



Modelling climate impact on floods with ensemble climate projections

Hannah L. Cloke,^{a,b,*} Fredrik Wetterhall,^b Yi He,^c Jim E. Freer^d and Florian Pappenberger^b

^aDepartment of Geography, King's College London, UK

^bEuropean Centre for Medium-Range Weather Forecasts, Reading, UK

^cTyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia, Norwich, UK

^dSchool of Geographical Sciences, University of Bristol, UK

*Correspondence to: H. L. Cloke, Department of Geography, King's College London, Strand, London WC2R 2LS, UK.
E-mail: hannah.cloke@kcl.ac.uk

The evidence provided by modelled assessments of future climate impact on flooding is fundamental to water resources and flood risk decision making. Impact models usually rely on climate projections from global and regional climate models (GCM/RCMs). However, challenges in representing precipitation events at catchment-scale resolution mean that decisions must be made on how to appropriately pre-process the meteorological variables from GCM/RCMs. Here the impacts on projected high flows of differing ensemble approaches and application of Model Output Statistics to RCM precipitation are evaluated while assessing climate change impact on flood hazard in the Upper Severn catchment in the UK. Various ensemble projections are used together with the HBV hydrological model with direct forcing and also compared to a response surface technique. We consider an ensemble of single-model RCM projections from the current UK Climate Projections (UKCP09); multi-model ensemble RCM projections from the European Union's FP6 'ENSEMBLES' project; and a joint probability distribution of precipitation and temperature from a GCM-based perturbed physics ensemble.

The ensemble distribution of results show that flood hazard in the Upper Severn is likely to increase compared to present conditions, but the study highlights the differences between the results from different ensemble methods and the strong assumptions made in using Model Output Statistics to produce the estimates of future river discharge. The results underline the challenges in using the current generation of RCMs for local climate impact studies on flooding. Copyright © 2012 Royal Meteorological Society

Key Words: GCM; RCM; hydrology; UKCP09; ENSEMBLES; hydrological model; perturbed physics; response surface

Received 15 December 2011; Revised 16 May 2012; Accepted 15 June 2012; Published online in Wiley Online Library

Citation: Cloke HL, Wetterhall F, He Y, Freer JE, Pappenberger F. 2012. Modelling climate impact on floods with ensemble climate projections. *Q. J. R. Meteorol. Soc.* DOI:10.1002/qj.1998

1. Introduction

Evidence from hydrological modelling impact studies based on climate model projections is often used to try to

understand effects on future flooding (e.g. Kay *et al.*, 2006; Bell *et al.*, 2007; Cloke *et al.*, 2010; Prudhomme *et al.*, 2010; Chen *et al.*, 2012). In using global climate model (GCM) projections for assessing future flood impact,

sources of uncertainty include: emission scenario; climate model structure and parametrization; climate projection downscaling and correction techniques; hydrological model structure and parametrization; and observations. The use of *ensembles* of models and techniques is one way to get a handle on and to represent these uncertainties (e.g. Murphy *et al.*, 2004; Stainforth *et al.*, 2007; Cloke and Pappenberger, 2009; Weisheimer *et al.*, 2011), and an ensemble of ensembles can be termed a *grand ensemble* (Pappenberger *et al.*, 2008).

Regional climate models (RCMs) are used to dynamically downscale GCM projections to make them more useful in climate impact studies. However, there remain a number of challenges in producing precipitation projections that are optimal for climate impact studies of flooding (Teutschbein and Seibert, 2010; Beven, 2011). RCMs have a relatively higher resolution than GCMs (~25 km compared to >200 km) while retaining the physical process representation of the climate. However this resolution is too coarse to capture the spatial resolution of precipitation required to effectively model the hydrological processes essential for determining flood risk (<1 km and often smaller as related to hydrological response units). There also remain challenges in correctly representing the physics of precipitation in RCMs and for compensating for missing larger-scale signals of extreme events in the GCMs. Although future increases in resolution will undoubtedly improve representation of precipitation, especially for convective-scale events, significant challenges will remain for the foreseeable future. More realistic precipitation fields for use in local river runoff studies can be produced by using statistical downscaling (Fowler *et al.*, 2007a) or by applying Model Output Statistics (MOS) (Maraun *et al.*, 2010). This correction procedure is also known as *calibration* or *bias correction* and focuses on corrections to moments of the climatology (rather than representation of forecast uncertainty, such as the calibration of ensemble spread and root mean square (RMS) error of the ensemble mean forecast, as is common in weather forecasting). It often includes simple corrections of the mean ('delta' approach), using cut-off thresholds, distribution corrections and increasing the variance (e.g. Yang *et al.*, 2010; Wetterhall *et al.*, 2012). MOS removes much of the model error in the precipitation, making it more useful in impact studies, which is considered by many to allow confidence in the examination of future changes in flow regimes in catchments from Europe and the wider world (e.g. Bell *et al.*, 2007; Fowler *et al.*, 2007b; Leander and Buishand, 2007; Akhtar *et al.*, 2009; Linde *et al.*, 2010; Marke *et al.*, 2011; Rojas *et al.*, 2011; Turco *et al.*, 2011). However, MOS can potentially also remove much of the spread in the driving variables, which could disrupt signals of climate change. There is also no guarantee that such statistics will be valid for future precipitation, especially if the physical processes of precipitation are expected to change. In addition, there are significant uncertainties in historical observations of precipitation used in MOS. Thus there remains an open question as to whether or not MOS should be applied in impact modelling.

The objective of this work is to evaluate the impacts on projected high river flows of differing ensemble approaches and the application of MOS to 'calibrate' precipitation from RCMs, using the example of assessing future climate impact of flood events in the Upper Severn catchment. In this paper we take the grand

ensemble approach and consider: (i) an ensemble of single-model RCM projections from the current UK Climate Projections (UKCP09); (ii) multi-model ensemble RCM projections from the ENSEMBLES project; and (iii) a joint probability distribution of precipitation and temperature from a GCM-based perturbed physics ensemble. The ensemble projections are used together with the HBV hydrological model for observed and future projections of high river discharge in the Upper Severn catchment in the Midlands Region of the UK. First, the UKCP09 and ENSEMBLES are used to directly force the HBV model for continuous future river discharge projections. Both uncorrected and MOS-corrected precipitation ensembles are generated together with consideration of hydrological model parameter uncertainty. Second, a continuous delta response surface technique is used with the HBV model in order to assess the perturbed physics ensemble climate model outputs, which is then compared to response surface results for UKCP09 and ENSEMBLES and also compared to the direct HBV runs.

2. Methods and datasets

2.1. The Upper Severn river catchment, UK, and observed datasets

The town of Shrewsbury, Shropshire, on the Upper River Severn has suffered a number of serious floods, which have caused severe damage to property and disruption to people's lives and livelihoods. Figure 1 depicts the Upper Severn catchment, which is approximately 4062 km², with urban, forest and agricultural land accounting for 3%, 7.1% and 48.5% of the area respectively, and loosely packed peat soil dominating the catchment. River levels are generally high in autumn and winter and low in summer (He *et al.*, 2009). The observed discharge data were provided by the UK Environment Agency (EA) Midlands region and cover 1950–2007. In particular, the gauging station at Montford is important for predicting flooding in the downstream town of Shrewsbury. The digital elevation model of the Upper Severn catchment was obtained from the NEXTMap Britain dataset through the UK NERC Earth Observation Data Centre. The observed precipitation and temperature data used in this study were from the gridded data on a 5 × 5 km grid provided by the UK Met Office. These were interpolated using daily observations as main input, incorporating geographical effects, latitude and longitude, altitude, coastal influence and urban land use through normalization with respect to the monthly 1961–1990 climate (Perry *et al.*, 2008). The spatial distribution of the observations is shown in Figure 1. The accumulated 5-day maximum precipitation (5dmax), expressed as mm day⁻¹, is a useful measure of flood-inducing precipitation in semi-humid meso-scale catchments such as the Severn and was used in this study. It is calculated as the annual maximum precipitation, as a mean over the entire catchment after the data have been filtered with a 5-day running mean filter. The rainfall statistics for the Upper Severn catchment upstream of both Montford and Buildwas are summarized in Table 1, as well as the statistics for the used RCMs.

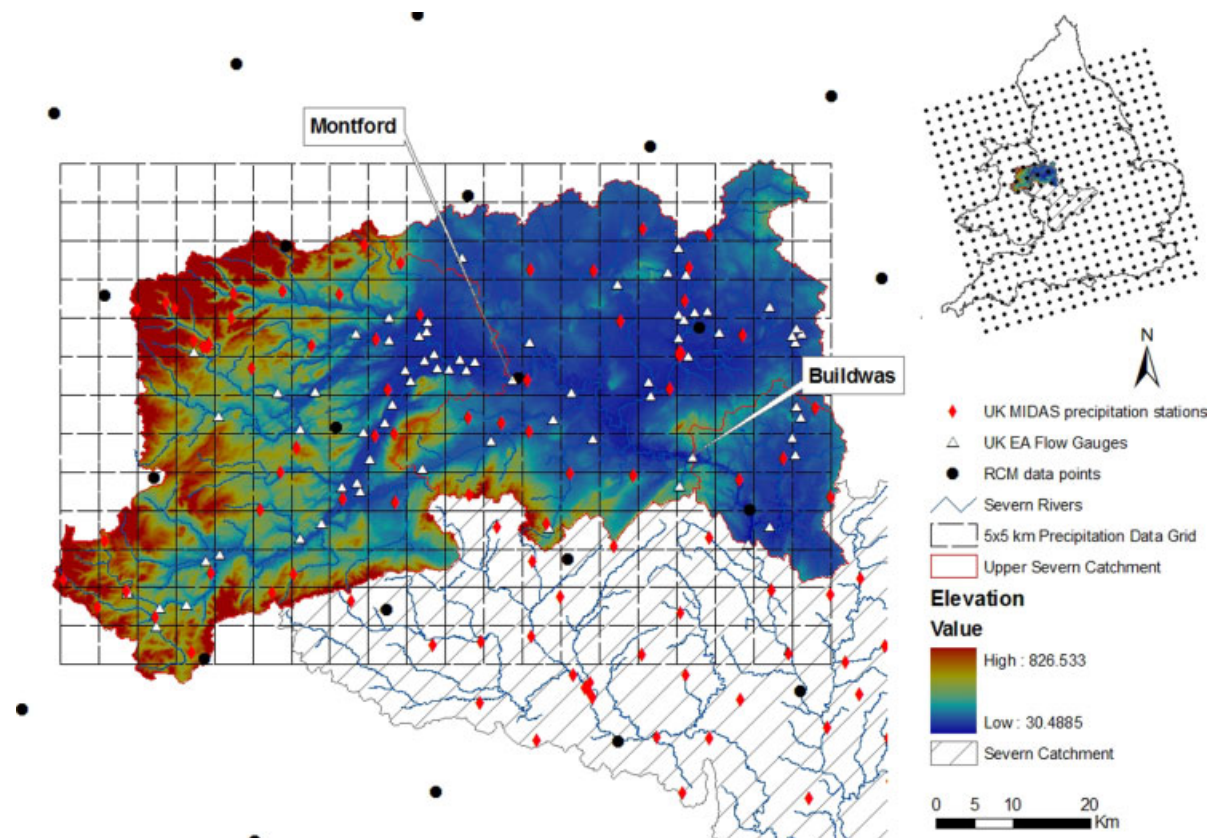


Figure 1. Upper Severn catchment located in the Midlands Region of England. Observed precipitation grid from UK Met Office at 5×5 km, RCM projection grid and Environment Agency river flow gauges are shown. This figure is available in colour online at wileyonlinelibrary.com/journal/qj

Table 1. Precipitation characteristics for observed and control period RCM precipitation for the Upper Severn catchment upstream of Montford and Buildwas. Uncorrected and MOS corrected values for the RCM precipitation are shown. Values show the mean and the standard deviation of the spread from the models.

