WATER RESOURCES RESEARCH, VOL. 45, W11411, doi:10.1029/2007WR006674, 2009

Click Here for Full Article

Using probabilistic climate change information from a multimodel ensemble for water resources assessment

L. J. Manning,[1] J. W. Hall,[1] H. J. Fowler,[1] C. G. Kilsby,[1] and C. Tebaldi[2]

Received 16 November 2007; revised 8 June 2009; accepted 27 July 2009; published 11 November 2009.

[1] Increasing availability of ensemble outputs from general circulation models (GCMs) and regional climate models (RCMs) permits fuller examination of the implications of climate uncertainties in hydrological systems. A Bayesian statistical framework is used to combine projections by weighting and to generate probability distributions of local climate change from an ensemble of RCM outputs. A stochastic weather generator produces corresponding daily series of rainfall and potential evapotranspiration, which are input into a catchment rainfall-runoff model to estimate future water abstraction availability. The method is applied to the Thames catchment in the United Kingdom, where comparison with previous studies shows that different downscaling methods produce significantly different flow predictions and that this is partly attributable to potential evapotranspiration predictions. An extended sensitivity test exploring the effect of the weights and assumptions associated with combining climate model projections illustrates that under all plausible assumptions the ensemble implies a significant reduction in catchment water resource availability.

Citation: Manning, L. J., J. W. Hall, H. J. Fowler, C. G. Kilsby, and C. Tebaldi (2009), Using probabilistic climate change information from a multimodel ensemble for water resources assessment, *Water Resour. Res.*, *45*, W11411, doi:10.1029/2007WR006674.

1. Introduction

[2] Large ensembles of climate model outputs provide the potential to explore some of the uncertainties in model-based predictions of future climate at both global and regional scales. The next generation of the UK Climate Impacts Programme scenarios (UKCP09) will be based upon more than 300 integrations of the general circulation model (GCM) HadCM3, 17 realizations of the regional climate model (RCM) HadRM3 and a Bayesian analysis structure that will be used to generate probability distributions of climate variables on a 25 km grid [*Murphy et al.*, 2007, 2009; *Collins et al.*, 2006]. A stochastic weather generator (WG) [*Kilsby et al.*, 2007] is to be used to generate daily time series outputs representative of future climates for impacts studies where sequencing and covariance of weather variables is of significance. Though this coupled system of models will in several respects be state of the art, it is important to recognize the implications of the various methodological choices and to present decision makers with information on the implications of those assumptions to avoid the misrepresentation of uncertainties [*Hall*, 2007].

[3] Studies of the impacts of modeled climate change on water resources and the surrounding uncertainties have been published previously [e.g., *Wilby and Harris*, 2006; *New et al.*, 2007; *Dessai and Hulme*, 2007; *Christensen and Lettenmaier*, 2007; *Nawaz and Adeloye*, 2006]. In this study we examine the effect of a number of methodological choices on the uncertainty estimates provided by the increasingly prevalent use of ensemble climate model outputs. Specifically we use the outputs from a multimodel ensemble experiment including several different RCMs [*Christensen et al.*, 2007], which serves to demonstrate the sensitivity of hydrological model outputs to the choice of both GCM and RCM. In addition, we introduce a method to allow the investigation of the sensitivity of predicted climate impact to uncertainties in combining output from different RCMs.

[4] To provide an application of practical significance, and to enable comparison with previous studies [*Wilby et al.*, 2006; *Wilby and Harris*, 2006], the methodological developments are presented in the context of an application to water resources management in the Thames catchment in the southern United Kingdom. Over 90% of the water abstracted from the River Thames supplies drinking water for London and for other population centers in the Thames valley, and the entire river is classified as "overabstracted" or "no water available" [*Environment Agency*, 2004]. In order to safeguard the existing drinking water supply, while at the same time maintaining water quality in the river, the Thames corridor abstraction management strategy [*Environment Agency*, 2004] prescribes conditions on the granting of new consumptive abstraction licenses. New licenses will permit abstraction only between November and March, and only on days when the gauged flow at Kingston, near the upper tidal limit, is above 1780 ML d^{-1} (20.6 m^3 s^{-1}), the median flow (Q50) calculated over the 15 years prior to the publication of the strategy. Examination of historical flow records show that if this regulation had been in place, abstraction would have been severely restricted during four of the last thirty five winters. The years with severe restriction would have been those with the lowest aggregated rainfall from June to March, i.e., the winters of 1975–1976 (when abstraction

[1]School of Engineering and Geosciences, Newcastle University, Newcastle upon Tyne, UK.
[2]Climate Central, Princeton, New Jersey, USA.

