 44.4% in CEE in autumn for the 1-day 5-year return value, 38.4% in CEE in autumn for the 1-day 25-year return value, 45.4% in ES in spring for the 10-day 5-year return value and, finally, 45.4% in ES in winter for the 10-day 25-year return value.

5.2. Future projections using a simple weighting scheme – comparison with equally weighted pooling

In Figures 11–14, results of weighted (column 1) *versus* unweighted (column 2) multi-model distributions are compared with pooled results from a 4×4 ensemble of RCM integrations: HIRHAM and RCAO driven by

DOI: 10.1002/joc

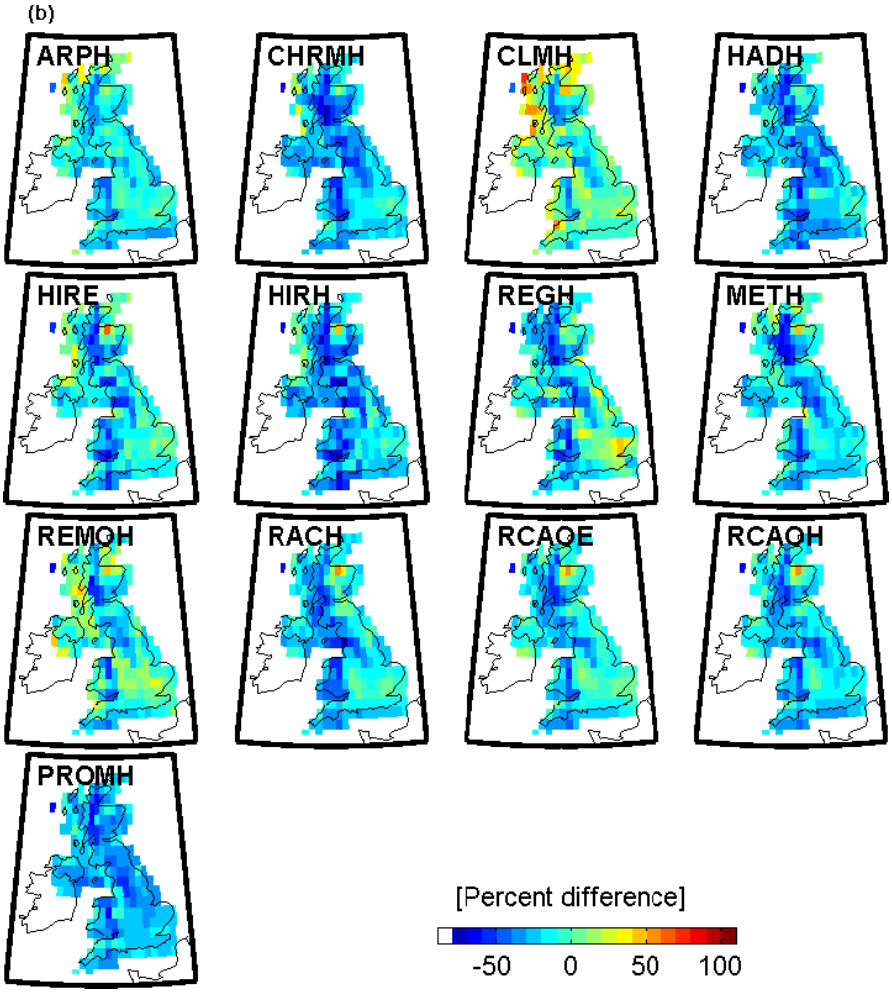

Figure 5. (*Continued*).

HadAM3H and ECHAM4/OPYC3 (columns 3 and 4, respectively, unweighted). Figures 11 and 12 show box plots of the estimated percent change in 1-day 5-year and 1-day 25-year return values, respectively, for each season, and Figures 13 and 14 show the same plots for the 10-day 5-year and 10-day 25-year return values.

In winter, the weighted multi-model ensemble distribution for 1-day 5-year precipitation extremes is wide but projects positive change in all regions (Figure 11, column 1). The most likely range of change (the 25th to 75th box plot quartiles) is from 15 to 30%, with slightly smaller increases projected for SWE and NI. The weighted and unweighted distributions are similar (Figure 11, columns 1 and 2), although there is a narrowing of the inter-quartile range (reduced uncertainty) for the weighted distribution. Differences between projections from Hadley-driven RCMs (Figure 11, column 3) and ECHAM-driven RCMs (Figure 11, column 4) are greater; respectively, projecting increases of ∼10–30% and ∼30–50% depending on region. For the 25-year event (Figure 12), the distribution indicates greater uncertainty and the contrast between RCM projections is larger; Hadley-driven RCMs project no change or reductions in the 1-day 25-year return value in NS and NI (Figure 12, column 3) but ECHAM-driven RCMs project larger increases in the 25-year than the 5-year return value, particularly in northern and western regions (Figure 12, column 4). Pooling the RCMs therefore produces wider multi-model ensemble distributions for the 25-year return period, although the weighted distributions (Figure 12, column 1) provide tighter constraints on the inter-quartile range of projected changes than the unweighted distributions (Figure 12, column 2) for all regions except perhaps NWE. The projected ranges for regional increases are 10–30% in Scotland, northern and central England, and 5–20% in southern England and Northern Ireland (NI).

The distributions of the multi-model ensemble are narrower for projected change in the 10-day 5-year return value in winter, with an estimated increase of 5–20% (lower in SWE) from the multi-model weighted ensemble (Figure 13, column 1 compared with Figure 11, column 1). Weighted and unweighted multi-model ensembles are very similar, suggesting perhaps that all RCMs reasonably simulate 10-day precipitation extremes. Projected increases are larger for ECHAM- than Hadley-driven RCMs, but the differences are not as great as for 1-day extremes (Figure 13, columns 3 and 4 compared with Figure 11, columns 3 and 4). At the 25-year return period, Hadley-driven RCMs project a gradient of change

DOI: 10.1002/joc

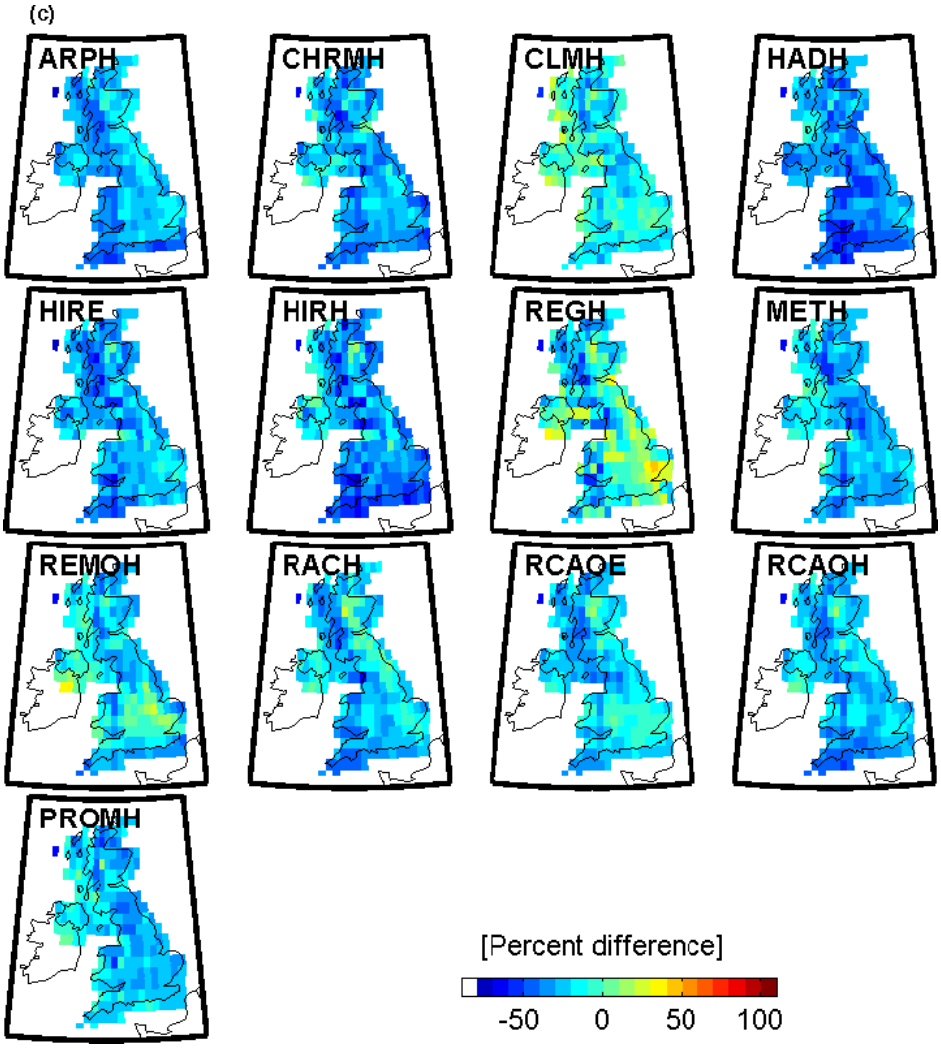

Figure 5. (*Continued*).

from a median increase of ~20% in northern Scotland to no change in southern England (Figure 14, column 3), whereas ECHAM-driven RCMs give a median increase of 20–30% across the UK (Figure 14, column 4). The multi-model ensemble (weighted or unweighted; columns 1 and 2) shows a gradient of change with higher increases in the north (~20%) than south (~0%) of the UK.

In spring, the multi-model ensemble (weighted and unweighted) suggests increases in the return values of 1- and 10-day 5-year events in all regions (Figures 11 and 13, columns 1 and 2). The RCMs project the most likely range of change to be 10–20% for the 1-day 5-year return value and 10–25% for the 10-day 5-year return value (except for SWE which is lower), with only small differences between the weighted and unweighted multi-model ensembles. For most UK regions the inter-quartile range is above the zero change line, except in SWE and NI. There is no clear distinction between projections from Hadley- and ECHAM-driven models (Figures 11 and 13, columns 3 and 4), although ECHAM-driven models project smaller increases than Hadley-driven models for 10-day events. Projected changes in 25-year return values are similar at 1 day for Hadley-driven models but uncertainty ranges are larger (Figure 12, column 3 compared with Figure 11, column 3). For ECHAM-driven RCMs (Figure 12, column 4), little change in the 1-day 25-year return value is projected for NEE, CEE and SEE; the largest increases (~20%) are projected for ES and NI. At both 1 and 10 days, the weighted and unweighted multi-model ensembles (Figure 12, columns 1 and 2) are very similar, although, for the 1-day 25-year return value the inter-quartile range is perhaps widened by weighting (Figure 12, columns 1 and 2). The ECHAM-driven RCMs project large uncertainty in changes to the 10-day 25-year return value; particularly in northern and western regions, due to large differences in changes projected from the individual RCMs for this season.

In summer, the multi-model ensembles are wide and weighting RCMs' results make little difference to estimated changes; perhaps all RCMs are equally bad at simulating summer precipitation extremes. Summer is the only season where the RCMs project a decrease in precipitation extremes. For the 1-day 5-year event, the most likely range of change is from −10 to 0% in North and East Scotland but from −5 to +15% changes in South

DOI: 10.1002/joc

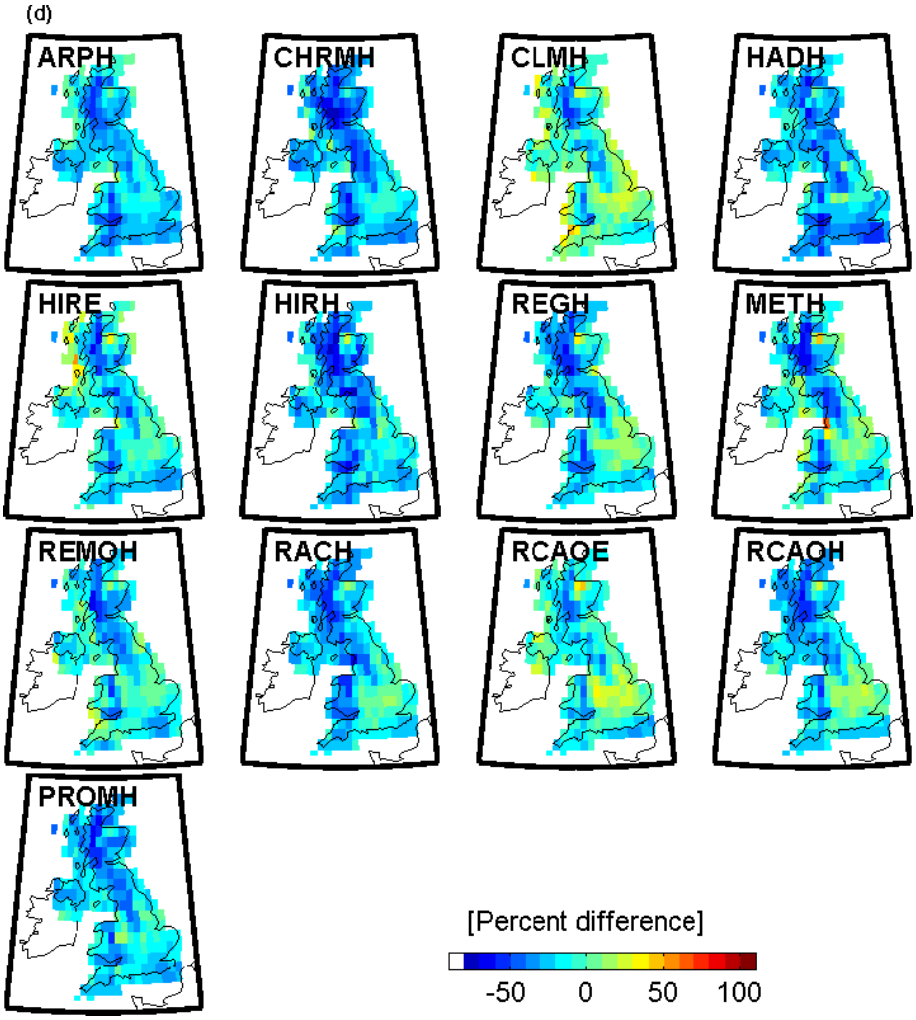

Figure 5. (*Continued*).