		Upstream Montford	Upstream Buildwas	
Mean daily precipitation	Observed	3.62	2.84	
	RCMs UKCP09	3.12 ± 0.26	2.55 ± 0.21	
	RCMs UKCP09 MOS	3.63 ± 0.3	2.84 ± 0.02	
	RCMs ENSEMBLES	3.56 ± 1.0	3.05 ± 0.68	
	RCMs ENSEMBLES MOS	3.79 ± 0.17	2.86 ± 0.01	
	Probability of wet day	Observed	0.66	0.63
Probability of wet day	RCMs UKCP09	0.79 ± 0.05	0.77 ± 0.04	
	RCMs UKCP09 MOS	$0.67 \pm < 0.01$	$0.64 \pm < 0.01$	
	RCMs ENSEMBLES	0.93 ± 0.04	0.92 ± 0.04	
	RCMs ENSEMBLES MOS	$0.66 \pm < 0.01$	$0.63 \pm < 0.01$	
	Precipitation cutoff for MOS correction	RCMs UKCP09	0.19 ± 0.08	0.18 ± 0.07
		RCMs ENSEMBLES	0.59 ± 0.28	0.59 ± 0.28
Mean annual 5-day maximum (5dmax)	Observed	87 ± 17	67 ± 14	
	RCMs UKCP09	77.6 ± 6.4	62.4 ± 5.0	
	RCMs UKCP09 MOS	92.0 ± 7.2	71.3 ± 5.6	
	RCMs ENSEMBLES	76 ± 19	68 ± 11	
	RCMs ENSEMBLES MOS	92.0 ± 3.1	74.6 ± 2.3	

2.2. Hydrological modelling and parameter uncertainty

The hydrological model used in this study was HBV (Bergström, 1992; Lindström *et al.*, 1997), which is a conceptual rainfall runoff model that is widely used for flood forecasting and climate impact assessment both in operations and research (e.g. Lidén and Harlin, 2000; Olsson and Lindström, 2008; van Pelt *et al.*, 2009; Arheimer *et al.* 2011). There are many versions of the HBV model, and the one implemented in this paper is based on HBV light

(Seibert, 2003). Meteorological variables that are required input are precipitation, 2 m temperature and potential evapotranspiration. Evapotranspiration was calculated with the McGuinness model, which requires only the mean daily temperature from the RCM output (McGuinness and Bordne, 1972) modified by Oudin *et al.* (2005):

$$PE = \frac{R_e T_a + X2}{\lambda \rho X1}, \text{ if } T_a + X2 > 0,$$

$$PE = 0 \text{ otherwise,} \quad (1)$$

Table 2. Bounds of the parameter values for the HBV model.

Parameter	Short name	Min.	Max.
Max. soil moisture content	FC	25	600
Limit for potential evapotranspiration	LP	0.001	1
Soil routine parameter	BETA	0.1	10
Percolation from the upper to lower box	PERC	0.01	500
Upper zone limit	UZL	0	1000
Recession coefficient from upper zone	k0	0	1
Recession coefficient from zone 1	k1	0	1
Recession coefficient from zone 2	k2	0	1
Transformation of runoff	MAXBAS	1	10
Evapotranspiration constant	X1	40	400
Evapotranspiration constant	X2	0	30

where R_e is extraterrestrial radiation (depending only on latitude and ordinal date), T_a is the mean daily temperature ($^{\circ}\text{C}$), λ is the latent heat flux, ρ is density of water and K_1 and K_2 are constants ($^{\circ}\text{C}$) that can be calibrated. This simple formulation of evapotranspiration has been found to be robust when applied in climate impact studies (Oudin *et al.*, 2005).

The HBV model (as all other hydrological models) is best used together with an uncertainty analysis framework (Seibert, 1997; Pappenberger and Beven, 2006). The model has a number of free parameters (Table 2). HBV treats snow melt with a degree-day factor combined with a threshold temperature. These parameters were not varied since their influence on the high flows in the test catchment was very small. The HBV model was calibrated over the period 1986–2006, covering the period when discharge observations were available. The initial parameter space was estimated by Monte Carlo simulation and visibly detecting upper and lower limits for the parameters from the resultant simulations. If no limit was detectable, the limits were set according to physical constraints or from literature. The final estimation of the optimal parameter space of the free parameters was tested using a base sample set generated through a quasi-random Lp Tau method, which is an efficient method to generate a quasi-random sequence for Monte Carlo experiments (Sobol, 1979, 1993; see implementation in Cloke *et al.*, 2008). The base sample was 200 000 simulations, which provides an adequate exploration of the parameter space while considering computational constraints. The top 100 behavioural parameter sets were selected for use in the response surface and RCM direct climate impact projections and these all achieved Nash–Sutcliffe efficiency (NSE) values over 0.895 (Nash and Sutcliffe, 1970) during the calibration period.

2.3. Climate model projections

2.3.1. RCM grand ensemble projections

In this paper two ensemble sets of RCMs have been considered: a multi-model RCM set provided by the

European Union's FP6 ENSEMBLES project (van der Linden and Mitchell, 2009) and the most recent UK Climate Projections (UKCP09; Murphy *et al.*, 2009), which provide an ensemble of HADRM3 RCM projections using the same model with different climate sensitivities achieved by varying uncertain parameters within the model formulation. In both cases scenario A1B was used and GCMs were used to force the RCMs at a 25×25 km resolution (Table 3). ENSEMBLES has 19 ensemble members and UKCP09 has 11 ensemble members run with boundary conditions from global GCMs. Since the RCM data were on a coarser grid than the observational grid, they were first interpolated to the same grid as the observational data using linear interpolation for temperature and nearest-neighbour interpolation for precipitation (in order to preserve the amount of precipitation falling over the catchment). We consider three periods: *observed* (1961–2000 – forced with ERA40 reanalysis; Uppala *et al.*, 2005); *control* (1961–2000 – forced with GCM control period); and *future* (2001–2100 – forced with GCM future projections).

2.3.2. Model output statistics

The MOS used in this study was a double-gamma distribution error correction (DBS; Yang *et al.*, 2010). The technique uses three steps: (i) a precipitation cut-off (<1 mm) for days with small amounts of precipitation, which changes the frequency of wet days (see Table 1); (ii) an estimation of the distributions of the observed and modelled precipitation on the remaining days; and (iii) a shift of the modelled precipitation using the distributions. The estimation of the cut-off frequency and the distribution parameters was undertaken for the period 1961–2000, and then applied to the model output of the scenario runs. The method has proven useful in hydrological impact studies, especially regarding the upper end of the precipitation distribution (Yang *et al.*, 2010):

$$\begin{cases} \hat{P} = F_{\text{obs}}^{-1}(\alpha_{\text{obs}}, \beta_{\text{obs}}, F_{\text{sim}}(P, \alpha_{\text{CTL}}, \beta_{\text{CTL}})) & \text{if } P < 95\text{th percentile value} \\ \hat{P} = F_{\text{obs}}^{-1}(\alpha_{\text{obs},95}, \beta_{\text{obs},95}, F_{\text{sim}}(P, \alpha_{\text{CTL},95}, \beta_{\text{CTL},95})) & \text{if } P \geq 95\text{th percentile value,} \end{cases} \quad (2)$$

where P is the simulated precipitation series, F_{obs}^{-1} denotes the inverse of the cumulative gamma distribution function, F_{sim} is the cumulative gamma distribution for the observations, and α and β are the gamma distribution parameters estimated over the control period 1961–2000 for observed values and simulated values respectively. The method uses two distributions to represent the bulk of the distribution (below 95th percentile) and the more intense precipitation events (above 95th percentile). The reason for this split is to better condition the top end of the distribution. The precipitation series were first subject to a reduction of the number of rainy days by using a cut-off threshold of precipitation to have the same number of rainy days in the control period for both the simulated and observed precipitation. The MOS was applied to the interpolated RCM grid on the 5×5 km resolution over the entire catchment, which means that local effects, such as orography, were implicitly accounted for. Here seasonal based corrections were not applied, as they would have been inappropriate

Table 3. RCM climate projections used in this study.

Grand ensemble index	Years simulated	GCM	RCM	Reference
UKCP09 1-11	1950–2099	HADCM3	HADRM3	Collins <i>et al.</i> (2001); Pope <i>et al.</i> (2000)
ENSEMBLES 1	1951–2050	ECHAM5	C4I-RCA3	Jones <i>et al.</i> (2004); Kjällström <i>et al.</i> (2005)
ENSEMBLES 2	1951–2098	HadCM3	C4I-RCA3	
ENSEMBLES 3	1951–2050	ARPEGE	CNRM-RM4.5	Somot <i>et al.</i> (2004); Radu <i>et al.</i> (2008)
ENSEMBLES 4	1951–2100	ARPEGE	DMI-HIRHAM5	Christensen <i>et al.</i> (2007)
ENSEMBLES 5	1951–2099	ECHAM5	DMI-HIRHAM5	
ENSEMBLES 6	1951–2098	HadCM3	ETHZ-CLM	Böhm <i>et al.</i> (2006)
ENSEMBLES 7	1951–2099	ECHAM5	ICTP-REGCM3	Giorgi and Mearns (1999)
ENSEMBLES 8	1951–2100	ECHAM5-r3	KNMI-RACMO2	Lenderink <i>et al.</i> (2003); van Meijgaard <i>et al.</i> (2008)
ENSEMBLES 9	1951–2050	BCM	METNO-HIRHAM	Haugen and Haakenstad (2006)
ENSEMBLES 10	1951–2100	ECHAM5	MPI-REMO	Jacob (2001)
ENSEMBLES 11	1951–2050	CGCM3	Ouranos-CGCM	Plummer <i>et al.</i> (2006)
ENSEMBLES 12	1961–2099	BCM	SMHI-RCA	Jones <i>et al.</i> (2004); Kjällström <i>et al.</i> (2005)
ENSEMBLES 13	1951–2100	ECHAM5	SMHI-RCA	
ENSEMBLES 14	1951–2100	HadCM3	SMHI-RCA	
ENSEMBLES 15	1951–2050	HadCM3Q0	UCLM_PROMES	Sanches <i>et al.</i> (2004)
ENSEMBLES 16	1951–2050	HadCM3Q0	VMGO-RRCM	Shkolnik <i>et al.</i> (2000)
ENSEMBLES 17	1951–2050	HadCM3	METNO-HIRHAM	Haugen and Haakenstad (2006)
ENSEMBLES 18	1951–2100	ARPEGE	CNRM-RM5	Somot <i>et al.</i> (2004); Radu <i>et al.</i> (2008)
ENSEMBLES 19	1951–2099	ECHAM5	DMI-HIRHAM5	Christensen <i>et al.</i> (2007)

because of lack of data (the upper 5 percentiles are already being considered).

2.3.3. Continuous delta response surface generation

A continuous delta response surface approach involves applying impact models to probabilistic outputs from climate models by identifying thresholds through sensitivity analysis and then constructing response surfaces (Jones, 2001; Wetterhall *et al.*, 2011). Scenario climate projection outputs are then superimposed on to the constructed response surface. The technique is very useful for visualizing two or three variables together, which in the case of hydrological studies mostly are temperature and precipitation variables. The caveat of the method is that any changes in the distribution of the variables of interest are not taken into account – merely a change in the mean. Thus there is an assumption that climate change only produces an overall scaling of precipitation and no change in the frequency and no change in the shape or spatial variation of the climatology. However, information on seasonal signals can be used.