0043-1397/09/2007WR006674$09.00

would have been permitted on 23% of the potentially available days), 1991–1992 (23% availability), 1996–1997 (25% availability), and 2004–2005 (38% availability).

[5] The structure of the paper reflects the multiple stages of the analysis required for climate impacts studies, and is as follows. In section 2 we consider the use of stochastic WGs for downscaling climate model outputs and examine the performance of a WG for the Thames catchment. Section 3 examines predictions of water availability in present and future climates based on an ensemble of GCM and RCM outputs. In section 4 we discuss the problem of weighting ensemble outputs and examine the implications of different weighting schemes for the Thames catchment, leading to an analysis of the sensitivity of the impact to the weighting scheme. At each stage in the analysis we present the relevant methodology, results and discussion so that the reader can evaluate intermediate outputs before proceeding to the next stage in the climate impacts study. The paper concludes in section 5.

2. Downscaling of Climate Model Outputs

[6] While statistical downscaling of GCM outputs is relatively inexpensive and has been widely used, dynamical downscaling using RCMs is more attractive if large numbers of jointly varying outputs are required [*Fowler et al.*, 2007a]. However, further downscaling may still be required to give more spatial detail, including the effects of local topography, and to provide the extended time series of variables needed for risk assessment studies. To this end, stochastic WGs have been used in a number of hydrological contexts [*Wilks and Wilby*, 1999], including the downscaling of future climate projections. The Environment Agency Rainfall and Weather Impacts Generator (EARWIG) is a WG that provides simulated rainfall and other weather variables over a 5km grid throughout Great Britain [*Kilsby et al.*, 2007]. The baseline rainfall time series is provided by a Neyman-Scott rectangular pulses rainfall model [*Cowpertwait*, 1991], whose statistics (rainfall mean, variance, skewness, 24 h lagged autocorrelation, and percentage of dry days) are derived from a 5 km grid of observed daily rainfall records throughout the United Kingdom, produced by interpolation from several thousand rain gauge records [*Perry and Hollis*, 2005]. Diurnal temperature mean and range are modeled as autoregressive processes, with coefficients determined by the time of year, and whether the current and previous day are wet or dry. Other derivative weather variables, such as vapor pressure, wind speed and sunshine duration, are generated using multiple regression models, using monthly data from the same grid, and daily data from a United Kingdom–wide network of 115 stations. Potential evapotranspiration is then calculated from these variables, using either the Food and Agriculture Organization of the United Nations (FAO) modified version of the Penman method [*Allen et al.*, 1994] or the MORECS method [*Hough and Jones*, 1997].

[7] In order to generate output representative of future climate conditions, multiplicative change factors for the daily rainfall statistics required by the Neyman-Scott rectangular pulses model and temperature standard deviation, and additive change factors for the daily mean temperature are obtained from the difference between current and future climates for each RCM grid cell. [*Kilsby et al.*, 2007]. These change factors are used to adjust the rainfall and temperature statistics which generate rainfall and weather time series. The derived variables are calculated as for the current climate.

[8] The use of change factors (CFs), often called the "perturbation method" [*Prudhomme et al.*, 2002] or "delta change" approach, assumes that the climate models represent relative change more accurately than absolute climate values, and relies on the climate model bias being constant through time [*Fowler et al.*, 2007a]. Normally, the CFs are applied to only the mean of observed time series and so no changes are made to the sequencing of weather or to its variability or extremes. However, here we apply the CFs through a stochastic method and to variance and extreme statistics, thus allowing new sequences of weather and changes to variability and extremes to be accounted for. In situations where climate models show noticeable bias in reproducing regional climate then their capacity to represent future may be questioned. For this reason, while using change factors in the WG, we adopt an approach, which is explained in more detail in section 4, that weights the climate model predictions according to their ability to reproduce observed climate. In this way models that are particularly biased are down weighted.

[9] In this study the models used for future climate projections are the UKCIP02 scenarios [*Hulme et al.*, 2002] based on the HadRM3H RCM, and 13 RCM integrations from PRUDENCE [*Christensen et al.*, 2007] from which the change factors for daily weather variables were extracted. The PRUDENCE models used are listed in Table 1, and biases in their representation of United Kingdom precipitation and temperature statistics have been assessed by *Blenkinsop and Fowler* [2007] and *Blenkinsop et al.* [2008], respectively. It should be noted that while the ensemble of model integrations represent 11 different RCMs, there are essentially only 2 driving GCMs represented here, as HadAM3H and HadAM3P are atmospheric models, used as an intermediate step between the GCM HadCM3 and the RCM. HadAM3H and HadAM3P are very similar, with slight differences in the details of cloud representation and thresholds for precipitation formation [*Moberg and Jones*, 2004]. The models are described by *Hagemann and Jacob* [2007].