Scotland, England and Wales and NI (Figure 11, columns 1 and 2). ECHAM-driven RCMs (Figure 11, column 4) show a much larger uncertainty range in predictions than Hadley-driven RCMs (Figure 11, column 3), although uncertainties vary considerably from region to region. For the 1-day 25-year event, Hadley-driven RCMs predict a gradient of change from ∼20% increase in North Scotland to zero change in SEE (Figure 12, column 3). ECHAM-driven RCMs predict much larger changes in western regions: ∼20% increase in NW and SW England and ∼30% increase in NI. The distributions of the multi-model weighted and unweighted ensembles project changes for the 1-day 25-year return value from −10 to +20% (Figure 12, columns 1 and 2). For the 10-day 5-year return value (Figure 13, columns 1 and 2), the multi-model ensemble projects decreases in extremes in northern regions but increases in southern regions. The uncertainty range is again much greater for ECHAM-driven RCM projections (Figure 13, columns 3 and 4). The spatial pattern of projected changes is similar for the 25-year event (Figures 13 and 14), although with smaller increases.

In autumn, the distributions of the weighted and unweighted multi-model ensembles consistently project increases for both the 1-day 5-year and 10-day 5-year return values, with the most likely range of change being 10–25% (except for NI which is lower) and 5–20%, respectively (Figures 11 and 13, columns 1 and 2), with narrow uncertainty ranges and weighting seemingly making little difference. However, predictions from Hadley- and ECHAM-driven RCMs for the 1-day 5-year return value are markedly different, with predicted increases of 0–20% and 20–60%, respectively, with the larger increases predicted for southern parts of the UK (Figure 11, columns 3 and 4). For the 10-day 5-year return value, this difference is not so marked but predicted increases for ECHAM-driven RCMs are still larger than those for Hadley-driven RCMs (Figures 11 and 13, columns 3 and 4). The distribution of the multi-model ensemble of changes in the 25-year return value is similar but wider (Figures 12 and 14, columns 1 and 2). For the 1-day 25-year event, ECHAM-driven RCMs predict very large changes; median estimates range from ∼20 to 60% with the largest increases predicted for NWE (Figure 12, column 4); these large increases are not similarly predicted at 10 days, where projected changes are ∼20–30% (Figure 14, column 4).

DOI: 10.1002/joc

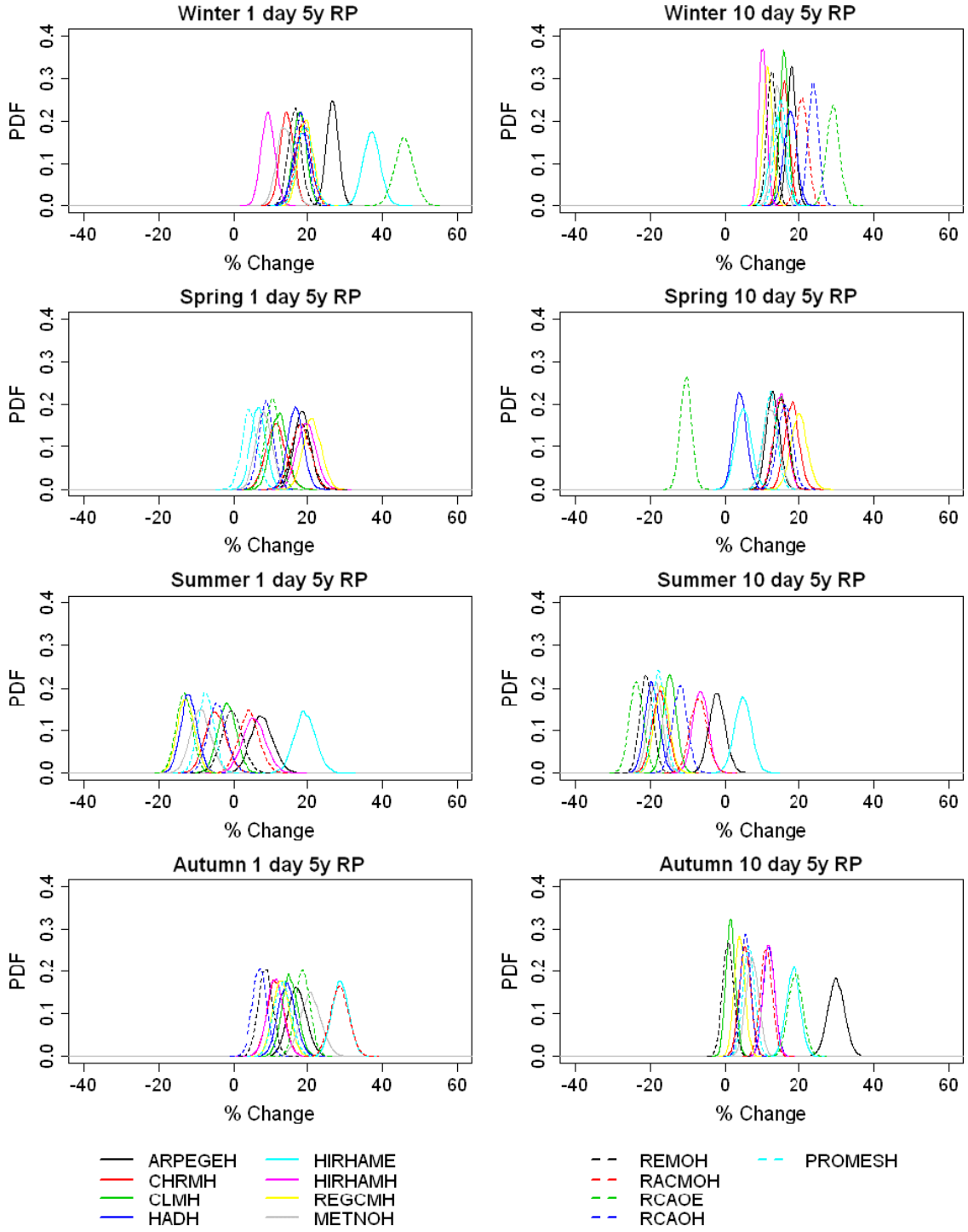

Figure 6. Estimates of percent change in the 1-day 5-year and 10-day 5-year return values, respectively, for each RCM and each season under the SRES A2 2071–2100 emissions scenario for the North Scotland region (NS). This figure is available in colour online at www.interscience.wiley.com/ijoc

6. Discussion and conclusions

Inter-model comparison projects, such as the EU-funded PRUDENCE and ENSEMBLES and their North American counterpart NARCCAP, have made available high-resolution climate model data from several RCMs to the climate impacts community. However, when faced with results from a range of models, it is difficult to interpret and communicate results, particularly when there is low agreement between individual models. In Fowler *et al.* (2007b) six RCMs from the PRUDENCE ensemble were used to investigate projections of regional change in UK annual precipitation extremes. To represent inter-model variability, bootstrapped samples of change in regional extremes for each model were pooled to provide an unweighted multi-model ensemble. However, it was clear that some RCMs failed to capture not only the expected geographical pattern of extremes but also their

DOI: 10.1002/joc

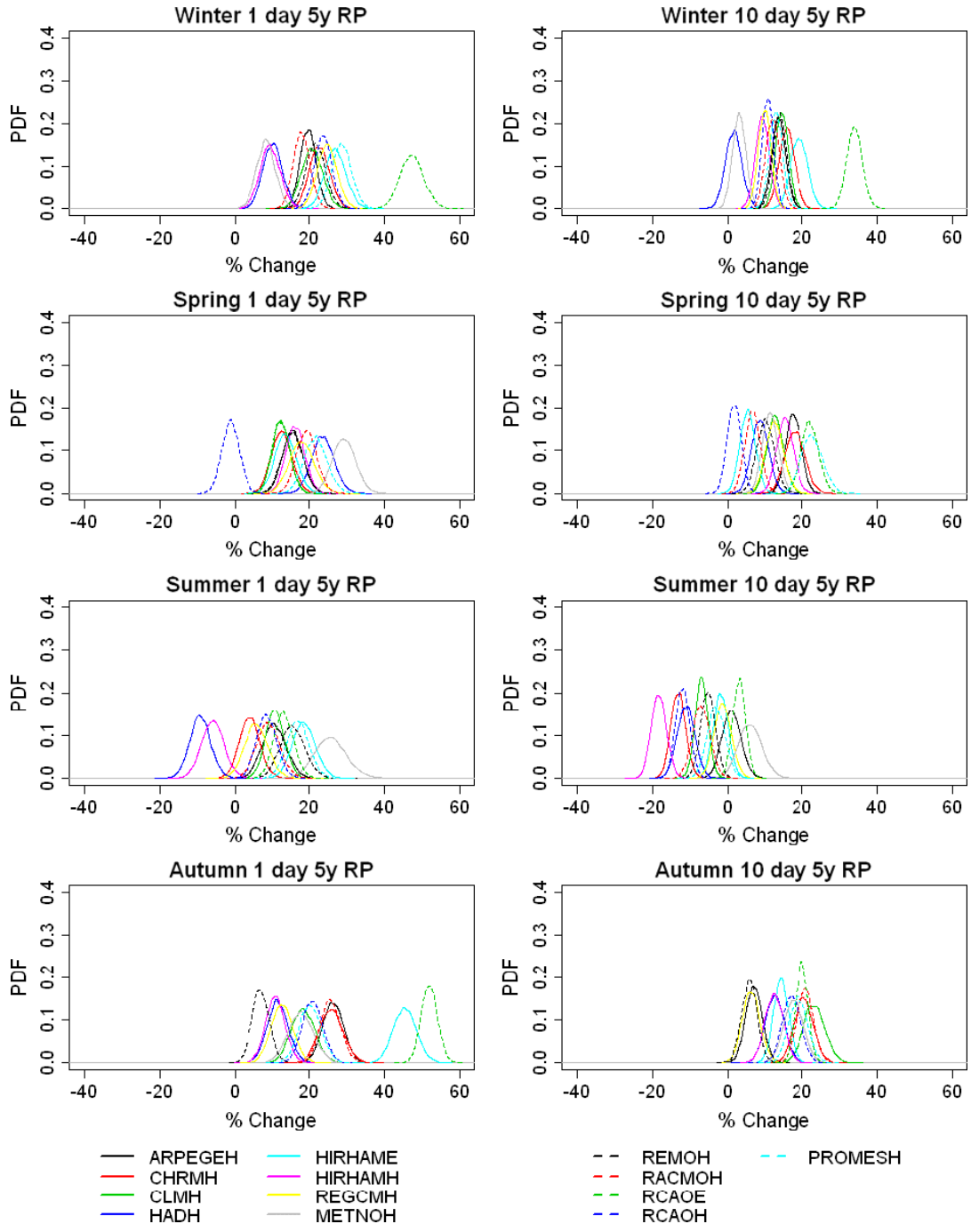

Figure 7. Estimates of percent change in the 1-day 5-year and 10-day 5-year return values, respectively, for each RCM and each season under the SRES A2 2071–2100 emissions scenario for the Northwest England region (NWE). This figure is available in colour online at www.interscience.wiley.com/ijoc

median magnitude, as seen in observations, suggesting that those RCMs were perhaps less suitable for examining regional change. Here, we addressed this issue by using a combination of the spatial structure and regionally averaged Rmed discrepancy of each RCM to weight its contribution to the multi-model ensemble. We assumed that RCMs that simulate extremes with dissimilar characteristics to observations are poorer at representing precipitation processes at the scales required for regional studies and hence should have less weight in the multi-model ensemble. Furthermore, we used a larger set of RCMs, 13 models instead of 6, and the analysis was performed separately for each season. Here precipitation was aggregated to the regional level, as estimates of return value magnitudes are improved due to the effect of data pooling and use of the regionally averaged Rmed (Fowler *et al.*, 2007b). Kendon *et al.* (2008) found that 3×3 grid cell pooling conveyed the greatest benefits to estimating return values, which is the approximate size of most UK rainfall regions.