Response surfaces were created by perturbing the temperature and precipitation observations that were used as input to the HBV hydrological model over the calibration period 1986–2006. The perturbation for temperature was constructed as an additive factor, varying from -1 to 8°C with an increment of 0.1°C . The perturbation for precipitation was constructed as multiplicative factor, ranging from 0.7 to 1.6 with an increment of 0.01 (Figure 2(a)). The ranges were selected to ensure coverage of the range of expected changes to precipitation and temperature. The perturbed series were constructed as annual mean increases of precipitation and temperature, but since the change in precipitation (and temperature) is not uniform over the year a response surface with a seasonal difference was also constructed. The seasonal differences were calculated as the mean seasonal differences from the UKCP09 and GCM perturbed physics runs, which predict an increase in winter precipitation and a decrease in summer precipitation (Figure 2(a)).

An evaluation of the modelled river runoff in a response surface can be used to evaluate how water resources are affected in a particular region. However, for flood-related studies high flows are of more interest. Therefore, the response surface was created by calculating the annual frequency of the maximum river runoff exceeding the Montford flood warning level of $331\text{ m}^3\text{ s}^{-1}$ (Figure 2(b)). The contour lines in Figure 2(b) correspond to the probability of the warning level being overtopped in any given year. Previously this technique has been used to assess both flooding and low water levels (Wetterhall *et al.*, 2011), but the focus in this paper was on flooding. Hydrological model parameter uncertainty was taken into consideration in the response surface by using several behavioural parameter sets (discussed in section 2.2). The response surface shows that an increase in precipitation increases the probability of floods in the future, which is expected. However, the annual and seasonal threshold exceedance probabilities change nonlinearly with precipitation increase. This underlines the importance of translating projected precipitation increases into flood flows with an understanding of hydrological processes and the use of a hydrological model, rather than just by extrapolating from precipitation behaviour. The decrease in probability because of an increase in temperature (and therefore potential increase in evapotranspiration) has much less effect on the flooding than the precipitation change. This would no doubt have a large effect on low flows in the summer, an effect which we recommend for pursuit in future research.

2.3.4. Perturbed physics ensemble (PPE) for use with response surfaces

The UK Met Office has constructed a joint probability distribution (JPD) of future changes in temperature and precipitation through an extensive experiment using GCMs, downscaling and statistical emulators (Harris *et al.*, 2010). These data were used with response surfaces similar to the method evaluated in Wetterhall *et al.* (2011). Based on the SRES-A1B scenario, the JPD combines transient HadCM3 simulation output with a combined perturbed physics and

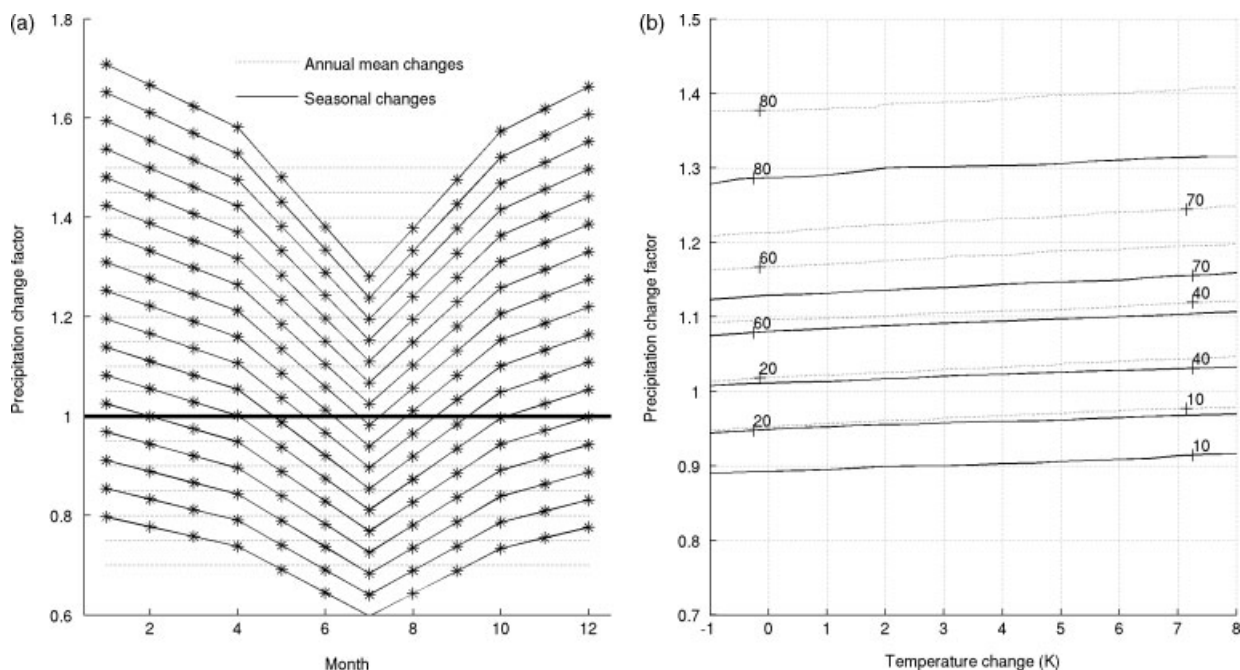


Figure 2. (a) Multiplicative factors used to calculate the response surfaces with the continuous delta change method. The dotted lines show the change factors with a constant annual change, and asterisks denote the perturbations with a seasonal change, estimated from the expected changes in the perturbed physics experiment. Increments for the annual perturbations in the figure are shown for each 0.05 level. (b) Response surface for the Environment Agency warning level ($331 \text{ m}^3 \text{ s}^{-1}$) for Montford river gauging station. The contour plot indicates the probability that the selected threshold is exceeded in a given year and takes into account parameter uncertainty in the HBV model.

emulator approach using equilibrium climate simulations. Observations were used to constrain the JPD as well as weightings based on the Coupled Model Intercomparison Project (CMIP3; IPCC, 2007) and HADRM3 European simulations. From the distributions 10 000 paired samples of temperature and precipitation were drawn, for grid boxes with areas of approximately $300 \times 300 \text{ km}$, which were provided as seasonal and annual changes over 20-year time slices over the 21st century. Details on the methodology can be found in Harris *et al.* (2010). On the response surfaces the JPD is displayed as contour plots where the outer limits were the 5 and 95 percentiles of the total runs.

3. Assessment of observed, control and future precipitation projections for the Upper Severn

Our assessment is formed of two parts. In this section the ERA-40 (Uppala *et al.*, 2005) forced RCM projections over 1961–2000 were assessed to test how well the models could reproduce observed precipitation events; then the future projections of precipitation are considered, followed by an analysis of the effects of the MOS. In section 4 the analysis is extended to modelled river flooding.

3.1. Observed: ERA40 forced RCM output

Flooding in the Severn catchment usually occurs during intense precipitation events from late summer to winter. Figure 3 shows that precipitation is well modelled by the RCMs during the dry season (February–July), but for the wet season (August–January) the precipitation is generally underestimated. Figure 3 also shows that the intra-annual variation is also underestimated by the RCMs, and this bias needs to be addressed in flood studies.

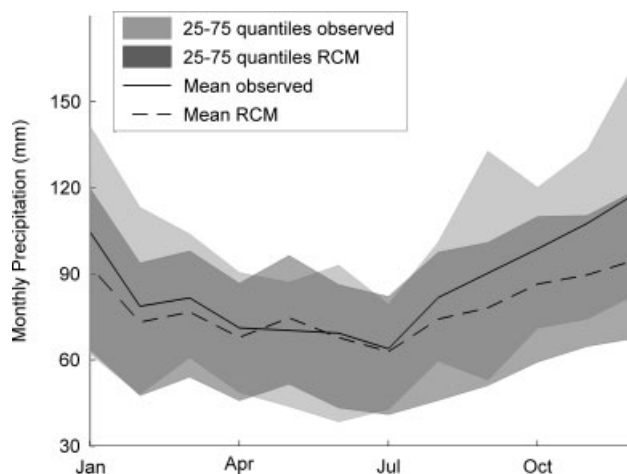


Figure 3. Annual variation of precipitation from the ERA40 forced RCM output used in the study. The dark-grey area shows the 25–75 percentiles of the RCMs and the light-grey area the 25–75 percentiles of observed precipitation. The solid line denotes the observed precipitation and the dotted black line the mean of the RCMs.

3.2. Control and future: GCM forced RCM output

Table 1 shows statistics comparing control period RCM precipitation and observed. Notably, the performance of the RCM precipitation is dependent on the size and geographical characteristics of the catchment. For the smaller Upper Severn (upstream of Montford), which has a large elevation gradient, both the mean daily precipitation and the mean annual 5-day maximum (5dmax) are underestimated for UKCP09 and ENSEMBLES and the number of rainy days is overestimated. However, for the larger Buildwas catchment with more lowland, although the number of rainy days is still overestimated and the UKCP09 is still underestimating

the precipitation, for the ENSEMBLES there is now an overestimation.

Figures 4 and 5 show precipitation cumulative distribution functions for the UKCP09 and ENSEMBLES RCMs. Figure 4 shows the results for the annual precipitation. The left panels show the control period and the right the 2071–2100 future projections. A comparison of the observed and uncorrected RCM precipitation demonstrates the offset between observed and RCM annual precipitation for the Severn. A comparison of panels (a) and (c) shows that the ENSEMBLES control period results show a significant change in the distribution shape, whereas the UKCP09 is more uniformly underestimated. This is related to the fact that UKCP09 is created by a single model with perturbations, whereas ENSEMBLES are multi-model results and thus likely to represent more variable model behaviours. The ENSEMBLES here shows a wet bias for larger precipitation values. The contribution of the spread in the precipitation from the GCM/RCM is mainly from the GCM. Figure 5 shows the 5-day maximum precipitation (5dmax) measure, which is an important indicator of flood-producing rainfall in the Severn catchment. A comparison of Figures 4 and 5 for both control and future periods demonstrates that the offset and under/overestimation of precipitation changes depending on the way the precipitation is assessed. Here both UKCP09 and ENSEMBLES show an underestimation and a change in distribution shape, widening at the lower end, although the ENSEMBLES is still further away from observed than UKCP09. The wet bias detected in the annual precipitation higher values for ENSEMBLES is no longer present in the 5dmax (Figure 5(c)), which is important when using these results for climate impact studies in flooding.

3.3. Model Output Statistics (MOS)

Here the DBS MOS technique is applied to the grand ensemble of RCM projections as described in section 2.3.2. Each individual GCM/RCM pair in the grand ensemble was corrected for the bias, using the period 1960–2000 as baseline for the observed climate. MOS was only applied to the precipitation output as floods are mainly driven by precipitation events. There is also bias in temperature (not shown), and in catchments where the floods are snow-melt driven, or in studies of low-flows, a MOS correction of temperature would be necessary (Yang *et al.*, 2010). Figures 4 and 5 show that the results after MOS correction are much closer to the observed precipitation. However, the higher ends of the distributions are not perfectly corrected, since many values exceed the observed values. This is a common problem with the application of MOS techniques and such corrections have to be done with care since it is a rather crude method to correct for RCM biases (see discussion in Maraun *et al.*, 2010).

It is also important that MOS does not affect the climate change signal, which can be seen by comparing corrected and uncorrected results in panels (b) and (d) of Figures 4 and 5 and also in the monthly changes shown in Figure 6. The change in precipitation before and after MOS is very similar. The stationarity of the MOS is not considered here and the assumption that the statistical relationships do not change in the future should not be considered insignificant. It will increase the uncertainty

over time, since the precipitation mechanisms, such as large-scale circulations, increased humidity, blocking frequency (blocking over mainland Europe can force Atlantic low pressures to take a certain path over the UK that can cause severe flooding), etc. might change (Maraun *et al.*, 2010). In general, future precipitation over the Upper Severn shown in Figure 6 demonstrates an increase in winter and a decrease in summer. There is quite a stark difference between the UKCP09 and ENSEMBLES results, which is much larger than the effects of MOS, although these are not insignificant, especially in the ENSEMBLES results for winter precipitation. In the last time slice, winter precipitation is higher for ENSEMBLES and summer precipitation is lower for UKCP09. Contrast these results with Figure 5, where 5-day maximum totals show ENSEMBLES to have the driest bias.