[10] While the EARWIG weather generator was established to generate point rainfall (and weather) series, an areally averaged time series is required for water resources assessments. It is well known [*Faulkner*, 1999] that areally averaged rainfall series demonstrate rather different characteristics from point series, in particular having reduced variance, and a reduced proportion of dry days. EARWIG was adapted for use for the Thames catchment (10,000 km^2) as follows. The rainfall parameterization for present-day climate was performed using statistics derived from the areal average of historical rainfall series across the catchment. For future climates, averaged change factors for the relevant EARWIG variables were generated by calculating the proportion of the catchment within each 50 km RCM grid square and calculating the change factors from the areally weighted average of the daily RCM output.

[11] *Kilsby et al.* [2007] demonstrated the ability of the stochastic WG used in this study to reproduce daily and monthly rainfall and weather statistics, and rainfall extremes, at Heathrow, near the downstream boundary of the catchment. In order to further validate the reproduction

Table 1. Climate Models Used From the PRUDENCE Ensemble

Model Acronym	Institution	RCM	Driving GCM
ARP.HAD	Meteo-France, France	Arpège	HadCM3
HADR.HAD	Hadley Centre, UK Meteorological Office	HadRM3P	HadAM3P
HIR.ECH	Danish Meteorological Institute (DMI)	HIRHAM	ECHAM4/OPYC3
HIR.HAD	Danish Meteorological Institute (DMI)	HIRHAM	HadAM3H
RCA.ECH	Swedish Meteorological and Hydrological Institute (SMHI)	RCAO	ECHAM4/OPYC3
RCA.HAD	Swedish Meteorological and Hydrological Institute (SMHI)	RCAO	HadAM3H
CHRM.HAD	Swiss Federal Institute of Technology (ETH), Zurich	CHRM	HadAM3H
CLM.HAD	GKSS, Institute for Coastal Research, Germany	CLM	HadAM3H
REMO.HAD	Max Planck Institute for Meteorology, Hamburg, Germany	REMO	HadAM3H
PROMES.HAD	Universidad Complutense de Madrid (UCM), Spain	PROMES	HadAM3H
REGCM.HAD	Abdus Salam International Centre for Theoretical Physics (ICTP), Italy	RegCM	HadAM3H
RACMO.HAD	Royal Netherlands Meteorological Institute (KNMI)	RACMO	HadAM3H
METNO.HAD	Norwegian Meteorological Institute	MetNo	HadAM3H

of historical temperature and potential evapotranspiration (PET) statistics, an ensemble of 20 EARWIG runs was generated, conditioned on the historical daily areal rainfall record for the Thames catchment from 1961 to 1990. The resulting summer (June–July–August) and winter (December–January–February) mean temperature was compared with a mean historical record for the Thames catchment. Similarly, the cumulative PET, calculated using the MORECS method [*Hough and Jones*, 1997], was compared with the historical MORECS PET record. The comparison is shown in Table 2, demonstrating that EARWIG reproduces the overall mean and standard deviation of summer and winter PET and temperature well. Figure 1, giving annual seasonal means of PET and temperature, shows that EARWIG is able to reproduce the variability range of summer and winter temperature and PET apart from the magnitude of individual historic extreme events (such as the summer of 1976). Total seasonal PET is well constrained by the historical rainfall pattern, but EARWIG does not reproduce PET well for the hot, dry summer of 1976. This shows a limitation of the use of a first-order autoregressive process in EARWIG to simulate temperature, and by association sunshine hours, thus failing to account for the intensification of heat and evaporation in long dry periods, which is being addressed in further work. However, the possible underestimation of PET in very hot summers is offset in the subsequent hydrological modeling because actual evaporation is limited by moisture supply rather than determined by PET.

[12] The rainfall and other weather variables modeled in EARWIG are based upon short-term weather statistics, and therefore cannot necessarily be expected to reproduce patterns of longer-term variability. To use EARWIG to reproduce climates exhibiting significant interannual correlation, the output would need to be conditioned to reflect that characteristic. However, analysis of the annual rainfall totals derived from the 133 year spatially averaged historical record for southeast England show no significant autocorrelation between annual rainfall totals at lags of 1, 2, or 3, indicating that at this site no improvement in the model would be achieved by introducing interannual correlation.