We derived RCM weights by combining two simple weighting schemes. The first weight (V) was derived

DOI: 10.1002/joc

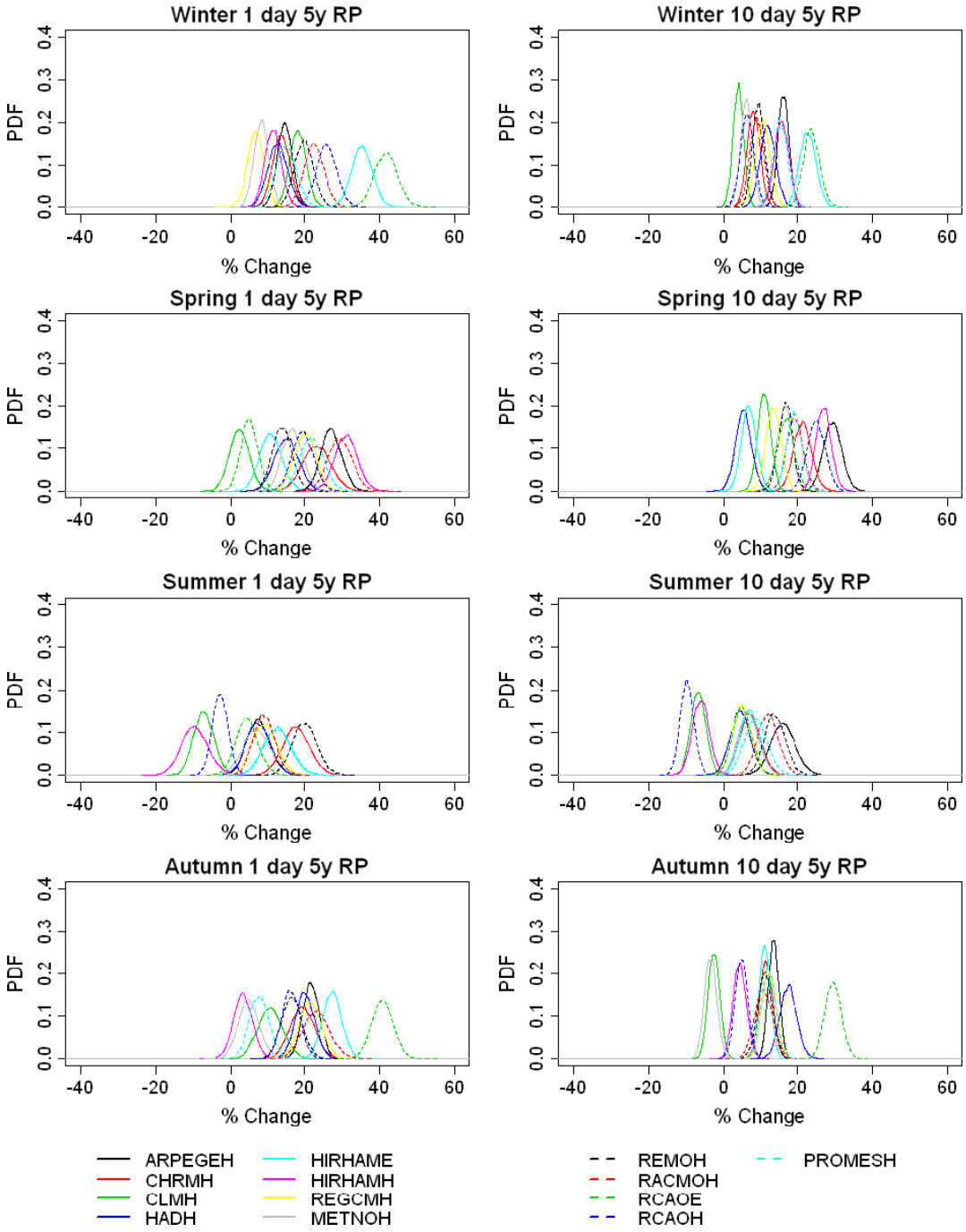

Figure 8. Estimates of percent change in the 1-day 5-year and 10-day 5-year return values, respectively, for each RCM and each season under the SRES A2 2071–2100 emissions scenario for the Southeast England region (SEE). This figure is available in colour online at www.interscience.wiley.com/ijoc

from experimental semi-variograms calculated for each RCM, accumulation, season and return period. Four different semi-variogram models were fitted to the experimental semi-variogram and the model with best fit was retained for further analysis. We then normalized the semi-variogram model parameters of sill and range and calculated the inverse distance between CTRL and UKMO data in sill-range parameter space to give weights. The second weighting scheme (R) was based on the discrepancy between the CTRL and UKMO regionally averaged Rmed. When combined, the weights scaled to unity were used to indicate the relative proportion of the regional RCM bootstrap sample to include in the multi-model ensemble distribution (shown in Tables II–V).

The weighting scheme identified regions and seasons where there were large discrepancies in model performance. For example, weights were more unevenly distributed in the regions CEE, SEE and NEE for the 1-day totals and ES and SEE for the 10-day totals. Similarly for seasons, the largest variability in weights was found for autumn (1 day) and summer (10 days). Due

DOI: 10.1002/joc

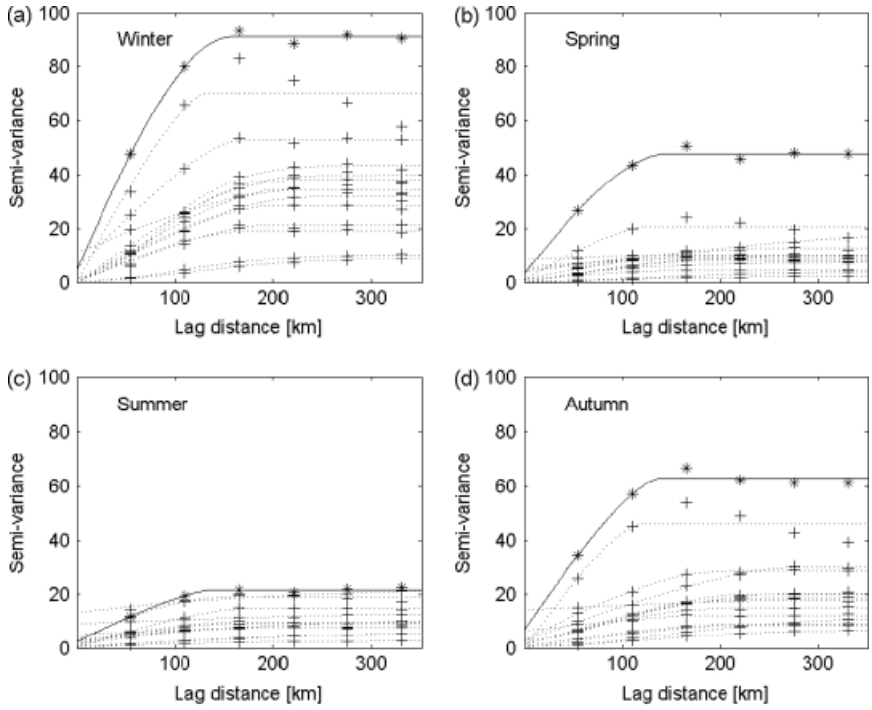

Figure 9. Experimental semi-variograms and fitted semi-variogram models for UKMO (asterisk and full line) and CTRLs (plus sign and dotted lines) for the 1-day 5-year return value during (a) winter, (b) spring, (c) summer and (d) autumn.

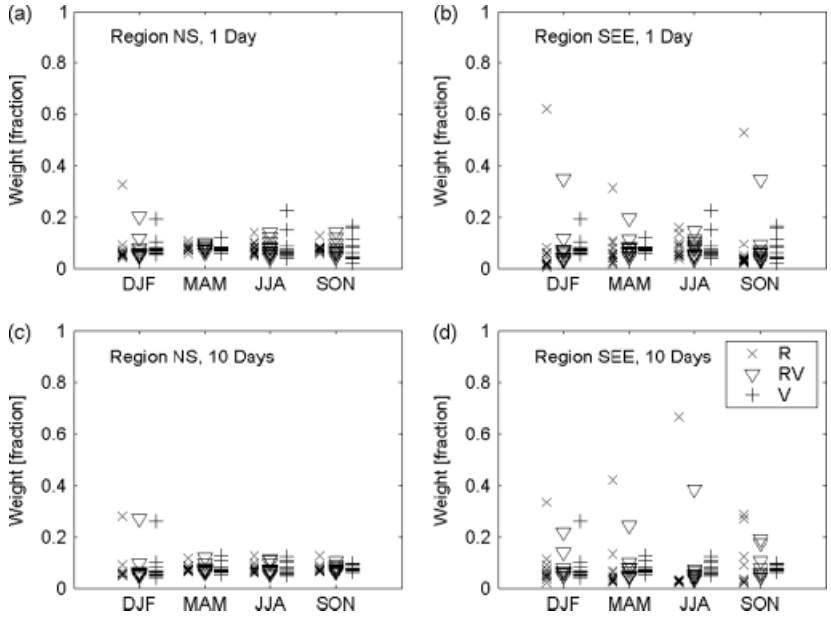

Figure 10. The distribution of weights for the 5-year return value for the 1- and 10-day totals for two rainfall regions: northern Scotland (NS) and Southeast England (SEE). For each season (x-axis), the magnitudes (y-axis) for three weighting schemes are visualized. The '×' markers denote the regionally averaged Rmed weights (R), '▽' markers denote the combined Rmed and semi-variogram parameter weights (RV) and the '+' markers denote the semi-variogram parameter weights (V).

to considerable variability across different regions and seasons, it is difficult to make an assessment of individual model performance. However, CLMH was most frequently assigned large weights in different regions or seasons and the largest individual weight (45.4%) was assigned to REMOH for the 10-day 5-year return value in ES in spring and for the 10-day 25-year return value in ES in winter.

Using the combined weights, multi-model ensembles of percent change were created for each region, season, accumulation period and return period (Figures 11–14). Identifying the effects of using the weighting scheme is difficult due to the very large inter-model differences among seasons and regions. However, the overall effect of using the weighting scheme seems to be a tightening of the regional distributions compared with the unweighted

DOI: 10.1002/joc

Table II. Model weights (%) for the 5-year return value.