4. Assessment of flood projections for Upper Severn

The HBV model was forced with ERA-40-RCM projections of ENSEMBLES and UKCP09 and then the grand-ensemble GCM/RCM projections. The runs with the ERA40-forced simulations are very useful in assessing the ability to model current climate and the effects of MOS on the simulated runoff. Note that the GCM/RCMs used in the ERA40-forced experiments are not exactly the same as in the future GCM/RCM simulations since the ENSEMBLES project did not provide the full matrix for both experiments, but the general performance of the RCMs can be assessed.

4.1. Observed: ERA40-RCM-forced river discharge simulations

Figure 7 shows the HBV-modelled river discharge for the November–December 2000 flood event on the River Severn at Montford (including HBV model parameter uncertainty). HBV has been forced by uncorrected (Figure 7(a)) and MOS-corrected (Figure 7(b)) ERA40-forced RCM projections. In comparison with the observed discharge, the MOS-corrected discharges capture the shape and dynamics of the 2000 flood event, but the spread is very large. The RCM ensemble median is well below the peak of the event for the MOS-corrected RCMs and the observations do not fall within the RCM ensemble. Considering the Montford flood warning threshold of $331 \text{ m}^3 \text{ s}^{-1}$, for the November peak this level is just exceeded by the RCM median and thus half of ensemble members reach this level, but the RCM median does not reach this level for the December peak and only the tails of the RCM distribution exceed this threshold. As these projections are forced by reanalysis and come from individual models (in contrast to a well-calibrated short-range ensemble prediction system) it can be inferred that there is a systematic underprediction in the simulation of discharge peaks by the RCMs. This could be very significant in assessing future flooding, and it is worth noting that the precipitation has already had MOS correction applied and thus this should not be a major cause of the underprediction (although the MOS is not perfect and may contribute to this in a small way). One way to constrain the uncertainty band and decrease the projected spread in climate impact studies is to weight the RCMs that perform better during the control period higher than the models that underperform. In Figure 7(c), the RCMs were weighted based on their NSE skill over the period 1986–2000. The reduction in the

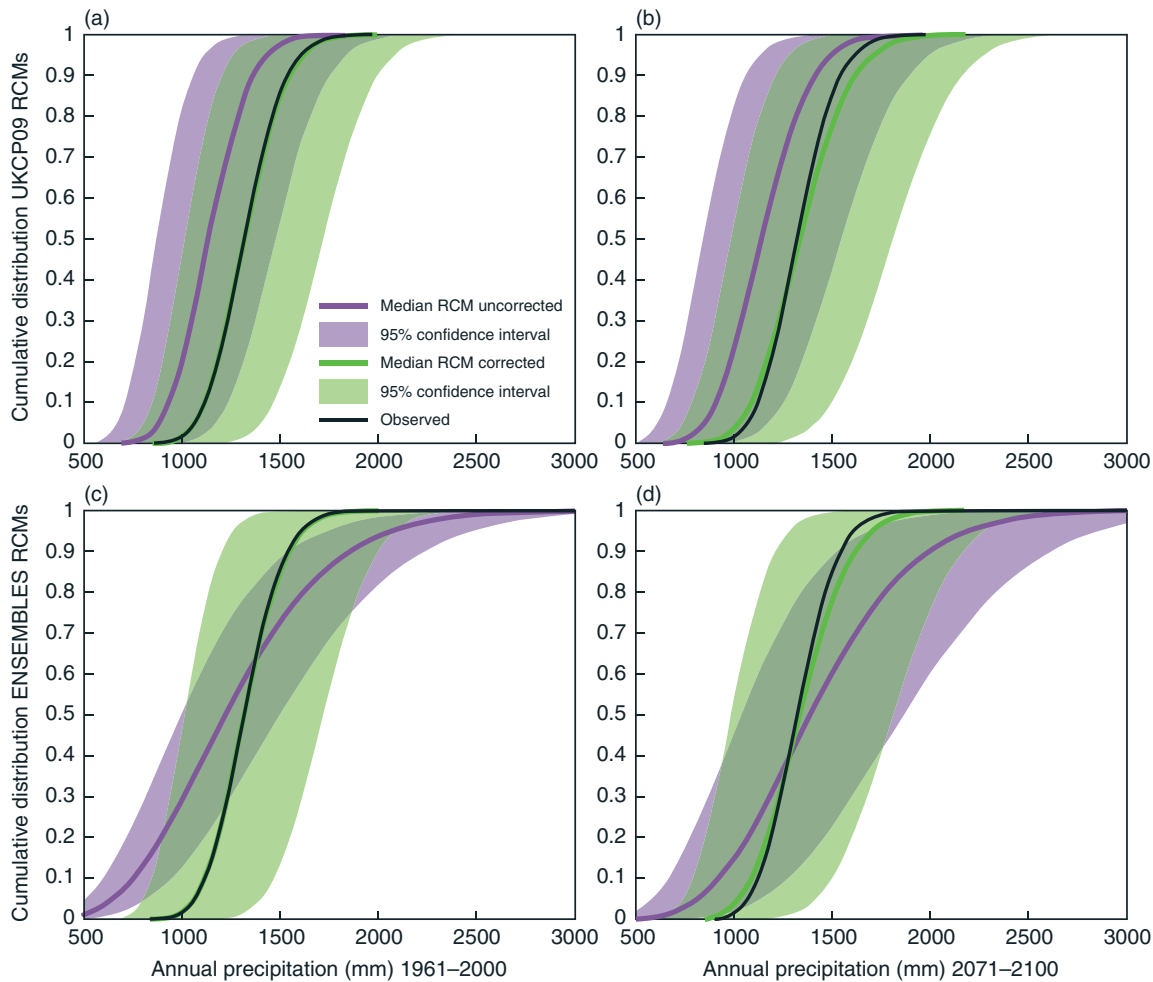


Figure 4. Annual precipitation cumulative distribution functions from the UKCP09 (upper panels) and ENSEMBLES (lower panels) RCMs. The left panels show the control period (1961–2000) and the right panels show the future projections (2071–2100). The black line shows the observed precipitation (also shown on the right panels for reference). The dark grey line (purple online) shows the uncorrected median and dark grey shading (purple online) shows the 95% confidence interval for the uncorrected. The light grey line (green online) shows the MOS corrected median values and the light grey shading (green online) shows the 95% confidence intervals for the MOS corrected. This figure is available in colour online at wileyonlinelibrary.com/journal/qj

uncertainty bands is evident from the figure, and in the later peaks the observed and ensemble distributions overlap much more than that seen in Figure 7(b). Any reduction in uncertainty presented can be useful in providing an ‘expert judgement’ to the stake holder; however, it discredits the models that behave less well during the control period but that might be useful as extra information in climate impact studies. It is also a very application-dependent outcome (Kjellström *et al.*, 2010) and so we do not continue this weighting in our analysis of future projections, as a more thorough consideration of the implications is beyond the scope of this analysis.

4.2. Future flood discharge projections with direct approach

The HBV model was then forced by the future RCM projections (again including parameter uncertainty). The precipitation inputs were included as both uncorrected and as MOS corrected. The MOS correction was undertaken for each individual GCM/RCM combination. Figure 8 shows the ensemble mean of the annual maximum discharge from HBV simulations averaged across time slices for the UKCP09 and ENSEMBLES. The results are separated into those simulations forced by uncorrected and those

forced by MOS-corrected projections. In all cases the mean annual maximum discharge is projected to increase across the time slices by around $40 \text{ m}^3 \text{ s}^{-1}$ by the end of the century. Two things are immediately apparent from Figure 8. (i) The ENSEMBLES mean annual maximum discharges are much lower than that for UKCP09 HadRM3, which is very interesting as this is opposite to the pattern seen in Figure 4 for the mean annual precipitation. This again highlights that the annual precipitation is not a very useful indicator for flood events, and the 5-day maximum precipitation (shown in Figure 5) and the calculated discharge are more useful. This supports the importance of using a hydrological perspective and modelling approach in these types of studies. (ii) The MOS correction is a very large correction, with the MOS-corrected discharge ensembles of the order of $100 \text{ m}^3 \text{ s}^{-1}$ higher than the uncorrected. This is a very significant impact on the projected discharge in that all time slices are above the observed maximum value of the mean annual maximum discharge. The distribution of mean annual maximum discharges across the time slices (gradient across the time slices) is also slightly altered by the MOS correction (see in particular the variability at the high end). This application demonstrates quite effectively what is a typical extreme difference in using MOS-corrected RCM

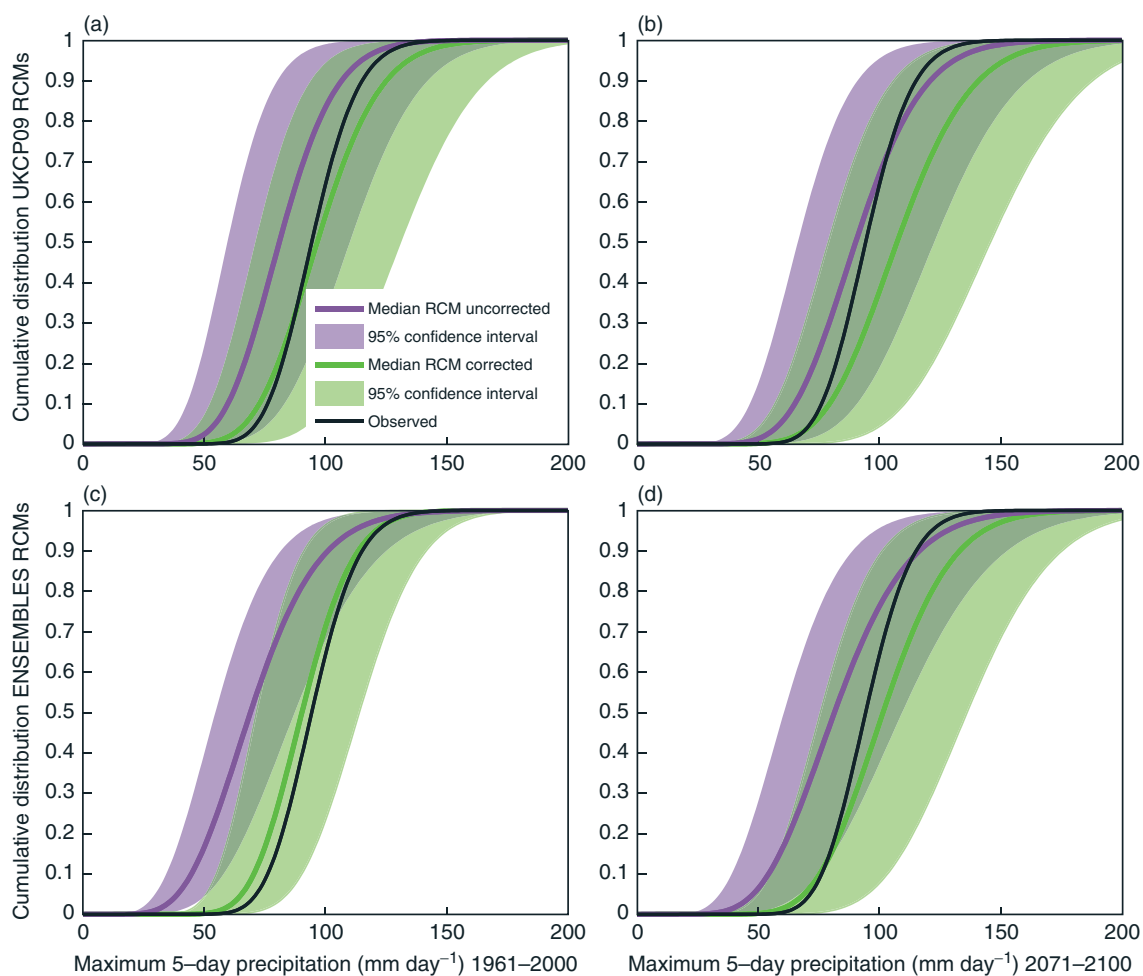


Figure 5. As Figure 4 but for 5-day maximum precipitation (mm day^{-1}). This figure is available in colour online at wileyonlinelibrary.com/journal/qj

precipitation for future discharge projections and illustrates the importance of understanding what the MOS techniques have done to the data.