Region	Season	ARPH	CHRMH	CLMH	HADH	HIRE	HIRH	REGH	METH	REMOH	RACH	RCAOE	RCAOH	PROMH
NS	Winter	5.5	6.9	20.2	11.9	5.8	4.8	5.9	7.1	8.2	5.5	5.3	5.1	8.1
	Spring	7.6	6.9	9.5	10.1	7.2	6.1	7.3	8.7	7.8	7.5	6.8	7	7.3
	Summer	4.9	6	10.5	14.2	5.3	4.2	7.8	8.2	6.8	8.5	6.5	5	12.3
	Autumn	7.3	5.8	14.3	4.4	6.3	4.7	6.6	6.5	12.2	7.4	9.9	6.6	8.1
ES	Winter	5.9	10.7	4.7	11.7	4.5	5	12.5	7.9	13.2	4.3	4.5	9.2	6.1
	Spring	6	4.8	13.1	8	5	4.1	6	7.2	5.5	19.5	5.3	5.3	10.1
	Summer	3.3	4.1	21	12.6	3.5	2.7	11	5.6	4.3	7	4.9	3.4	16.5
	Autumn	5.2	10.1	8.5	3.7	6.6	3.5	10.8	8	9.6	12.8	7.6	8	5.5
SS	Winter	6	7	11.7	13.1	6.3	5.7	6.7	9.3	8.3	5.8	6	5.6	8.5
	Spring	7.4	7.1	9.4	9.8	7.7	6.1	7.5	9.6	8	7.2	6.7	6.9	6.5
	Summer	5.2	6.7	9.4	15.3	5.3	4.4	8.4	7.6	7.2	7.9	6	5.3	11.2
	Autumn	7.4	5.5	14.8	5	5.9	4.7	7.1	7.3	11.6	7	9.3	5.9	8.4
NEE	Winter	6	4.2	5.2	10.3	14	3.7	4.6	26.8	5	4.4	5.8	3.8	6.1
	Spring	7.3	6.5	11.2	8.6	6.5	5.3	9.6	8.3	7.8	10.8	6.9	6.3	4.9
	Summer	3.9	4.5	6.8	13	4.2	3.1	27.4	5.4	4.8	7.9	5.4	4.1	9.6
	Autumn	4.8	2.4	13.1	1.5	2.6	2.5	8.7	34.1	9	5.5	6.7	3.4	5.5
NWE	Winter	8.2	6.3	9.7	13.7	7.2	5.6	6.6	7.1	6.5	7.3	6.9	5.7	9.2
	Spring	10.2	6.4	9.8	9.9	6.9	5.7	7.9	7.3	6.9	8.1	7.3	7.5	6.3
	Summer	6	5.5	7.9	15.1	5.2	4.1	11.3	6.3	5.9	9.3	6.3	5.6	11.4
	Autumn	7.4	4.5	15.6	5.1	5.5	4.3	6.9	6.7	11.4	9	9.4	5.9	8.4
CEE	Winter	3.6	7.9	4.4	11.1	3.4	3.6	5.1	3.9	4.9	34.4	4	4.9	8.8
	Spring	4.8	4.3	5.6	10.7	4.4	3.6	5.9	36.5	5.3	7.2	4.2	4	3.5
	Summer	9.9	5.6	10.5	15	5	3.5	8.8	6.9	5.2	8.5	6.4	4.6	10.1
	Autumn	4.6	2.3	8	4.1	2.8	2.1	3.3	2.2	44.4	8.8	8.7	3.1	5.6
SEE	Winter	3.6	34.8	4.3	11.7	3.7	6.8	4.1	4	6.5	4.1	5.6	3.2	7.5
	Spring	6	5.6	7.9	11.4	6.5	4.3	6.2	19.6	8.2	6.7	8.5	4.7	4.4
	Summer	8.7	5.3	11.5	14.7	4.7	3.7	8	7.7	5.3	9.4	5.5	5	10.5
	Autumn	6.3	3.5	34.3	3	4.6	3.1	5.6	6.8	9.5	5.9	7.3	4	6
SWE	Winter	7.9	6.5	9.3	13.7	6.9	5.4	6.6	7	7.1	8.4	6.8	5.9	8.6
	Spring	8.4	6.4	9.4	10.3	7.6	5.8	6.6	8.8	7.3	9.5	7.2	7	5.6
	Summer	6.7	5.8	8.4	14.9	5.3	4.5	9.1	7.3	6.1	9	6.4	5.7	10.8
	Autumn	8	4.7	13	5.4	5.8	4.6	7.4	7.4	11.9	8.1	9.9	6	7.8
NI	Winter	6.1	5	4.4	12	9.1	3.2	18.8	5	7.9	5.4	3.6	13.1	6.5
	Spring	6.3	5	5.9	9.8	6.7	4.3	12.3	6.1	7.9	9.5	10.6	10	5.6
	Summer	4.4	5	12.2	13.2	3.5	2.8	18.3	5.2	5.2	8.7	4.3	4.3	12.8
	Autumn	5.8	3.3	25.5	6.3	5.2	3	6.1	6	10.9	6.6	8.2	4.5	8.6

Column *Region* gives rainfall region and column *Season* gives meteorological season. Values are given for the 1-day event duration.

DOI: 10.1002/joc

Table III. Model weights (%) for the 25-year return value.

Region	Season	ARPH	CHRMH	CLMH	HADH	HIRE	HIRH	REGH	METH	REMOH	RACH	RCAOE	RCAOH	PROMH
NS	Winter	5.8	7.3	19.8	9.7	6.3	5.1	6.5	6.6	8.1	5.9	5.4	5.8	7.8
	Spring	7.4	7.1	9.2	9.2	7.9	6.2	7	8.5	7.7	7.4	6.2	7.4	8.8
	Summer	6.5	7.2	11.3	11.4	5.6	4.3	9.6	8.9	7.2	5.1	6.2	4.8	12
	Autumn	7.3	6.7	8.9	4.5	6	11.6	13.4	5.7	6.2	10.9	6.6	6.8	5.5
ES	Winter	6.2	11.1	4.3	9.5	5	5.3	13.1	7.3	13.1	4.7	4.6	9.9	5.8
	Spring	5.8	5	12.8	7.1	5.6	4.2	5.7	7	5.4	19.4	4.6	5.6	11.6
	Summer	4.9	5.3	21.8	9.8	3.8	2.9	12.8	6.3	4.7	3.6	4.6	3.2	16.2
	Autumn	5.2	11	3.2	3.8	6.3	10.4	17.6	7.2	3.6	16.4	4.2	8.2	2.9
SS	Winter	6.4	7.4	11.3	11	6.8	6	7.3	8.8	8.2	6.2	6.2	6.3	8.2
	Spring	7.3	7.3	9.1	8.9	8.4	6.3	7.2	9.4	7.9	7	6.1	7.2	8
	Summer	6.7	7.9	10.2	12.6	5.6	4.6	10.2	8.3	7.6	4.5	5.8	5.1	10.9
	Autumn	7.4	6.4	9.5	5.1	5.5	11.6	13.8	6.5	5.6	10.6	6	6.1	5.8
NEE	Winter	6.4	4.7	4.8	8.2	14.5	4	5.2	26.3	4.8	4.8	5.9	4.5	5.7
	Spring	7.1	6.7	10.9	7.7	7.2	5.4	9.3	8.1	7.7	10.7	6.2	6.7	6.4
	Summer	5.5	5.7	7.6	10.2	4.5	3.2	29.2	6	5.2	4.5	5.2	3.8	9.3
	Autumn	4.8	3.3	7.8	1.7	2.3	9.4	15.4	33.4	3	9.1	3.3	3.6	2.8
NWE	Winter	8.5	6.8	9.3	11.6	7.7	6	7.2	6.6	6.4	7.8	7	6.4	8.8
	Spring	10	6.6	9.5	9	7.5	5.8	7.6	7.1	6.8	7.9	6.6	7.8	7.8
	Summer	7.5	6.8	8.8	12.3	5.5	4.2	13.1	7	6.3	5.9	6	5.4	11.1
	Autumn	7.4	5.4	10.2	5.2	5.1	11.3	13.7	5.9	5.4	12.6	6.1	6.1	5.7
CEE	Winter	4	8.3	4	9	3.9	4	5.7	3.4	4.7	34.8	4.1	5.6	8.4
	Spring	4.7	4.5	5.3	9.8	5.1	3.8	5.6	36.3	5.2	7.1	3.5	4.4	5
	Summer	11.5	6.8	11.3	12.2	5.3	3.7	10.5	7.6	5.6	5.1	6.1	4.4	9.8
	Autumn	4.6	3.2	2.7	4.2	2.5	9.1	10	1.4	38.4	12.3	5.4	3.3	2.9
SEE	Winter	4	35.2	3.9	9.6	4.2	7.2	4.7	3.5	6.4	4.5	5.7	3.9	7.2
	Spring	5.9	5.8	7.6	10.5	7.1	4.4	5.9	19.4	8.1	6.6	7.9	5	5.9
	Summer	10.3	6.5	12.3	11.9	4.9	3.8	9.8	8.3	5.7	6	5.3	4.8	10.2
	Autumn	6.3	4.4	29	3.1	4.3	10	12.3	6.1	3.5	9.4	3.9	4.2	3.4
SWE	Winter	8.2	6.9	8.9	11.6	7.4	5.8	7.2	6.5	6.9	8.8	6.9	6.6	8.2
	Spring	8.3	6.6	9.1	9.4	8.2	6	6.3	8.6	7.2	9.4	6.5	7.3	7.1
	Summer	8.3	7	9.2	12.1	5.6	4.6	10.8	8	6.6	5.6	6.1	5.5	10.5
	Autumn	8	5.6	7.7	5.5	5.5	11.5	14.1	6.7	5.9	11.7	6.5	6.2	5.1
NI	Winter	6.4	5.4	4	9.9	9.6	3.6	19.4	4.5	7.8	5.8	3.7	13.8	6.2
	Spring	6.1	5.2	5.6	8.9	7.4	4.5	11.9	5.9	7.8	9.4	9.9	10.3	7.1
	Summer	6	6.3	13	10.4	3.7	2.9	20.1	5.9	5.6	5.3	4	4.1	12.5
	Autumn	5.8	4.2	20.2	6.5	4.9	10	12.8	5.2	4.9	10.2	4.8	4.6	6

Column *Region* gives rainfall region and column *Season* gives meteorological season. Values are given for the 1-day event duration.

DOI: 10.1002/joc

Table IV. Model weights (%) for the 5-year return value.

Region	Season	ARPH	CHRMH	CLMH	HADH	HIRE	HIRH	REGH	METH	REMOH	RACH	RCAOE	RCAOH	PROMH
NS	Winter	6.4	6.7	27	6.5	5.6	4.8	5.5	6.2	9.7	5.2	5.2	5.4	5.9
	Spring	8.1	6.9	12.2	7.5	7.2	5.9	6.6	7.3	9.9	6.9	7.1	7.1	7.4
	Summer	6.1	6.7	11.6	7.3	6.1	5.4	7.2	7.5	11.1	7.5	7.2	6.8	9.7
	Autumn	7.3	7.1	10.8	8.2	8.6	6.3	6.7	7.6	9.2	6.6	7.5	6.8	7.2
ES	Winter	7.1	6.2	13.4	5.3	18.9	16.9	5.8	5.1	5.8	3.3	3.1	3.2	5.8
	Spring	6.9	3.7	7	4.2	3.9	2.9	4.4	3.8	45.4	4.2	4.9	4.4	4.2
	Summer	3.8	4	9.8	4.8	3.5	3	37.1	4.1	8.9	4.9	5.2	4.5	6.5
	Autumn	5.3	4.7	8.6	6.3	7.7	4.6	6	5.4	9.2	6.7	7.4	21.1	7
SS	Winter	7	6.6	21.8	7.1	5.8	5.2	6	8.5	10.6	5	5.2	5.3	5.9
	Spring	7.8	6.8	12	7.5	7.4	6	6.8	7.9	9.9	6.7	7	7.1	7.2
	Summer	6	6.6	11	7.5	6.6	5.9	7.6	7.4	10.5	7.4	7.2	7.2	9.2
	Autumn	8	6.9	10	8.5	8.9	6.5	6.6	8.6	8.9	6	7.6	6.3	7.3
NEE	Winter	7	4.2	14.8	4.9	4.7	3.4	5.1	8.4	16.5	6.4	6.7	13.8	4.1
	Spring	8.5	5.3	12.3	5.7	6.6	4.9	8.3	6.5	14.7	7.7	7.4	6.6	5.6
	Summer	5.8	5.6	8.9	6.1	5.9	5.2	16.1	6	10	7.9	7.7	7	7.8
	Autumn	5.1	4.3	14.3	5.7	6.9	4.2	5.1	5.7	7.6	4.8	26	5.4	5
NWE	Winter	7.9	5.2	28.2	6.3	5.2	4.2	5.2	5.3	8	5.7	6.8	6.6	5.4
	Spring	10.9	5.8	12.4	6.5	7.2	5.1	6.8	6.5	9.2	7.7	7.4	8.2	6.3
	Summer	6	6.3	11.2	6.6	5.8	5.2	8.6	6	10.1	9	8.8	7.8	8.5
	Autumn	7.4	6.1	10.4	7.9	9.1	5.9	6.7	6.6	8.2	6.9	11.3	6.9	6.8
CEE	Winter	4.3	13.4	13.7	10.7	3.7	3.7	4.6	3.7	7.1	7.5	4.2	6	17.3
	Spring	7.6	5.1	16.6	7.1	11.6	4.5	6.7	9.5	8	5.9	5.6	5.6	6.1
	Summer	6.7	6.1	13.6	7.3	5.9	5	8.3	7.4	9.7	7.5	7.3	6.8	8.4
	Autumn	4.5	5.9	4.7	18.2	5.2	4.9	4.3	3.8	17	7.6	4.5	10.6	8.8
SEE	Winter	4.6	6.1	14.1	6.9	6.1	7.8	6.4	5.1	21.8	4.8	6.8	4.8	4.8
	Spring	7.8	4.9	8.4	7.1	24.5	4.3	5.5	10	7.8	5.2	4.7	4.7	5.3
	Summer	4.3	4.4	38.5	5.6	3.8	3.4	5.2	5.2	7.5	5.6	4.9	4.9	6.7
	Autumn	17.4	4.6	10.6	6.2	19.2	4.6	4.5	8	5.8	4.4	5.6	4.3	4.8
SWE	Winter	8.2	5.8	20.1	7	6.2	4.8	5.6	6	9.2	6.4	7.8	6.7	5.9
	Spring	9	5.7	15.1	6.4	7	5	6.2	6.7	9.8	7.8	7.5	7.7	5.9
	Summer	6.6	6.3	10.4	7	5.9	5.4	8.4	6.6	10.4	8.6	8.3	7.6	8.4
	Autumn	7.8	5.8	10.5	7.7	10.1	5.7	6.6	6.9	8.1	6.6	11.4	6.5	6.3
NI	Winter	5.1	4.8	19.4	5.5	5.1	3.5	5.3	5	9.2	4.4	14.3	13.6	4.6
	Spring	7.4	5.1	15.8	5.6	6.2	4.2	7.5	5.4	17.7	5.8	7.3	6.8	5.2
	Summer	4.1	4.7	10.4	5	4	3.6	8.7	4.8	31	6.1	5.2	5.4	7.1
	Autumn	5.5	5.1	14.9	6.9	7.6	4.7	6	5.8	7.9	5.4	19	5.6	5.6

Column *Region* gives rainfall region and column *Season* gives meteorological season. Values are given for the 10-day event duration.