4.3. Future projections with PPE JPD response surfaces

Figure 9 shows the annual changes in the probability that the flood warning level is exceeded at Montford. Output from the PPE JPD was overlain on the response surface (shown in Figure 2), along with the RCM projections. Even though the climate projections used in the grand ensemble and in the PPE JPD are different (both in terms of the model used and the fact that the PPE JPD is for a much larger area), their comparison is useful in that if results reinforce each other then it may support their robustness. The contours show the probability of exceeding the threshold and the colour plot shows the density of runs from the PPE. The annual changes suggest that they are roughly within the same range as the RCM projections, which could indicate that the GCM PPE is likely to be useful in determining the effects on flood warning levels. For example, for the PPE JPD the probability that the flood warning level is exceeded is by 2030–2050 already up to 40% in the centre of the distribution and extending towards 20% and 60%. The UKCP09 is for the same period, centring on 40% and extending from 30% to 50%. The ENSEMBLES are centred on 50% and extending from 30% to 60%.

However, looking at Figure 10, where the results for only the winter season are shown (DJF), there is a starker difference. The spread is larger and the differences between the three components (PPE JPD, ENSEMBLES and UKCP09) are more noticeable. Looking again at the example of the 2030–2050 time slice period, the PPE JPD has a higher ensemble mean than the other ensembles at around 50% and a higher spread extending from 15% to 75%. The ENSEMBLES centre on 40% and extend from 10% to 60%. The UKCP09 centres on 25% and extends from 10% to 40%. Thus, from these results, it is clear that (i) it is essential to consider the seasonal results rather than just the annual mean projections as they can be very different, and (ii) within the seasonal results, and also to a lesser extent within the annual results, the different ensembles considered (PPE JPD, ENSEMBLES and UKCP09) have quite different projections of future flooding.

4.4. Comparing ensemble approaches for determining future flood hazard change

Figure 11 shows a direct comparison of all the ensemble methods used in the previous sections for projected changes in the 2-year return period of flood warning level at Montford. The left-hand panel shows the ENSEMBLES, the middle panel the UKCP09 and the right-hand panel the combined results. RCMs have been used for direct runs with the HBV model as in section 4.2 (but have also been used

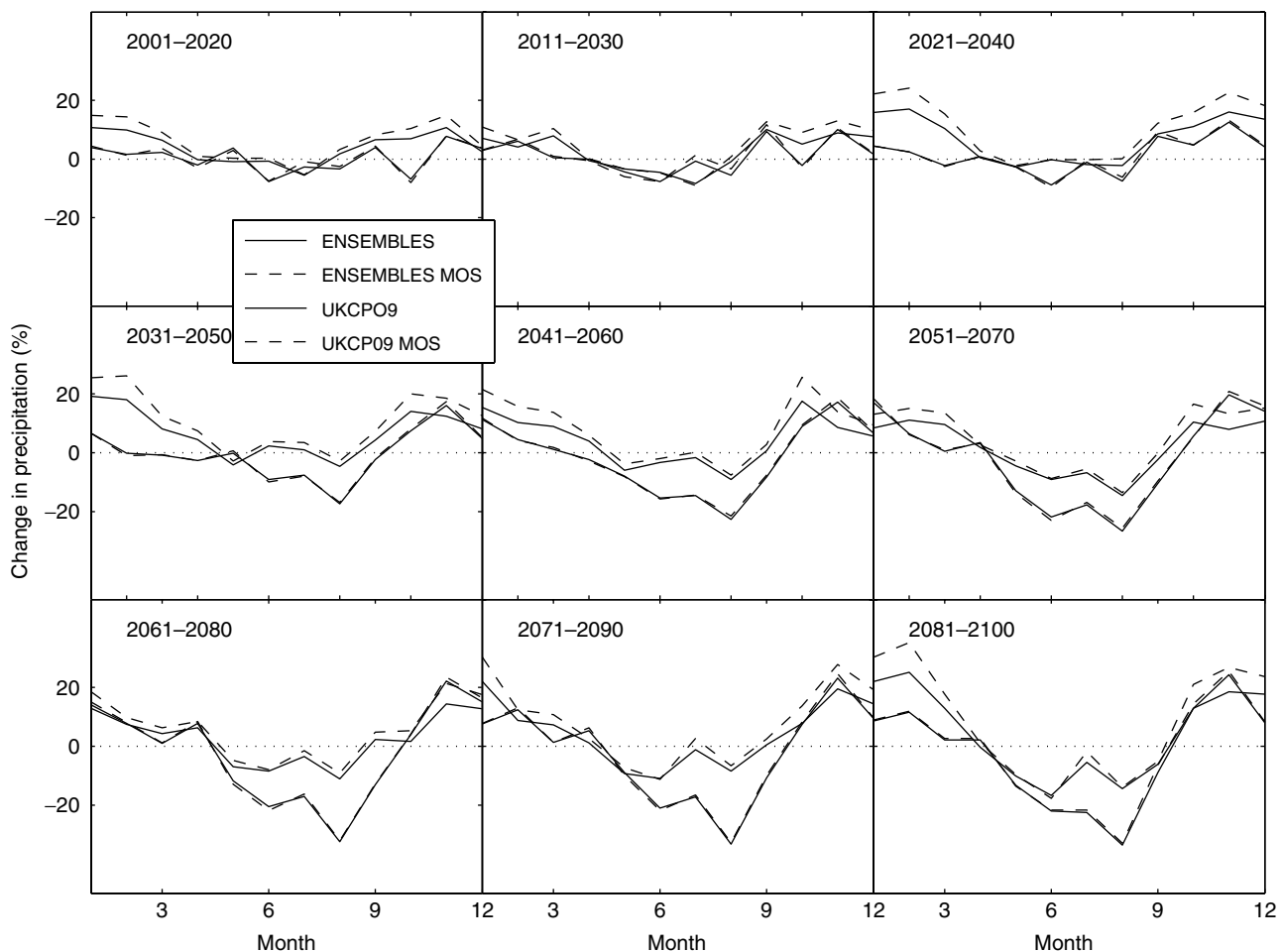


Figure 6. Modelled changes in monthly precipitation for 20-year time slices between 2001 and 2100 compared to the control period, 1961–2000. The lower lines are UKCP09 and the upper lines are ENSEMBLES. The solid lines are uncorrected projections and the dashed lines are MOS-corrected projections.

with the response surface technique. Both uncorrected and MOS-corrected runs have been shown for both direct and response surface results. The PPE JPD and the interquartile range is also shown for these results in order to bound the other results (the uncertainties in the RCM results are not shown here, for clarity, but have been explored in previous figures).

As would be expected from Figure 10, the largest change is noted for the PPE for the first half of the 21st century, whereas there is a sign of decreasing flood hazard towards the end of the century. The RCMs projected on the response surfaces (black lines in Figure 11) show a more moderate increase, and the ENSEMBLES RCMs (Figure 11(a)) differ from the UKCP09 RCMs (Figure 11(b)). The former indicates an increase towards the end of the century – something that is not the case in the UKCP09 runs. The UKCP09 follows a similar path to the perturbed physics with a decrease towards the end of the century; however, the RCMs in general indicate a much less pronounced change in comparison with today's climate. This result concurs with the direct runs using the RCMs as direct input to the hydrological model. Contrary to the response surface RCMs, the direct RCMs indicate that UKCP09 models give a higher estimate of the flood events in comparison to the direct runs with the ENSEMBLES RCMs. MOS increases the effect of the direct runs much more than for the response surfaces. The effects of MOS, where the high precipitation events are treated especially, are seen in

the modelling of high flows. Response surfaces deal with mean changes over an entire season, and the MOS has very little effect on these results. However, the two approaches do give very similar results in terms of the projected change in flooding over the catchment. There is a larger increase projected in the GCMs, but this result should be taken with great care since it is the projected change over a very large area.

5. Discussion

5.1. Ensemble techniques and uncertainty

Projections from ensemble modelling are useful for exploring uncertainty in climate models. These projections are being used in many impact modelling and decision-making studies, including those looking at climate impact on flood hazard. Here we have used projections from three different ensemble techniques in a flood hazard climate impact study: RCM projections from single perturbed model (UKCP09), RCM multi-model projections (ENSEMBLES) and compared these with GCM-perturbed physics projections (also using the RCMs) used with response surfaces (PPE JPD). The modelled control period and projections of future precipitation and flood hazard (and the uncertainties of these projections) certainly varied between methods (e.g. Figures 10 and 11). In the

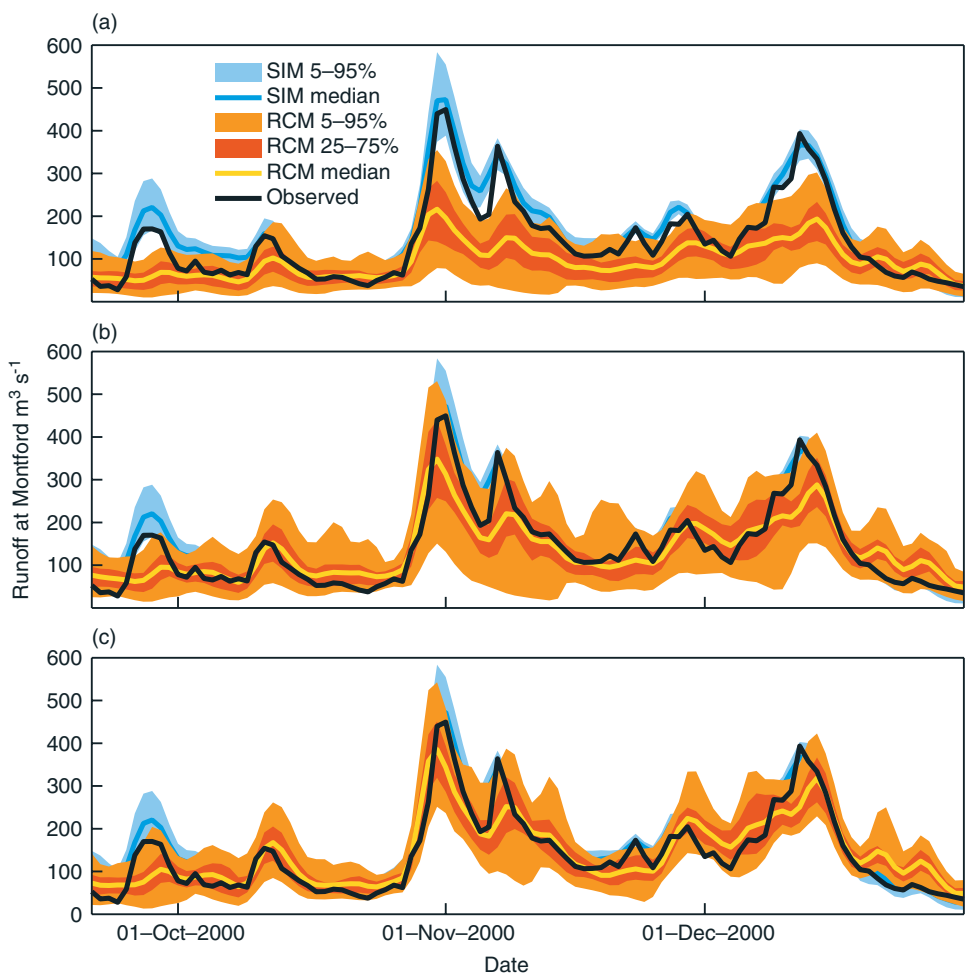


Figure 7. HBV flood discharge predictions at Montford for the November–December 2000 flood event using observed precipitation (black line) and ERA-40 forced RCM projections (RCM). HBV parameter uncertainty (SIM) for the observed data is also shown. (a) Uncorrected RCM output and observation uncertainty; (b) MOS-corrected RCM output; (c) MOS-corrected RCM output weighted according to performance over the control period. This figure is available in colour online at wileyonlinelibrary.com/journal/qj

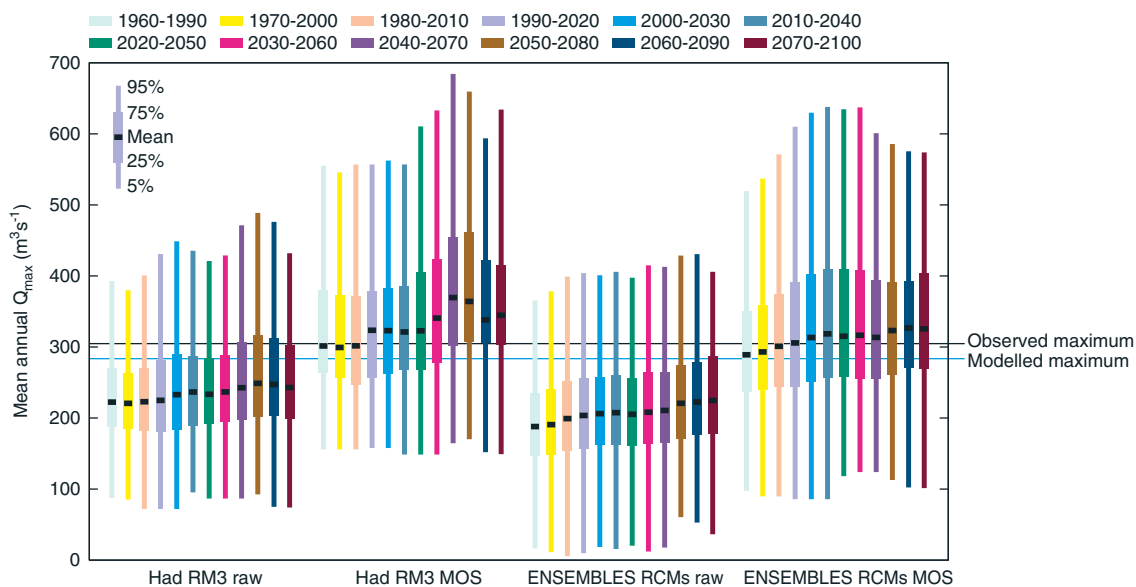


Figure 8. Ensemble mean of the annual maximum discharge from HBV simulations using uncorrected (raw) and MOS-corrected UKCP09 and ENSEMBLES projections. The lower horizontal line is the HBV results using the modelled maximum with observed precipitation, and the upper horizontal line is the observed annual maximum discharge (both averaged over 1986–2006). The coloured bars are annual maximum over 30-year time slices from 1991–2020, 2000–2030 etc. up until 2071–2100. This figure is available in colour online at wileyonlinelibrary.com/journal/qj

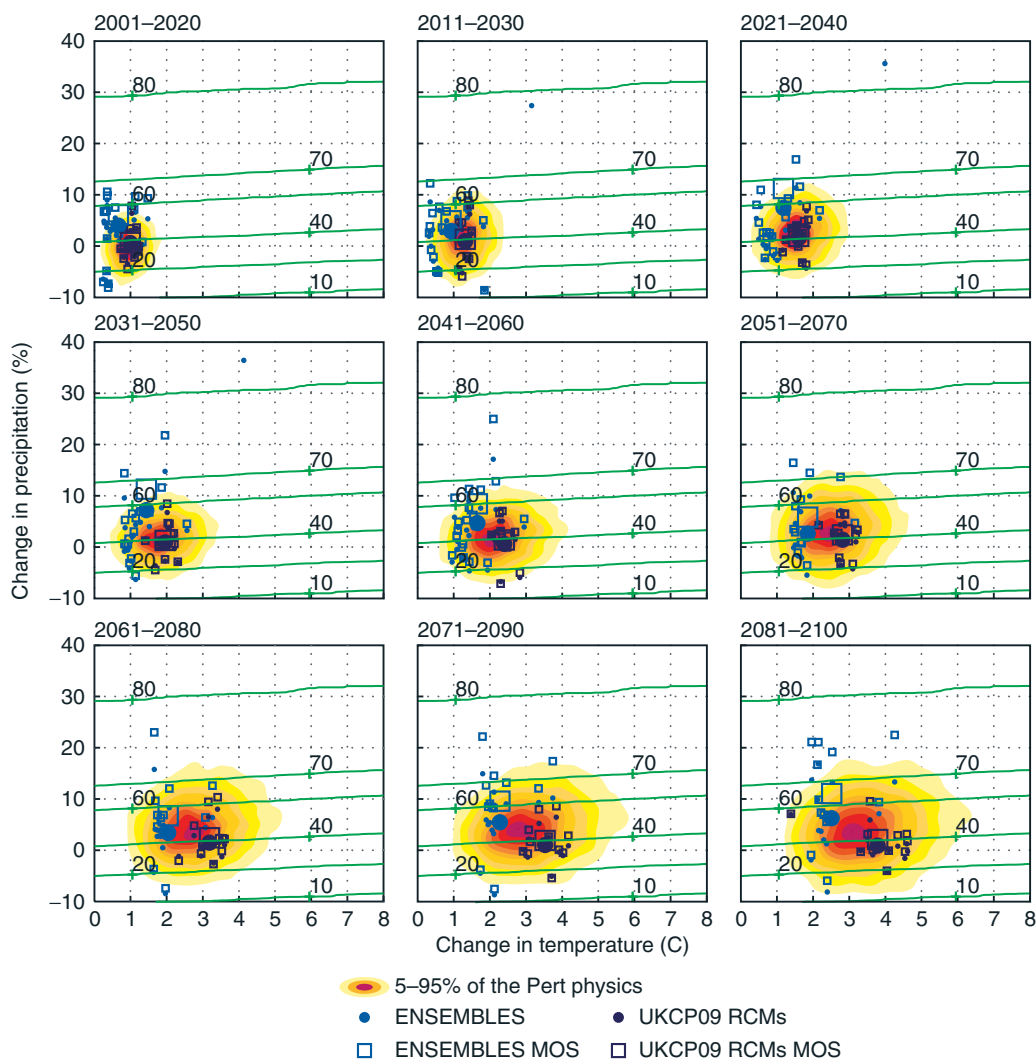


Figure 9. Perturbed physics experiment for the probability that the flood warning level is exceeded for the Montford catchment. The contours show the probability of exceeding the threshold and the shaded plot the density of runs from the perturbed physics experiment. The thicker dots denote the mean of the groups of RCMs. The squares indicate the RCMs after MOS. This figure is available in colour online at wileyonlinelibrary.com/journal/qj

control period assessment, there is an underestimation of precipitation in the Upper Severn catchment evident in the (uncorrected) UKCP09 results with a more varied bias in the ENSEMBLES (Figures 4 and 5). Overall, for this particular catchment, the combined results point towards an increase in flood hazard, but the results are not significantly different from the current situation. Certainly, the results highlight the very large (and well-known) uncertainties associated with projections of future climate whichever ensemble technique is used. In particular, the PPE shows a strong trend in the increase of flood hazard, but also a larger spread in the ensemble, and therefore a larger uncertainty. Similar results were found in Wetterhall *et al.* (2011) for response surfaces on a Swedish catchment. We highlight again that in the response surface method there is a crucial assumption that climate change only produces an overall scaling of the precipitation, whereas a hydrological model forced by GCM/RCM predictions can, at least in principle, respond to more detailed changes in climatology.

In terms of flood hazard, the difference in uncertainty between the techniques is likely to come from both the resolution (RCMs are downscaling the simulations to a domain which is in comparison with the catchment size) but also that the GCMs in the perturbed experiment sample

the model/scenario uncertainty to a much higher extent. The RCMs from the UKCP09 do sample some of the model uncertainty from the GCM (Figures 10 and 11), but it is a very limited sample. An obvious problem in studies of climate model uncertainty are the few numbers of realizations possible, due to computational constraints. However, it would be very useful for the climate impact community if a large grand ensemble of high-resolution RCMs could be created, consisting of a multi-model approach combined with perturbed physics for each model (multi-model each with many perturbed physics ensemble). This would enable a comprehensive understanding of the RCM uncertainty that is cascading into impact modelling. However, this will not completely solve the problems of the bias evident in the results or other issues surrounding the interpretation of the ensemble results for decision making.

5.2. To MOS or not to MOS

Our assessment of RCM projections has shown that there are deficiencies in the prediction of significant rainfall events over catchment scales which is in line with the findings of others (e.g. Rivington *et al.*, 2008; Themessl *et al.*, 2012) and, more importantly, that this degrades

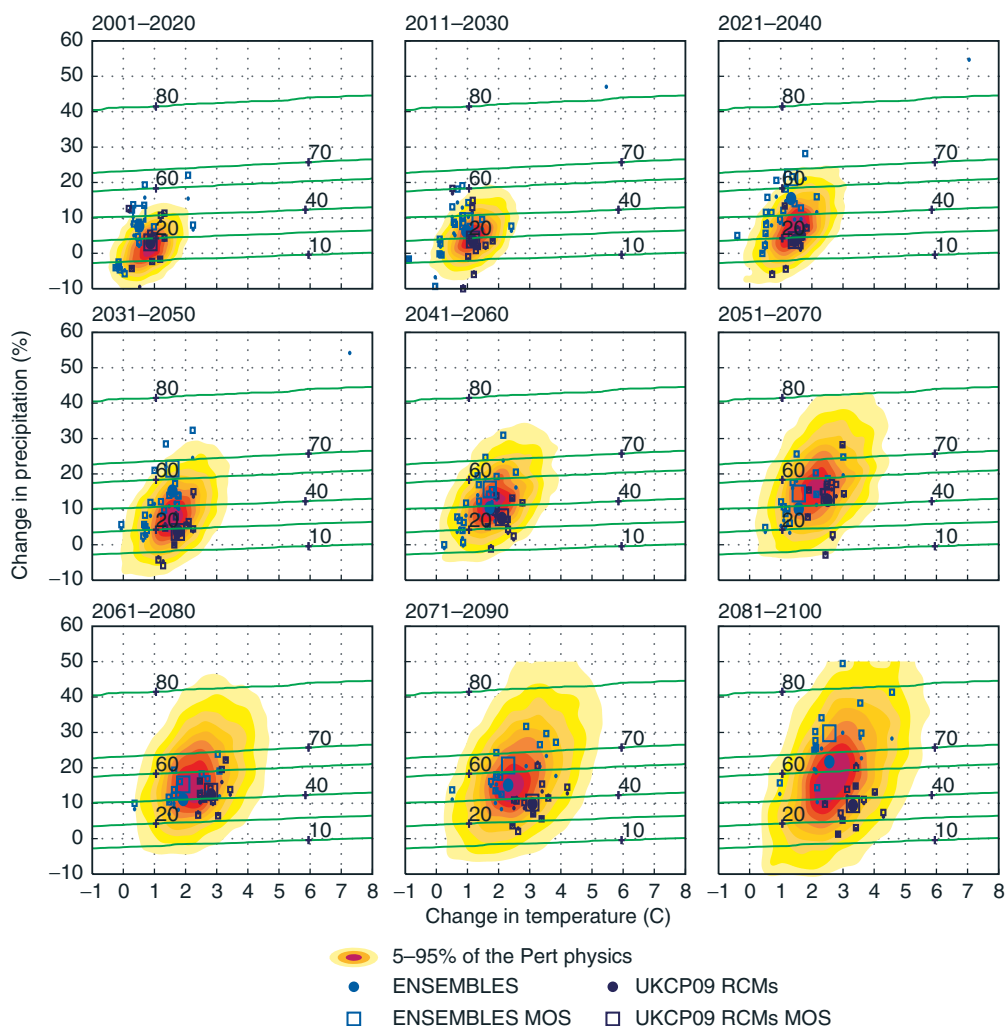


Figure 10. As Figure 9, but for the projected changes over the winter months (December–February). This figure is available in colour online at wileyonlinelibrary.com/journal/qj