DOI: 10.1002/joc

Table V. Model weights (%) for the 25-year return value.

Region	Season	ARPH	CHRMH	CLMH	HADH	HIRE	HIRH	REGH	METH	REMOH	RACH	RCAOE	RCAOH	PROMH
NS	Winter	6.4	6.9	22.9	6.7	5.5	5.1	5.7	6.4	12.1	5.3	5.3	5.5	6.2
	Spring	7.6	7	12.6	7.3	6.4	6.4	6.9	7.2	9.9	7	7	6.9	7.9
	Summer	5.9	5.6	12.8	8.7	5.5	5.4	6.6	7.8	11.1	7	6.7	6.4	10.7
	Autumn	6.8	6.9	10.1	9.9	8.5	6.3	6.6	7.6	9.8	6.6	7.1	6.7	7.1
ES	Winter	7.1	6.4	9.3	5.5	18.7	17.2	6.1	5.4	8.2	3.4	3.2	3.4	6.1
	Spring	6.4	3.9	7.3	4	3.1	3.4	4.7	3.7	45.4	4.3	4.8	4.2	4.7
	Summer	3.6	2.9	11	6.2	2.9	2.9	36.6	4.5	8.8	4.5	4.8	4	7.4
	Autumn	4.8	4.5	7.9	8	7.6	4.6	5.9	5.5	9.8	6.7	7	21	6.9
SS	Winter	7	6.8	17.7	7.3	5.7	5.5	6.3	8.8	13	5.2	5.3	5.4	6.3
	Spring	7.3	6.9	12.3	7.3	6.6	6.5	7.1	7.8	9.9	6.7	6.9	6.9	7.7
	Summer	5.8	5.5	12.2	8.9	6	5.8	7	7.7	10.4	6.9	6.7	6.7	10.1
	Autumn	7.6	6.7	9.3	10.1	8.8	6.5	6.5	8.6	9.5	5.9	7.2	6.1	7.2
NEE	Winter	6.9	4.4	10.7	5.1	4.6	3.6	5.4	8.6	18.9	6.5	6.7	14	4.4
	Spring	8.1	5.4	12.6	5.6	5.8	5.3	8.6	6.4	14.7	7.8	7.3	6.4	6.1
	Summer	5.6	4.5	10.1	7.6	5.3	5.1	15.6	6.3	9.9	7.4	7.3	6.5	8.7
	Autumn	4.6	4.1	13.6	7.3	6.7	4.1	5	5.7	8.2	4.8	25.5	5.3	4.9
NWE	Winter	7.8	5.4	24.1	6.5	5.1	4.5	5.4	5.5	10.4	5.8	6.9	6.8	5.8
	Spring	10.4	6	12.8	6.3	6.4	5.6	7.1	6.4	9.2	7.7	7.3	8	6.8
	Summer	5.8	5.3	12.3	8.1	5.2	5.1	8.1	6.3	10.1	8.5	8.4	7.3	9.5
	Autumn	7	5.9	9.7	9.5	8.9	5.8	6.5	6.6	8.8	6.9	10.8	6.8	6.8
CEE	Winter	4.3	13.7	9.6	10.8	3.6	4	4.9	3.9	9.5	7.7	4.3	6.1	17.6
	Spring	7.1	5.3	16.9	6.9	10.9	5	7	9.4	8	5.9	5.6	5.4	6.6
	Summer	6.5	5	14.8	8.8	5.3	4.9	7.7	7.7	9.7	7	6.9	6.3	9.4
	Autumn	4	5.7	4	19.9	5.1	4.9	4.2	3.8	17.6	7.6	4.1	10.4	8.8
SEE	Winter	4.5	6.4	10	7.1	6	8	6.6	5.3	24.3	4.9	6.9	4.9	5.2
	Spring	7.3	5	8.7	6.9	23.7	4.8	5.8	9.9	7.8	5.2	4.6	4.6	5.8
	Summer	4.1	3.4	39.7	7.1	3.2	3.3	4.6	5.5	7.5	5.1	4.4	4.4	7.6
	Autumn	17	4.4	10	7.8	19	4.5	4.3	8	6.4	4.4	5.2	4.2	4.8
SWE	Winter	8.2	6.1	16	7.2	6.1	5.1	5.9	6.3	11.7	6.6	7.8	6.9	6.3
	Spring	8.5	5.9	15.5	6.2	6.2	5.5	6.5	6.6	9.8	7.8	7.5	7.5	6.4
	Summer	6.4	5.3	11.6	8.4	5.3	5.3	7.8	7	10.4	8.1	7.9	7.2	9.4
	Autumn	7.3	5.6	9.8	9.3	10	5.6	6.5	6.9	8.7	6.6	11	6.3	6.3
NI	Winter	5.1	5	15.3	5.7	5	3.8	5.6	5.3	11.6	4.5	14.4	13.8	5
	Spring	6.9	5.2	16.1	5.4	5.5	4.7	7.8	5.3	17.7	5.8	7.2	6.6	5.7
	Summer	3.9	3.6	11.6	6.5	3.4	3.6	8.1	5.1	30.9	5.6	4.7	4.9	8.1
	Autumn	5	4.9	14.2	8.6	7.4	4.7	5.8	5.8	8.5	5.4	18.6	5.4	5.6

Column *Region* gives rainfall region and column *Season* gives meteorological season. Values are given for the 10-day event duration.

DOI: 10.1002/joc

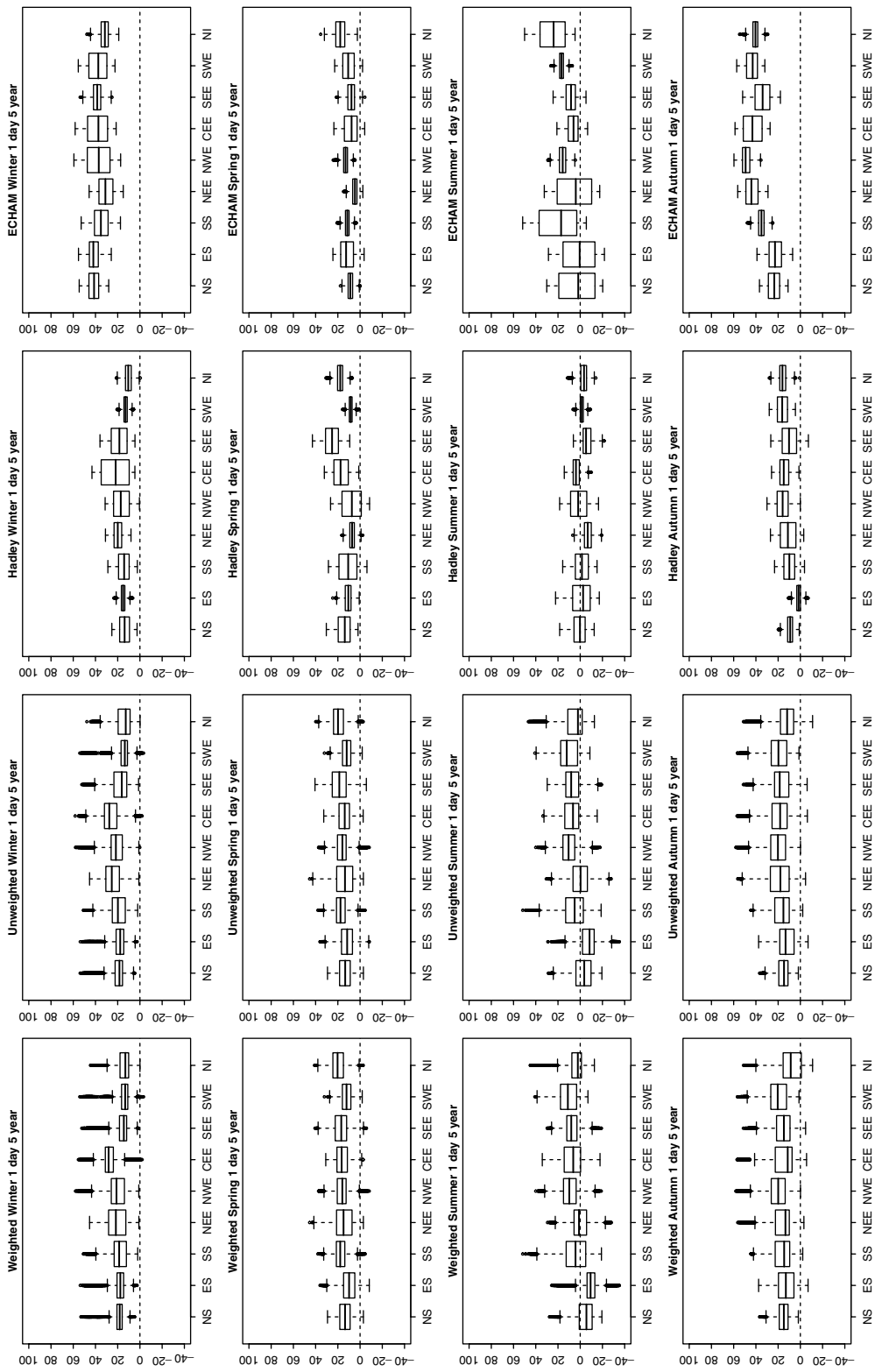

Figure 11. Estimates of percent change in the 1-day 5-year return value for the SRES A2 2071–2100 scenario for each of the nine UK homogenous rainfall regions. Each row of results shows a different season, from top to bottom: winter (DJF), spring (MAM), summer (JJA) and autumn (SON). For each row: column 1 shows pooled results from all RCMs with weighting estimates from semi-variograms; column 2 shows pooled results from all RCMs assuming equal weighting; column 3 shows pooled results for HIRH and RCAOH and column 4 shows pooled results for HIRE and RCAOE. 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. A dashed line indicates the zero change line.