further when quantifying extreme rainfall events that are the cause of floods – the interest of this paper. For example, in Figure 5 the large underprediction can be seen in the differences in the cumulative distribution functions of the 5-day maximum precipitation between the observed data and both the UKCP09 and ENSEMBLES RCM projections, and for example Figure 7(a) highlights the issues with this when these uncorrected projections are used to force a hydrological model for a flood event, and the river discharge is significantly underpredicted. Thus the solution to such bias is usually to advocate the application of MOS in order to adjust RCM projections, which is considered by many to allow confidence in the examination of future changes in flow regimes.

The influence of the application of MOS has been tested here. What is clear is that MOS can (to a certain extent) do what it is supposed to do, namely match distributions of observed and modelled precipitation, allowing simulated river discharges to be more compatible with those forced by observed rainfall products. This in turn then significantly alters future projections of precipitation and river discharge, supposedly to give a more accurate picture. Even in data-sparse regions there is nothing stopping such a methodology being applied – it would just mean that the observed would need to be treated more cautiously and the resolutions of the RCM’s may be lower. These factors would have the result

of increasing the final uncertainties further. Typically, the methodology would be applied only where data are available to constrain the hydrological model predictions, but in fact if the hydrological model parameters could be regionalized this could also be applied in ‘ungauged’ basins.

However, there are some credibility issues in undertaking such transformations and then relying on results for interpreting future climate. First, RCMs are dynamically downscaled physically based models of the climatological processes, albeit a simplification due to the resolving resolution. Thus transforming an output – in this case rainfall, arbitrarily without understanding what that means for the coupled fluxes and processes in the model – is effectively ‘throwing away the physics’ and therefore reasons why these models might provide feasible futures. Second, there is an assumption that such a transformation is stationary over multiple decades and hence holds as a reliable predictor of the bias in future scenarios. There are many reasons why this may not be a scientifically defensible position to take, in the same way that any nonlinear complex set of processes is unlikely to maintain a simplistic transformation of the output variables under changing conditions.

The more important question at this stage is what is the alternative to using MOS-corrected RCM projections and what is scientifically credible in order to express potential

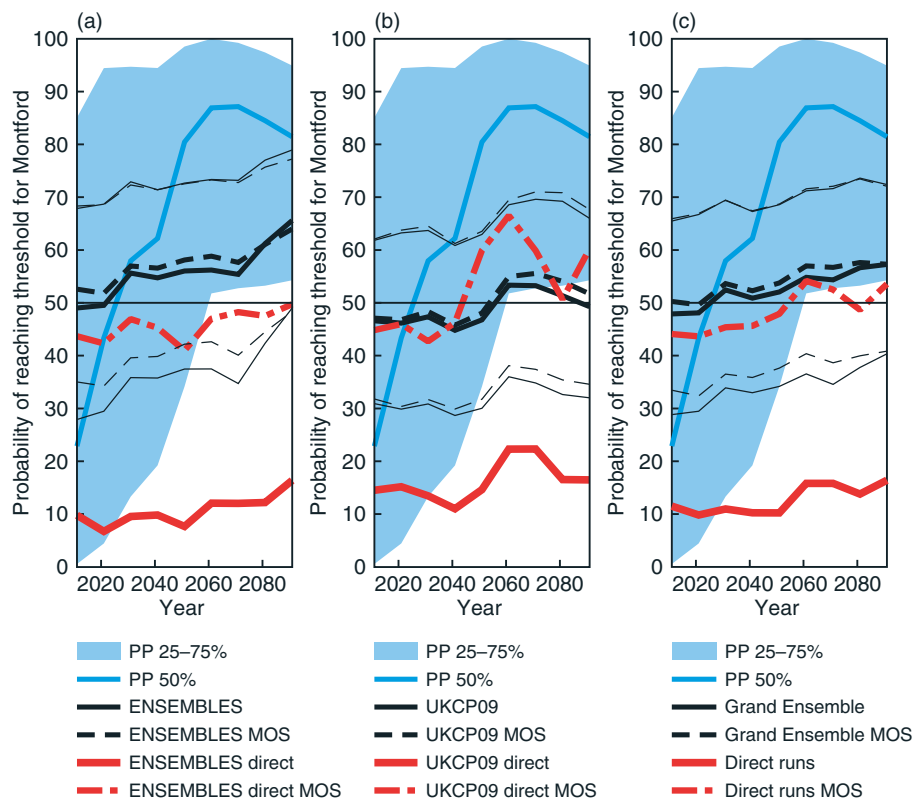


Figure 11. Changes in the 2-year return period of flood warning level at Montford (horizontal line). The left-hand plot shows the ENSEMBLES results; the middle plot shows the UKCP09 results; the right-hand plot shows the combined ENSEMBLES-UKCP09 grand ensemble. The median and 25–75% percentiles of the perturbed physics experiment used with the response surface technique are also shown. The black lines show the results for the RCMs used with the response surface technique and the dark grey lines (red online) show the direct simulations using the RCMs. The solid lines are uncorrected and the dashed lines are MOS corrected results. This figure is available in colour online at wileyonlinelibrary.com/journal/qj

changes in flood hazard at the catchment scale in the future. Certainly, our results have shown that by taking a grand ensemble of potential changes the future becomes very uncertain and that this uncertainty spread of annual maximum discharge increases with the use of MOS. But if we show that our current RCMs do not adequately characterize the most important variable for quantifying flooding impacts then why would we use them? Assuredly, there is substantially more work to be done here before they are used to define the limits of future changes. We could adopt different strategies and not try to defend the indefensible by using what could be seen as bias-corrupted RCM projections. We could use RCMs with the response surface method we have shown here, which may be a good compromise between maintaining important downscaled processes and features that are not evident in the GCM but also using information that is not affected by MOS while still maintaining links to observations (e.g. the change shown in the median lines in Figure 11). However, the assumptions of stationarity in the distributions and only changes in the mean are quite important. A close consideration of the results we have presented shows that with the two ways in which we have used the RCMs the flood hazard signal predicted is not completely different, which could be indicative of either a more robust signal from using the two methods together or of course alternatively they could both be equally wrong.

In the longer term, climate mitigation policies might be better evidenced by studies considering narratives of feasible futures (Beven, 2011), or ‘catchment change scenarios’, which combine scenarios of land use, rainfall and many other more ‘human factors’ (such as flood defence, agricultural

and development practice, financial stability and population movements) with a focus on risk-based outcomes (such as the changing number of people vulnerable to flooding). There is also an opportunity to take the pragmatic route and maintain concepts of freeboards from our current observed continuous simulations of discharge (e.g. adding 20% on to current observed discharge time series), at least for comparative purposes.

However, for those current and future studies where use of climate projections is for some reason seen as mandatory, and use of MOS is considered the way forward, we strongly advocate: (i) also running uncorrected projections so that the impact of MOS on results is clear; (ii) considering catchment-based indicators of precipitation, such as 5-day maximum precipitation (or whatever suits the catchment and discharge levels you are studying); and (iii) considering running alternative evidence streams alongside MOS-corrected RCM projections, be this response surfaces or other change factor analysis or freeboard estimates. If you must use MOS then don’t use it alone.

6. Conclusions

A grand ensemble of projections from a number of GCM/RCMs was used to force the HBV hydrological model and analyse the resulting future flood projections for the Upper Severn, UK, and the impact and implications of applying MOS techniques to precipitation fields was examined. The impact of hydrological model parameter uncertainty was taken into account. The resultant grand ensemble of future river discharge projections was compared

with a response surface technique combined with perturbed physics ensemble climate model outputs. The ensemble distribution of results shows that future risk of flooding in the Upper Severn increases compared to present conditions, particularly with regard to the probability of exceeding the flood warning threshold at Montford, but the study also highlights the large uncertainties in results and the strong assumptions made in using MOS to produce the estimates of future discharge. MOS has a clear effect on the results when the RCM output was used directly in combination with the hydrological model.

Because of the potential problems associated with using MOS without understanding the effects, here uncorrected and MOS-corrected precipitation products and modelled discharge have been presented together. We strongly advocate this for all future climate impact studies in order to make transparent what the RCMs are actually producing in terms of precipitation and the effects that MOS are having on the data. We also challenge the routine use of MOS in climate impact studies. The inability of the RCMs to produce realistic precipitation which can be used in local climate impact studies on flooding, even in present conditions, is a serious issue, and this should be a focus for future development. We advocate using multiple evidence streams, including using grand ensembles of RCMs and different ensemble techniques in analysing future flooding. We also acknowledge that climate model projection techniques should be combined with alternative strategies such as 'catchment change scenarios' in order to present the most robust understanding of future flood risk.

Acknowledgements

The authors were funded by NERC FREE grant number NE/E002242/1 and NERC Storm Risk Mitigation, project DEMON, grant number NE/I005366/1. Thanks to the Environment Agency of England and Wales, the UK Met Office Hadley Centre and FP6 ENSEMBLES for data provision, advice and assistance. Thanks to Glenn McGregor and Matt Wilson for initial ideas on project.

References

- Akhtar MN, Ahmad, Booij MJ. 2009. Use of regional climate model simulations as input for hydrological models for the Hindukush–Karakorum–Himalaya region. *Hydrology and Earth System Sciences* **13**: 1075–1089.
- Arheimer B, Lindström G, Olsson J. 2011. A systematic review of sensitivities in the Swedish flood-forecasting system. *Atmospheric Research* **100**: 275–284.
- Bell VA, Kay AL, Jones RG, Moore RJ. 2007. Use of grid-based hydrological model and regional climate model outputs to assess changing flood risk. *International Journal of Climatology* **27**: 1657–1671.
- Bergström S. 1992. *The HBV model: its structure and applications*. RH No.4, SMHI Hydrology, Norrköping, Sweden.
- Beven KJ. 2011. I believe in climate change but how precautionary do we need to be in planning for the future? *Hydrological Processes* **25**: 1517–1520.
- Böhm U, Kücken M, Ahrens W, Block A, Hauffe D, Keuler K, Rockel B, Will A. 2006. Clm – the climate version of Im: brief description and long-term applications. *COSMO Newsletter* **6**.
- Chen H, Xu C-Y, Guo S. 2012. Comparison and evaluation of multiple GCMs, statistical downscaling and hydrological models in the study of climate change impacts on runoff. *Journal of Hydrology* **434–435**: 36–45.
- Christensen OB, Drews M, Christensen JH, Dethloff K, Ketelsen K, Hebestadt I, Rinke A. 2007. *The HIRHAM Regional Climate Model Version 5*. Technical Report 06-17, Danish Climate Centre, Danish Meteorological Institute, Copenhagen.
- Cloke HL, Pappenberger F. 2009. Ensemble flood forecasting: a review. *Journal of Hydrology* **375**: 613–626.
- Cloke HL, Pappenberger F, Renaud J-P. 2008. Multi-method global sensitivity analysis (MMGSA) for modelling floodplain hydrological processes. *Hydrological Processes* **22**: 1660–1674.
- Cloke HL, Jeffers C, Wetterhall F, Pappenberger F, Byrne T, Lowe J. 2010. Climate impacts on river flow: projections for the Medway catchment UK with UKCIP and CATCHMOD. *Hydrological Processes* **24**: 3476–3489.
- Collins M, Tett SFB, Cooper C. 2001. The internal climate variability of HadCM3, a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics* **17**: 61–81.
- 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, Kilsby CG, Stunell J. 2007b. Modelling the impacts of projected future climate change on water resources in northwest England. *Hydrology and Earth System Sciences* **11**: 1115–1126.
- Giorgi F, Mearns LO. 1999. Introduction to special section: regional climate modeling revisited. *Journal of Geophysical Research* **104**: 6335–6352.
- Harris GR, Collins M, Sexton DMH, Murphy JM, Booth BBB. 2010. Probabilistic projections for 21st century European Climate. *Natural Hazards and Earth Systems Science* **10**: 2009–2020.
- Haugen JE, Haakenstad H. 2006. *Validation of HIRHAM version 2 with 50 km and 25 km resolution*. RegClim General Technical Report No. 9; 159–173.
- He Y, Wetterhall F, Cloke HL, Pappenberger F, Wilson M, Freer J, McGregor G. 2009. Tracking the uncertainty in flood alerts driven by grand ensemble weather predictions. *Meteorological Applications* **16**: 91–101.
- IPCC. 2007. *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. Cambridge University Press: Cambridge, UK.
- 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.
- Jones CG, Willén U, Ullerstig A, Hansson U. 2004. The Rossby Centre regional atmospheric climate model part I: model climatology and performance for the present climate over Europe. *Ambio* **33**: 199–210.
- Jones RN. 2001. An environmental risk assessment/management framework for climate change impact assessments. *Natural Hazards* **23**: 197–230.
- Kay AL, Jones RG, Reynard NS. 2006. RCM rainfall for UK flood frequency estimation II: climate change results. *Journal of Hydrology* **318**: 163–172.
- Kjällström E, Bärring L, Gollvik S, Hansson U, Jones C, Samuelsson P, Rummukainen M, Ullerstig A, Willén U, Wyser K. 2005. *A 140-year simulation of European climate with the new version of the Rossby Centre regional atmospheric climate model (RCA3)*. SMHI Reports Meteorology and Climatology 108, SMHI, SE-60176 Norrköping, Sweden.
- Kjällström E, Bøberg F, Castro M, Christensen JH, Nikulin G, Sánchez E. 2010. Daily and monthly temperature and precipitation statistics as performance indicators for regional climate models. *Climate Research* **44**: 135–150.
- Leander R, Buishand TA. 2007. Resampling of regional climate model output for the simulation of extreme river flows. *Journal of Hydrology* **332**: 487–496.
- Lenderink G, van den Hurk B, van Meijgaard E, van Ulden AP, Cuijpers AH. 2003. *Simulation of present-day climate in RACMO2: first results and model developments*. KNMI Technical Report 252.
- Lidén R, Harlin J. 2000. Analysis of conceptual rainfall–runoff modelling performance in different climates. *Journal of Hydrology* **238**: 231–247.
- Linde AHT, Aerts J, Bakker AMR, Kwadijk JCI. 2010. Simulating low-probability peak discharges for the Rhine basin using resampled climate modeling data. *Water Resources Research* **46**: W03512.
- Lindström G, Johansson B, Persson M, Gardelin M, Bergström S. 1997. Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology* **201**: 272–288.
- Maraun D, Wetterhall F, Ireson AM, Chandler RE, Kendon EJ, Widmann M, Brienen S, Rust HQ, Sauter T, Themessl M, Venema VKC, Chun KP, Goodess CM, Jones RG, Onof C, Vrac M, Thiele-Eich I. 2010. Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics* **48**: RG3003.
- Marke T, Mauser W, Pfeiffer A, Zängl G. 2011. A pragmatic approach for the downscaling and bias correction of regional climate simulations:

- evaluation in hydrological modeling. *Geoscientific Model Development* **4**: 759–770.
- McGuinness JL, Bordne EF. 1972. *A comparison of lysimeter-derived potential evapotranspiration with computed values*. Technical Bulletin 1452, Agricultural Research Service, US Department of Agriculture, Washington, DC.
- Murphy JM, Sexton DMH, Barnett DN, Jones GS, Webb MJ, Collins M, Stainforth DA. 2004. Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature* **430**: 768–772.
- Murphy JM, Sexton DMH, Jenkins GJ, Booth BBB, Brown CC, Clark RT, Collins M, Harris GR, Kendon EJ, Betts RA, Brown SJ, Humphrey KA, McCarthy MP, McDonald RE, Stephens A, Wallace C, Warren R, Wilby R, Wood RA. 2009. *UK Climate Projections Science Report: Climate change projections*. Met Office Hadley Centre, Exeter, UK.
- Nash JE, Sutcliffe JV. 1970. River flow forecasting through conceptual models. Part I: A discussion of principles. *Journal of Hydrology* **10**: 282–290.
- Olsson J, Lindstrom G. 2008. Evaluation and calibration of operational hydrological ensemble forecasts in Sweden. *Journal of Hydrology* **350**: 14–24.
- Oudin L, Hervieu F, Michel C, Perrin C, Andréassian V, Anctil F, Loumagne C. 2005. Which potential evapotranspiration input for a rainfall-runoff model? Part 2: Towards a simple and efficient PE model for rainfall-runoff modelling. *Journal of Hydrology* **303**: 290–306.
- Pappenberger F, Beven KJ. 2006. Ignorance is bliss: or 7 reasons not to use uncertainty analysis. *Water Resources Research* **42**: W05302.
- Pappenberger F, Bartholmes J, Thielen J, Cloke HL, de Roo A, Buizza R. 2008. New dimensions in early flood warning across the globe using GRAND ensembles. *Geophysical Research Letters* **35**: L10404.
- Perry M, Hollis D, Elms M. 2008. *The generation of daily gridded datasets for temperature and precipitation, covering the UK for the period 1960 to 2006*. Internal report, Met Office, Exeter, UK.
- Plummer D, Caya D, Côté H, Frigon A, Biner S, Giguère M, Paquin D, Harvey R, de Elia R. 2006. Climate and climate change over North America as simulated by the Canadian Regional Climate Model. *Journal of Climate* **19**: 3112–3132.
- Pope VD, Gallani ML, Rowntree PR, Stratton R. 2000. The impact of new physical parameterizations in the Hadley Centre climate model: HadAM3. *Climate Dynamics* **16**: 123–146.
- Prudhomme C, Wilby RL, Crooks S, Kay AL, Reynard NS. 2010. Scenario-neutral approach to climate change impact studies: application to flood risk. *Journal of Hydrology* **390**: 198–209.
- Radu R, Déqué M, Somot S. 2008. Spectral nudging in a spectral regional climate model. *Tellus A* **60**: 898–910.
- Rivington M, Miller D, Matthews KB, Russell G, Bellocchi G, Buchan K. 2008. Evaluating regional climate model estimates against site-specific observed data in the UK. *Climatic Change* **88**: 157–185.
- Rojas R, Feyen L, Dosio A, Bavera D. 2011. Improving pan-European hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations. *Hydrology and Earth System Sciences* **15**: 2599–2620.
- Sanchez E, Gallardo C, Gaertner MA, Arribas A, Castro M. 2004. Future climate extreme events in the Mediterranean simulated by a regional climate model: a first approach. *Global and Planetary Change* **44**: 163–180.
- Seibert J. 1997. Estimation of parameter uncertainty in the HBV model. *Nordic Hydrology* **28**: 247–262.
- Seibert J. 2003. Reliability of model predictions outside calibration conditions. *Nordic Hydrology* **34**: 477–492.
- Shkolnik IM, Meleshko VP, Pavlova TV. 2000. Regional hydrodynamical model for climate studies over Russia. *Russian Meteorology and Hydrology* **4**: pp. 32–49.
- Sobol IM. 1979. On the systematic search in a hypercube. *SIAM Journal of Numerical Analysis* **16**: 790–793.
- Sobol IM. 1993. Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiment* **1**: 407–414.
- Somot S, Spiridonov V, Marquet P, Déqué M. 2004. Climate version of the LAM ALADIN: first validation of a 10-year simulation over Europe. In *Regional Climate Model Workshop*, Prague, 29 November–3 December 2004.
- Stainforth DA, Downing TE, Washington R, Lopez A, New M. 2007. Issues in the interpretation of climate model ensembles to inform decisions. *Philosophical Transactions of the Royal Society A* **365**: 2163–2177.
- Teutschbein C, Seibert J. 2010. Regional climate models for hydrological impact studies at the catchment scale: a review of recent modeling strategies. *Geography Compass* **4**: 834–860.
- Themessl MJ, Gobiet A, Heinrich G. 2012. Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal. *Climatic Change* **112**: 449–468.
- Turco M, Quintana-Segui P, Llasat MC, Herrera S, Gutiérrez JM. 2011. Testing MOS precipitation downscaling for ENSEMBLES regional climate models over Spain. *Journal of Geophysical Research-Atmospheres* **116**: D18109.
- UK Met Office. 2011. MIDAS Land Surface Stations data (1853–current), British Atmospheric Data Centre, 2006, <http://badc.nerc.ac.uk/data/ukmo-midas>
- Uppala SM, Kållberg PW, Simmons AJ, Andrae U, Bechtold VDC, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Berg LVD, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen L, Janssen PAEM, Jenne R, McNally AP, Mahfouf J-F, Morcrette J-J, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J. 2005. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society* **131**: 2961–3012.
- van der Linden P, Mitchell FJB (eds). 2009. *ENSEMBLES: Climate Change and its Impacts: Summary of Research and Results from the ENSEMBLES Project*. Met Office Hadley Centre, Exeter, UK.
- van Meijgaard E, Meijgaard van Ulft LH, van de Berg WJ, Bosveld FC, van den Hurk BJJM, Lenderink G, Siebesma AP. 2008. *The regional atmospheric climate model RACMO, version 2.1*. KNMI Technical Report TR-302, De Bilt, Netherlands.
- van Pelt SC, Kabat P, ter Maat HW, van den Hurk BJJM, Weerts AH. 2009. Discharge simulations performed with a hydrological model using bias corrected regional climate model input. *Hydrology and Earth System Sciences* **13**: 2387–2397.
- Weisheimer A, Palmer TN, Doblus-Reyes FJ. 2011. Assessment of representations of model uncertainty in monthly and seasonal forecast ensembles. *Geophysical Research Letters* **38**: L16703.
- Wetterhall F, Graham LP, Andreasson J, Rosberg J, Yang W. 2011. Using ensemble climate projections to assess probabilistic hydrological change in the Nordic region. *Nat Hazards and Earth System Sciences* **11**: 2295–2306.
- Wetterhall F, Pappenberger F, He Y, Freer J, Cloke HL. 2012. Conditioning model output statistics of Regional Climate Model precipitation on circulation patterns. *Nonlinear Processes in Geophysics* (in press).
- Yang W, Andreasson J, Graham LP, Olsson J, Rosberg J, Wetterhall F. 2010. Distribution based scaling to improve the usability of RCM regional climate model projections for hydrological climate change impacts studies. *Hydrology Research* **41**: 211–229.