DOI: 10.1002/joc

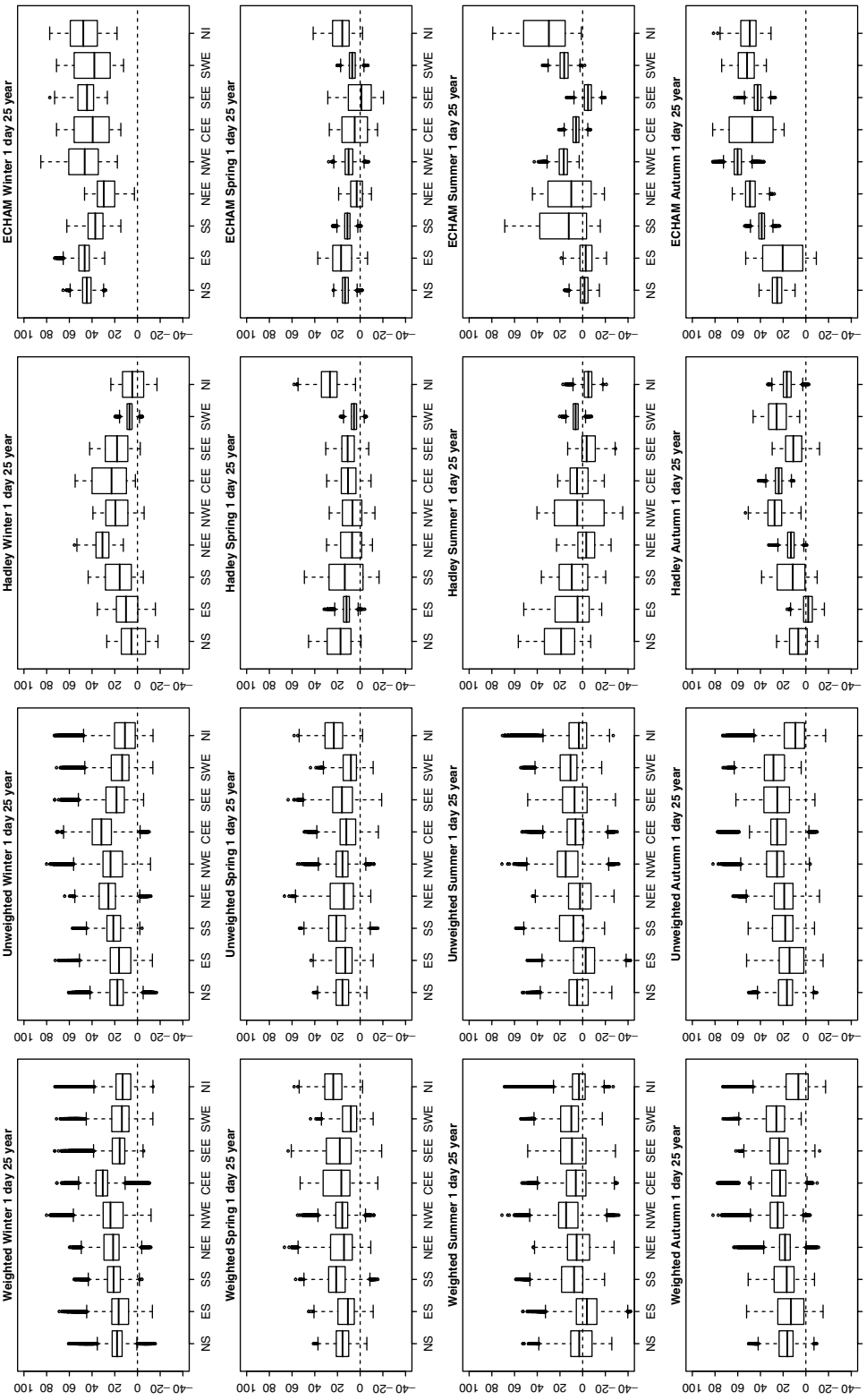

Figure 12. Estimates of percent change in the 1-day 25-year return value for the SRES A2 2071–2100 scenario for each of the nine UK homogenous rainfall regions. Each row of results shows a different season, from top to bottom: winter (DJF), spring (MAM), summer (JJA) and autumn (SON). For each row: column 1 shows pooled results from all RCMs with weighting estimates from semi-variograms; column 2 shows pooled results from all RCMs assuming equal weighting; column 3 shows pooled results for HIRH and RCAOH and column 4 shows pooled results for HIRE and RCAOE. Box plot details are the same as for Figure 11.

DOI: 10.1002/joc

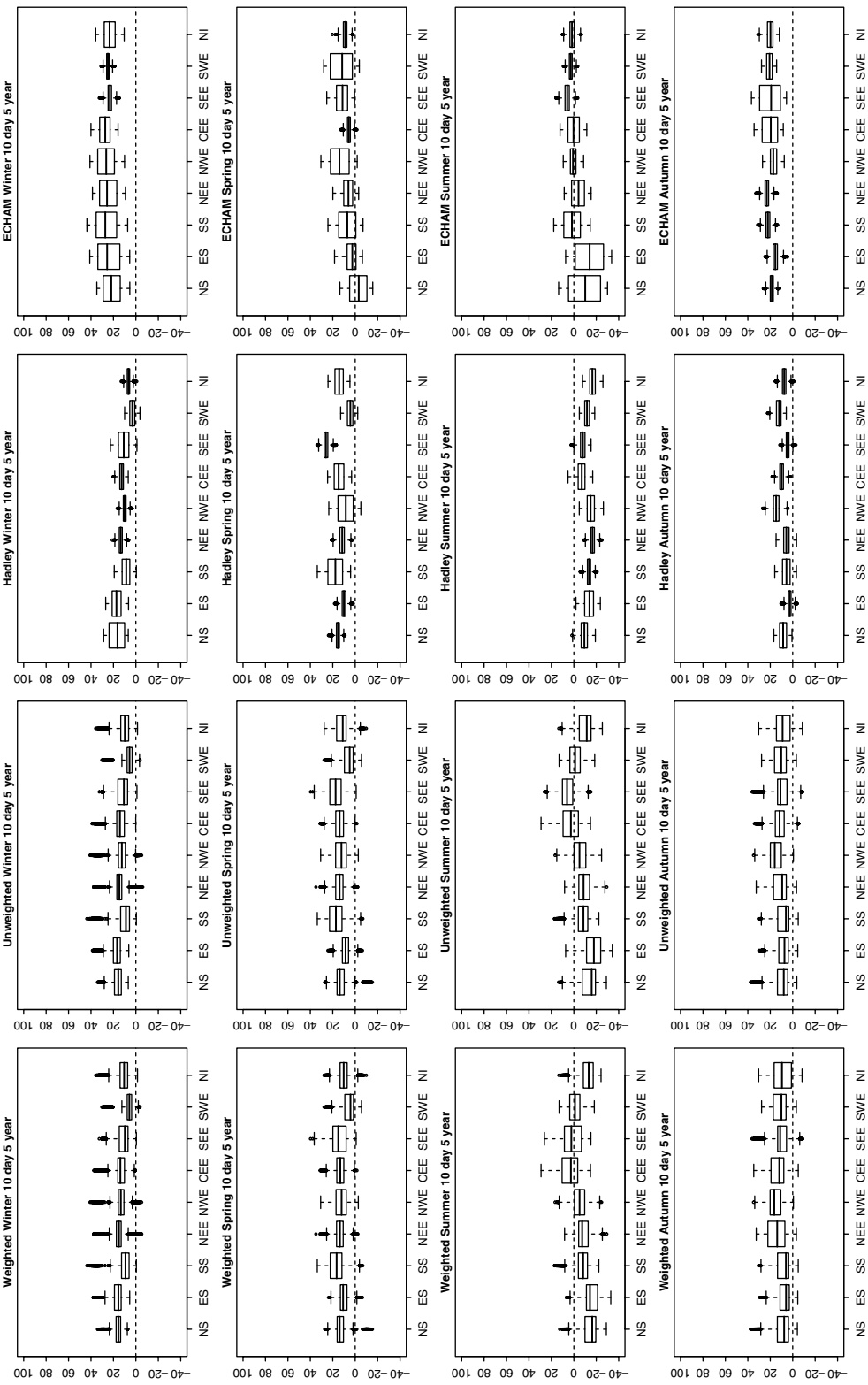

Figure 13. Estimates of percent change in the 10-day 5-year return value for the SRES A2 2071–2100 scenario for each of the nine UK homogenous rainfall regions. Each row of results shows a different season, from top to bottom: winter (DJF), spring (MAM), summer (JJA) and autumn (SON). For each row: column 1 shows pooled results from all RCMs with weighting estimates from semi-variograms; column 2 shows pooled results from all RCMs assuming equal weighting; column 3 shows pooled results for HIRH and RCAOH and column 4 shows pooled results for HIRE and RCAOE. Box plot details are the same as for Figure 11.

DOI: 10.1002/joc

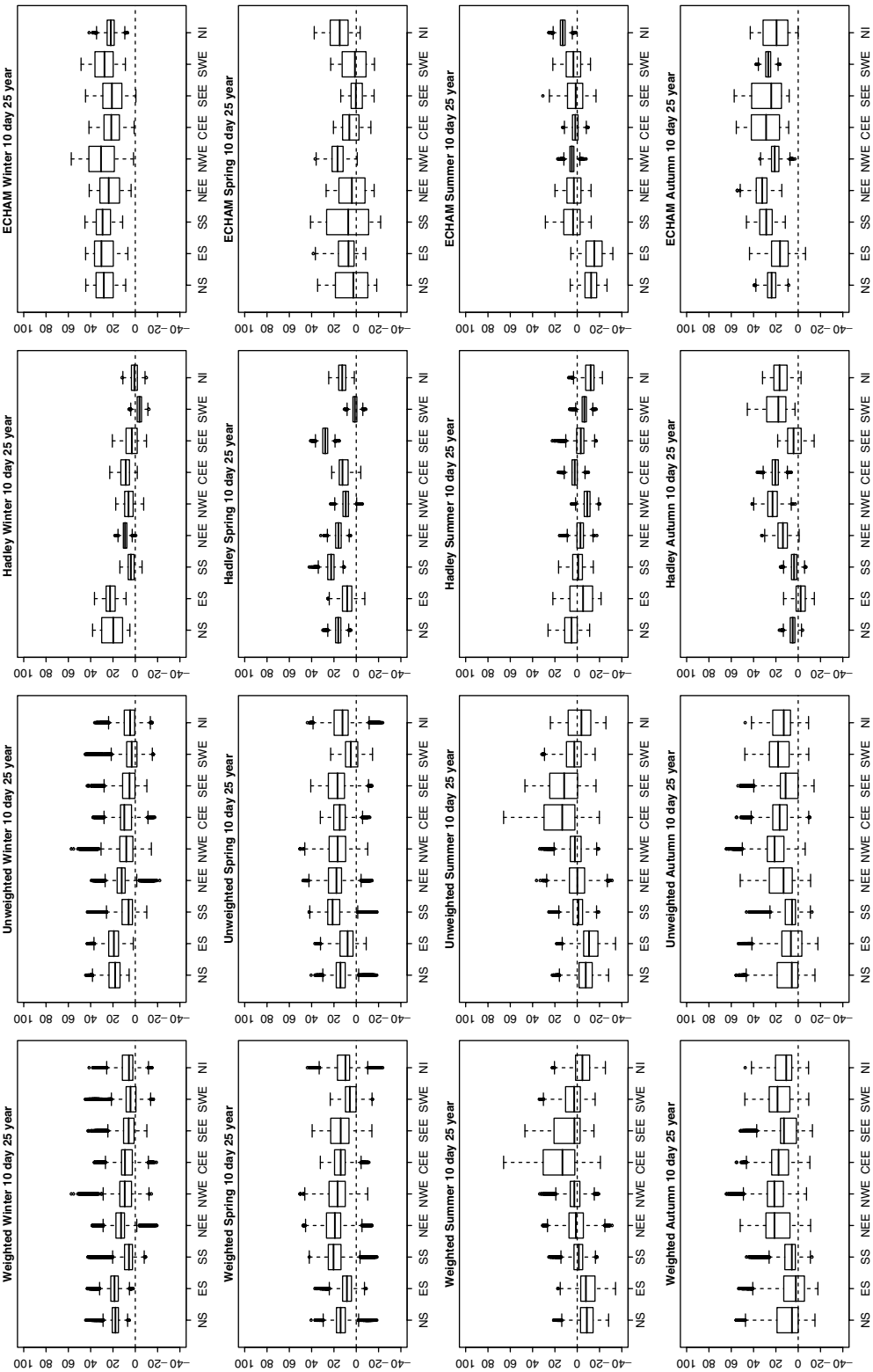

Figure 14. Estimates of percent change in the 10-day 25-year return value for the SRES A2 2071–2100 scenario for each of the nine UK homogenous rainfall regions. Each row of results shows a different season, from top to bottom: winter (DJF), spring (MAM), summer (JJA) and autumn (SON). For each row: column 1 shows pooled results from all RCMs with weighting estimates from semi-variograms; column 2 shows pooled results from all RCMs assuming equal weighting; column 3 shows pooled results for HIRH and RCAOH and column 4 shows pooled results for HIRE and RCAOE. Box plot details are the same as for Figure 11.

DOI: 10.1002/joc

distributions, see for example summer for the 10-day 25-year return value (Figure 14). Irrespective of weighting scheme, the distributions associated with the 25 return values tend to be wider than those for the 5-year return value.

Regional differences are most pronounced in the weighted multi-model ensembles during summer. Although most distributions span the 0% change line in summer, for Scotland, NEE and NI, a significant proportion of each distribution shows an expected decrease in extremes, particularly for longer duration events. However, since RCMs cannot adequately simulate summer precipitation extremes, we have little confidence in their projections. In other seasons, there is less regional variability and the expected change is generally of the same sign; predominately positive changes are predicted from the multi-model ensemble, although less so for 10-day events. We have much confidence in the multi-model projections for winter as extreme precipitation in this season is reasonably well simulated by RCMs. In winter, the multi-model ensemble suggests changes of 15–30% in the 1-day 5-year event across the UK under the SRES A2 2070–2100 emissions scenario; similar increases are projected, albeit with larger uncertainty, for the higher 25-year return period. For longer duration extremes, distributions are narrower and so we have greater confidence in the projections. For the 10-day 5-year event in winter, increases of 5–20% are projected across the UK, with a magnitude of change for the 25-year return period similar to that for the 1-day event. In spring, for the 5-year return period the projected range of change is from 10 to 20% in the 1-day return value and 10–25% in the 10-day return value; projections are similar for the 25-year return period. In autumn, the range of change is for 10–25% increases in the 1-day 5-year return value and 5–20% increases in the 10-day 5-year return value. In autumn, it is notable that ECHAM-driven RCMs predict much larger increases than Hadley-driven RCMs, but the ensemble membership, where 11 of 13 RCMs are driven by Hadley models, implies that the multi-model ensemble is skewed towards more conservative increases.

The rather limited differences in the weighted and unweighted multi-model ensembles of percent change may be a consequence of a lack of model independence (Tebaldi and Knutti, 2007) and has been noted in other studies (e.g. Manning *et al*., in press). The comparison of RCMs driven by lateral boundary conditions from ECHAM4/OPYC with those using Hadley Centre GCMs (HadAM3H/P or HadCM3) suggests that a large proportion of uncertainty in the distributions stems from the driving lateral boundary conditions. Furthermore, many RCMs contain structural similarities, for example, METNO is the Norwegian version of HIRHAM, and many other RCMs share parameterization schemes. Consequently, although we have shown that semi-variogram parameters and regionally averaged Rmed discrepancy can be used to weight the combination of RCMs in a multi-model ensemble and this is shown to influence predictions of extreme precipitation in winter and spring, our results suggest that the largest uncertainty in projections of change to future precipitation extremes is linked to the driving lateral boundary conditions of the RCM (see Figures 11–14, columns 3 and 4). Using a combined weight solely on the basis of the spatial variability of extremes and their regionally averaged Rmed discrepancy in relation to observations is therefore not entirely satisfactory. Here, we have used an application-specific weighting; however, it is likely that a multi-scale approach to weighting, assessing not only the simulation of synoptic-scale regional climate, but also the simulation of continental-scale and global modes of variability, may be more appropriate. Hence, the uncertainty bounds shown here are in many respects still conservative despite the relatively large number of RCMs contributing to the multi-model ensembles.

Nevertheless, importantly for policy makers, the multi-model ensembles of change project increases in extreme precipitation for most UK regions in winter, spring and autumn. This change is physically consistent with warmer air in the future climate being able to hold more moisture. The use of multi-day extremes and return periods also showed that short-duration extreme precipitation is projected to increase more than longer-duration extreme precipitation, where the latter is associated with narrower uncertainty ranges. However, for both types of precipitation event there is considerable uncertainty as to the magnitude of change. In summer there is less confidence in RCM projections and results are more varied; for most regions ensembles span the 0% change line indicating potential for both increases and decreases in extremes.

Acknowledgements

Data have been provided through the PRUDENCE data archive, funded by the EU through contract EVk2-CT2001-00132. Data are 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. Thanks to Mr Andy Smith, PhD student at Newcastle, for producing the aggregated 50-km observed precipitation data and Dr Stephen Blenkinsop for re-gridding the RCM data to the common CRU grid for the European Union FP6 Integrated Project AquaTerra (Project no. 505428) under the thematic priority sustainable development, global change and ecosystems. This work was supported by an NERC post-doctoral Fellowship award to Dr Hayley Fowler (2006-2009) NE/D009588/1. Thanks to the two anonymous reviewers whose comments helped to greatly improve this manuscript.

DOI: 10.1002/joc

A1. Technical appendix

A1.1. Regional frequency analysis

RFA usually follows a two part *index-flood* procedure, which is a convenient way of pooling statistics from different samples (Hosking and Wallis, 1997) and can be used for any type of data. If the data are available at N sites (here grid cells), with site i having sample size n_i and observed data X_{ij}, $j = 1, \ldots, n$. Then $X_i(F)$, $0 < F < 1$, forms the frequency distribution's quantile function at site i. In an index-flood procedure, the sites (grid cells) must form a homogeneous region, with identical frequency distributions at the N sites (grid cells) apart from the site-specific scaling factor, the *index-flood* (Hosking and Wallis, 1997).

This is normally checked using the 'discordancy measure', D_i (Hosking and Wallis, 1997), which compares the L-moment ratios of a site (grid cell) with those of the pooling group as a whole, hence identifying sites (grid cells) that are unusual relative to the pooling group. A high value of the discordancy measure indicates that a site (grid cell) may be discordant within the pooling group, or that the record contains a few unusual precipitation events. If M is the number of sites (grid cells) in the pooling group and $\boldsymbol{u}_i$ is a vector of the L-moment ratios at site i, then (after Robson and Reed, 1999) (A1):

$$u_i = (t_1, t_2, t_3)^{\mathrm{T}} \tag{A1}$$

where t_1 is L-CV, t_2 the L-skewness, t_3 the L-kurtosis and superscript T denotes the transpose of a vector. Thus defining two matrices $\boldsymbol{U}$ (Equation (A2)) and $\boldsymbol{A}$ (Equation (A3)) as:

$$\boldsymbol{U} = \frac{1}{M}\sum_{i=1}^{M}\boldsymbol{u}_i \tag{A2}$$

$$\boldsymbol{A} = \sum_{i=1}^{M}(\boldsymbol{u}_i - \boldsymbol{U})(\boldsymbol{u}_i - \boldsymbol{U})^{\mathrm{T}} \tag{A3}$$

Then the discordancy measure D_i for site (grid cell) i is given by (Equation (A4)):

$$D_i = \frac{1}{3}M(\boldsymbol{u}_i - \boldsymbol{U})^{\mathrm{T}}\boldsymbol{A}^{-1}(\boldsymbol{u}_i - \boldsymbol{U}) \tag{A4}$$

where $\boldsymbol{A}^{-1}$ is the inverse of matrix $\boldsymbol{A}$.

Critical values of the discordancy measure for each site (grid cell) in a pooling group based on a 10% significance level are suggested by Hosking and Wallis (1997). For a pooling group with more than 15 members, as in this study, values of D_i higher than 3.0 show possible discordancy. The UK rainfall regions have previously been checked for homogeneity using observed site data by Fowler and Kilsby (2003a).

The index-flood procedure may then be defined as (from Hosking and Wallis, 1997) Equation (A5):

$$X_i(F) = \mathrm{Rmed}_i\ x(F), \quad i = 1, \ldots, N \tag{A5}$$

where Rmed_i is the index-flood (here it is the median of the seasonal maxima frequency distribution for an individual grid cell), and $x(F)$ is the regional growth curve, a quantile function identical at every site (grid cell) within that region.

The site-specific *index-flood variable*, Rmed_i, is naturally estimated for each site (grid cell) as the median [as in the Flood Estimation Handbook (IH, 1999)] of the SM dataset at site (grid cell) i.

Secondly, the *regional growth curve*, $x(F)$, $0 < F < 1$ is derived, using a pooled analysis of the dimensionless rescaled data, $x_{ij} = X_{ij}/\mathrm{Rmed}_i$, $j = 1, \ldots, n_i$, $I = 1, \ldots, N$. Here, L-moments are used to derive the *regional growth curve*. The L-moment ratios of L-CV, L-Skewness and L-Kurtosis are derived for each site (grid cell) within a region and then combined by regional averaging (as Hosking and Wallis, 1997). Thus, giving an example formula for L-CV (Equation (A6)):

$$\mathrm{LCV}_{\mathrm{pooled}} = \sum_{i=1}^{N} w_i \mathrm{LCV}_i \tag{A6}$$

where N is the number of sites (grid cells) in the pooling group and the weight w_i is an effective record length at the ith site (grid cell) defined by Equation (A7):

$$w_i = \frac{n_i}{\sum_{i=1}^{N} n_i} \tag{A7}$$

The denominator is the total number of station-years of record in the pooling group, while the numerator is the number of station-years at the ith site (grid cell). Since in this analysis all grid cells have the same record length, all of the weights, w_i, are equal. The L-Skewness and L-Kurtosis moment ratios are derived in the same way (see Hosking and Wallis, 1997, for details).

L-moments are then used to fit the GEV distribution for each standardized SM dataset by matching the sample L-moments to the distribution L-moments.

The GEV distribution has three parameters and is described by Equation (A8):

$$x(F) = \xi + \frac{\alpha}{k}[1 - (-\ln\ F)^k]\ (k \neq 0) \tag{A8}$$

where ξ is the location parameter, α the scale parameter, k the shape parameter and F refers to a given quantile.

A *regional growth curve* was fitted for each region using the regionally averaged L-moment ratios. The fitted growth curve is given by Equation (A9):

$$x(F) = 1 + \frac{\beta}{k}[(\ln\ 2)^k - (-\ln\ F)^k] \tag{A9}$$

where

$$\beta = \frac{\alpha}{[\xi + (\alpha/k)][1 - (\ln\ 2)^k]} \tag{A10}$$

DOI: 10.1002/joc

The parameter k is estimated from the L-skewness (Hosking *et al*., 1985) (Equation (A11)):

$$k \approx 7.8590c + 2.9554c^2 \tag{A11}$$

where

$$c = \frac{2}{3 + t_3} - \frac{\ln 2}{\ln 3} \tag{A12}$$

The parameter β is estimated using L-CV (Hosking and Wallis, 1997) as Equation (A13):

$$\beta = \frac{kt_2}{t_2[\Gamma(1+k) - (\ln 2)^k] + \Gamma(1+k)(1-2^{-k})} \tag{A13}$$

where Γ denotes the gamma function, t_2 the L-CV L-moment ratio and t_3 the L-Skewness L-moment ratio.

Quantile estimates at site (grid cell) i can then be obtained by combining the estimates of Rmed_i and $x(F)$ as Equation (A14):

$$X_i(F) = \text{Rmed}_i \; x(F) \tag{A14}$$

References

Alexander LV, Zhang X, Peterson TC, Caesar J, Gleason B, Klein Tank AMG, Haylock M, Collins D, Trewin B, Rahimzadeh F, Tagipour A, Ambenje P, Rupa Kumar K, Revadekar J, Griffiths G. 2006. Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research* **111**: D05109, DOI:10.1029/2005JD006290.

Arribas A, Gallardo C, Gaertner MA, Castro M. 2003. Sensitivity of Iberian Peninsula climate to land degradation. *Climate Dynamics* **20**: 477–489.

Beniston M, Stephenson DB, Christensen OB, Ferro CAT, Frei C, Goyette S, Halsnaes K, Holt T, Jylhä K, Koffi B, Palutikof J, Schöll R, Semmler T, Woth K. 2007. Future extreme events in European Climate: An exploration of regional climate model projections. *Climatic Change* **81**(Suppl. 1): 81–95.

Blenkinsop S, Fowler HJ. 2007. Changes in drought frequency and severity over the British Isles projected by the PRUDENCE regional climate models. *Journal of Hydrology* **342**: 50–71.

Brunetti M, Buffoni L, Maugeri M, Nanni T. 2000. Precipitation intensity trends in northern Italy. *International Journal of Climatology* **20**: 1017–1031.

Castro M, Fernández C, Gaertner MA. 1993. Description of a meso-scale atmospheric numerical model. In *Mathematics, Climate and Environment*, Díaz JI, Lions JL (eds). Masson: Paris; 230–253.

Chilès J-P, Delfiner P. 1999. *Geostatistics: Modelling Spatial Uncertainty*. Wiley Series in Probability and Statistics, Wiley-Interscience publication: New York; 695.

Christensen JH, Carter TR, Rummukainen M, Amanatidis G. 2007. Evaluating the performance and utility of regional climate models: the PRUDENCE project. *Climatic Change* **81**(Suppl. 1): 1–6.

Christensen JH, Christensen OB, Lopez P, van Meijgaard E, Botzet M. 1996. The HIRHAM4 regional atmospheric climate model, DMI Technical Report 96-4. Available from DMI, Lyngbyvej 100, Copenhagen Ø.

Christensen OB, Christensen JH, Machenhauer B, Botzet M. 1998. Very high-resolution regional climate simulations over Scandinavia – present climate. *Journal of Climate* **11**: 3204–3229.

Christensen JH, Christensen OB, Schultz JP. 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.

Covey C, Achuta Rao KM, Cubasch U, Jones P, Lambert SJ, Mann ME, Phillips TJ, Taylor KE. 2003. An overview of results from the Coupled Model Intercomparison Project. *Global and Planetary Change* **37**: 103–133.

Déqué M, Marquet P, Jones RG. 1998. Simulation of climate change over Europe using a global variable resolution general circulation model. *Climate Dynamics* **14**: 173–189.

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

Döscher R, Willén U, Jones C, Rutgersson A, Meier HEM, Hansson M, Graham LP. 2002. The development of the coupled regional ocean-atmosphere model RCAO. *Boreal Environmental Research* **7**: 183–192.

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

Ekström M, Fowler HJ, Kilsby CG, Jones PD. 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. *Journal of Hydrology* **300**: 234–251.

Elía R, Caya D, Côté H, Frigon A, Biner S, Giguère M, Paquin D, Harvey R, Plummer D. 2008. Evaluation of uncertanties in the CRCM-simulated North American climate. *Climate Dynamics* **30**: 113–132.

Fowler HJ, Blenkinsop S, Tebaldi C. 2007a. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology* **27**: 1547–1578.

Fowler HJ, Ekström M, Blenkinsop S, Smith AP. 2007b. Estimating change in extreme European precipitation using a multi-model ensemble. *Journal of Geophysical Research-Atmospheres* **112**: D18104, DOI:10.1029/2007JD008619.

Fowler HJ, Ekström M, Kilsby CG, Jones PD. 2005. New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1. Assessment of control climate. *Journal of Hydrology* **300**: 212–233.

Fowler HJ, Kilsby CG. 2003a. A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000. *International Journal of Climatology* **23**: 1313–1334.

Fowler HJ, Kilsby CG. 2003b. Implications of changes in seasonal and annual extreme rainfall. *Geophysical Research Letters* **30**: 1720, DOI:10.1029/2003GL017327.

Fowler HJ, Tebaldi C, Blenkinsop S. 2008. Probabilistic estimates of climate change impacts on flows in the River Eden, Cumbria. In: Sustainable Hydrology for the 21st Century, Proc. 10th BHS National Hydrology Symposium, Exeter, 416–423.

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

Frei C, Schöll R, Fukutome S, Schmidli J, Vidale PL. 2006. Future change of precipitation extremes in Europe: An intercomparison of scenarios from regional climate models. *Journal of Geophysical Research* **111**: D06105, DOI:10.1029/2005JD005965.

Germann U, Joss J. 2001. Variograms of radar reflectivity to describe the spatial continuity of Alpine precipitation. *Journal of Applied Meteorology* **40**: 1042–1059.

Giorgi F, Francisco R. 2000. Evaluating uncertainties in the prediction of regional climate change. *Geophysical Research Letters* **27**: 1295–1298.

Giorgi F, Huang Y, Nishizawa K, Fu C. 1999. A seasonal cycle simulation over eastern Asia and its sensitivity to radiative transfer and surface processes. *Journal of Geophysical Research* **104**: 6403–6423.

Giorgi F, Marinucci MR, Bates GT. 1993a. Development of a second generation regional climate model (REGCM2). Part I: Boundary layer and radiative transfer processes. *Monthly Weather Review* **121**: 2794–2813.

Giorgi F, Marinucci MR, Bates GT, DeCanio G. 1993b. Development of a second generation regional climate model (REGCM2). Part II: Convective processes and assimilation of lateral boundary conditions. *Monthly Weather Review* **21**: 2814–2832.

Goovaerts P. 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology* **228**: 113–129.

Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TM, Mitchell JFB, Wood RA. 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 Dynamics* **16**: 147–166.

Gringarten E, Deutsch CV. 2001. Teacher's aide, Variogram Interpretation and Modeling. *Mathematical Geology* **33**: 507–534.

Hagemann S, Botzet M, Machenhauer B. 2001. The summer drying problem over south-eastern Europe: Sensitivity of the limited area model HIRHAM4 to improvements in physical parametrization and resolution. *Physics and Chemistry of the Earth Part B-Hydrology Oceans and Atmosphere* **26**: 391–396.

Hanssen-Bauer I, Førland E, Haugen JE, Tveito OE. 2003. Temperature and precipitation scenarios for Norway: comparison of results from dynamical and empirical downscaling. *Climate Research* **25**: 15–27.

Holawe F, Dutter R. 1999. Geostatistical study of precipitation series in Austria: time and space. *Journal of Hydrology* **219**: 70–82.

Hosking JRM, Wallis JR. 1997. *Regional Frequency Analysis: An Approach Based on L-Moments*. Cambridge University Press: Cambridge; 224.

Hosking JRM, Wallis JR, Wood EF. 1985. Estimation of the generalised extreme-value distribution by the method of probability-weighted moments. *Technometrics* **27**: 251–261.

Isaaks EH, Srivastava RM. 1989. *Applied Geostatistics*. Oxford University Press: New York; 561.

Jacob D. 2001. A note to the simulation of the annual and inter-annual variability of the water budget over the Baltic Sea drainage basin. *Meteorology and Atmospheric Physics* **77**: 61–73.

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

Johns TC, Gregory JM, Ingram WJ, Johnson CE, Jones A, Lowe JA, Mitchell JFB, Roberts DL, Sexton DHL, Stevenson DS, Tett SFB, Woodage MJ. 2003. Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emission scenarios. *Climate Dynamics* **20**: 583–612.

Jones RG, Murphy JM, Hassell DC, Woodage MJ. 2005. A high resolution atmospheric GCM for the generation of regional climate scenarios, Hadley Centre Technical Note 63, Met Office, Exeter.

Jones RG, Noguer M, Hassell DC, Hudson D, Wilson SS, Jenkins GJ, Mitchell JFB. 2004a. *Generating high resolution climate change scenarios using PRECIS*, Tech. report available from Met. Office, Hadley Centre, Exeter, 35.

Jones CG, Ullerstig A, Willén U, Hansson U. 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.

Kendon EJ, Rowell DP, Jones RG, Buonomo E. 2008. Robustness of future changes in local precipitation extremes. *Journal of Climate* **21**: 4280–4297, DOI: 10.1175/2008JCLI2082.1.

Krishnamurti TN, Kishtawal CM, Zhang Z, Larow T, Bachiochi D, Williford E, Gadgil S, Surendran S. 2000. Multimodel ensemble forecasts for weather and seasonal climate. *Journal of Climate* **13**: 4196–4216.

Kyriakidis PC, Kim J, Miller NL. 2001. Mapping of precipitation from rain gauge data using atmospheric and terrain characteristics. *Journal of Applied Meteorology* **40**: 1855–1877.

Lenderink G, van den Hurk B, van Meijgaard E, van Ulden AP, Cuijpers J. 2003. Simulation of present-day climate in RACMO2: first results and model developments, *KNMI Technical Report 252*, Available from KNMI, Postbus 201, 3730 AE, De Bilt, Netherlands, 24.

Lopez A, Tebaldi C, New M, Stainforth D, Allen M, Kettleborough J. 2006. Two approaches to quantifying uncertainty in global temperature changes. *Journal of Climate* **19**: 4785–4796.

Lüthi D, Cress A, Davies HC, Frei C, Schär C. 1996. Interannual variability and regional climate simulations. *Theoretical and Applied Climatology* **53**: 185–209.

Manning L, Hall JW, Fowler HJ, Kilsby CG, Tebaldi C. Using probabilistic climate change information from a multi-model ensemble for water resources assessment. *Water Resources Research* (in press).

Meier HEM, Döscher R, Faxén T. 2003. A multiprocessor coupled ice-ocean model for the Baltic Sea. Application to the salt inflow. *Journal of Geophysical Research* **108**(C8): 3273, DOI:10.1029/2000JC000521.

Meehl GA, Arblaster JM, Tebaldi C. 2005. Understanding future patterns of increased precipitation intensity in climate model simulations. *Geophysical Research Letters* **32**: L18719, DOI:10.1029/2005GL023680.

Meehl GA, Stocker TF, Collins WD, Friedlingstein P, Gaye AT, Gregory JM, Kitoh A, Knutti R, Murphy JM, Noda A, Raper SCB, Watterson IG, Weaver AJ, Zhao Z-C. 2007. Global climate projections. In *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds). Cambridge University Press: Cambridge, New York.

Moberg A, Jones PD. 2004. Regional Climate Models simulations of daily maximum and minimum near-surface temperatures across Europe compared with observed station data for 1961–1990. *Climate Dynamics* **23**: 695–715.

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

Pal JS, Small EE, Eltahir EAB. 2000. Simulation of regional – scale water and energy budgets: Representation of subgrid cloud and precipitation processes within RegCM. *Journal of Geophysical Research* **105**: 29579–29594.

Pardo-Igúzquiza E. 1998. Comparison of geostatistical methods for estimating the areal average climatological rainfall mean using data on precipitation and topography. *International Journal of Climatology* **18**: 1031–1047.

Pardo-Igúzquiza E, Grimes DIF, Teo C-K. 2006. Assessing the uncertainty associated with intermittent rainfall fields. *Water Resources Research* **42**: W01412, DOI: 10.1029/2004WR003740.

Perry M, Hollis D. 2005a. The development of a new set of long-term climate averages for the UK. *International Journal of Climatology* **25**: 1023–1039.

Perry M, Hollis D. 2005b. The generation of monthly gridded datasets for a range of climatic variables over the UK. *International Journal of Climatology* **25**: 1041–1054.

Pope VD, Gallani ML, Rowntree PR, Stratton RA. 2000. The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3. *Climate Dynamics* **16**: 123–146.

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

Ribero PJ Jr, Diggle PJ. 2001. geoR: A package for geostatistical analysis. *R-NEWS* **1**: 1609–3631.

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

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

Robson A, Reed D. 1999. Flood Estimation Handbook. Volume 3. Statistical Procedures for Flood Frequency Estimation. Institute of Hydrology: Wallingford.

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

Shao J, Tu D. 1995. *The Jackknife and Bootstrap*. Springer-Verlag: New York; 516.

Steppeler J, Doms G, Schättler U, Bitzer HW, Gassmann A, Damrath U, Gregoric G. 2003. Meso-gamma scale forecasts using the nonhydrostatic model LM. *Meteorology and Atmospheric Physics* **82**: 75–96.

Tebaldi C, Hayhoe K, Arblaster JM, Meehl GA. 2006. Going to the extremes; An intercomparison of model-simulated historical and future changes in extreme events. *Climatic Change* **79**: 185–211.

Tebaldi C, Knutti R. 2007. The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A* **365**: 2053–2075.

Tebaldi C, Mearns LO, Nychka D, Smith RL. 2004. Regional probabilities of precipitation change: A Bayesian analysis of

multimodel simulations. *Geophysical Research Letters* **31**: L24213, DOI:10.1029/2004GL021276.

Tebaldi C, Smith RL, Nychka D, Mearns LO. 2005. Quantifying uncertainty in projections of regional climate change: a Bayesian approach to the analysis of multi-model ensembles. *Journal of Climate* **18**: 1524–1540.

Tiedtke M. 1989. A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly Weather Review* **117**: 1779–1800.

Tiedtke M. 1993. Representation of clouds in large-scale models. *Monthly Weather Review* **121**: 3040–3061.

Trenberth KE, Jones PD, Ambenje P, Bojariu R, Easterling D, Klein Tank A, Parker D, Rahimzadeh F, Renwick JA, Rusticucci M, Soden B, Zhai P. 2007. Observations: surface and atmospheric climate change. In *Climate Change 2007: The Physical Science Basis*, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds). Cambridge University Press: Cambridge, New York.

Vidale PL, Lüthi D, Frei C, Seneviratne S, Schär C. 2003. Predictability and uncertainty in a regional climate model. *Journal of Geophysical Research* **108**(D18): 4586, DOI: 10.1029/2002JD002810.

Webster R, Oliver M. 2001. *Geostatistics for Environmental Scientists*. Wiley: Chichester; 271.

Wigley TML, Lough JM, Jones PD. 1984. Spatial patterns of precipitation in England and Wales and a revised, homogeneous England and Wales precipitation series. *Journal of Climatology* **4**: 1–25.

Zelterman D. 1993. A semi-parametric bootstrap sample technique for simulating extreme order statistics. *Journal of the American Statistical Association* **88**: 477–485.