[13] Summer and winter mean temperature and precipitation change predicted by the PRUDENCE models for the Thames basin, as well as model bias for current climate, are listed in Table 3. All models predict a rise in temperature in both winter and summer, while precipitation is predicted to increase in winter, and decrease in summer. It is noticeable that the models driven by ECHAM4 predict a higher increase in summer temperature than the other models; this has been demonstrated in more detail by *Blenkinsop and Fowler* [2007]. The predicted changes in Table 3 refer to the 2080s, and the A2 emissions scenario [*Nakicenovic et al.*, 2000]. Since RCM output is not available for all time periods and emissions scenarios, predictions for other time periods and scenarios are calculated by scaling these outputs on relative global temperature change from the driving GCM. Scaling factors for HadCM3-derived models are given by *Hulme et al.* [2002], and equivalents for ECHAM4 were calculated using the change factors from the IPCC Web site, http://www.ipcc-data.org/cgi-bin/ddcvis/gcmcf/. Further details on the application of scale factors are given by *Kilsby et al.* [2007].

3. Ensemble Prediction of Flow Series

[14] The flow series in the Thames catchment has been modeled using the CATCHMOD rainfall-runoff model, which is a water balance model used for water resource planning by the Environment Agency, and has been described in detail elsewhere [*Wilby et al.*, 1994; *Davis*, 2001]. CATCHMOD is a lumped parameter conceptual model, which allows for the subdivision of the catchment into a number of zones, according to its geological and surface runoff characteristics. The parameterization of this model for the Thames basin used here is that described by *Wilby* [2005], and involves three zones, representing clay, limestone and urban regions. Input to the model is in the

Table 2. Comparison Between Historical Data and Model Output of Daily Mean and Standard Deviation of Temperature and Potential Evapotranspiration for the Thames Catchment When Forced by the Historical Rainfall Record[a]

	Historical Data 1961–1990	EARWIG Model Output
Summer temperature (deg C)	15.8 (2.6)	15.4 (2.6)
Winter temperature (deg C)	4.2 (3.6)	4.3 (3.4)
Daily PET, summer (mm)	3.0 (1.26)	2.7 (0.85)
Daily PET, winter (mm)	0.62 (0.56)	0.53 (0.44)

[a]Standard deviation is given in parentheses. Potential evapotranspiration (PET) is calculated using the MORECS method. Summer is June–July–August, and winter is December–January–February.

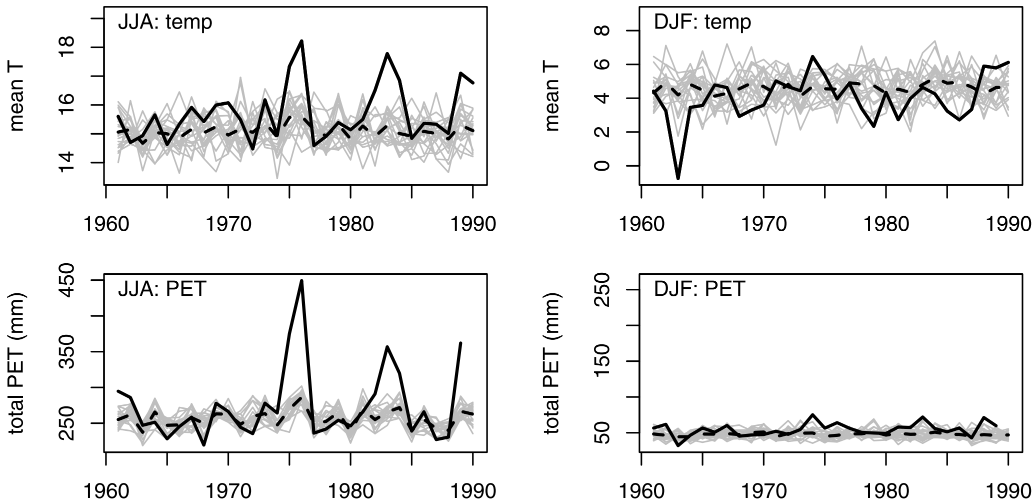

Figure 1. Comparison of historical record (seasonal mean) with WG output with output from 20 runs of the WG, conditioned by observed rainfall (Thames catchment mean). (top) Temperature compared with annual mean observed temperature for the Thames catchment. (bottom) Potential evapotranspiration compared with MORECS PET record. Solid black lines, historical record; gray lines, simulated data; dashed black lines, mean of WG outputs.

form of time series of daily rainfall and potential evapotranspiration representative of the entire catchment and the output is a time series of the daily natural flow at Kingston. This is the flow which would exist in the absence of artificial abstractions, and the model has been calibrated with reference to the daily historical flow, corrected for abstraction. However, since the application requires calculation of the gauged flow at Kingston for simulated weather series, the modeled output flow series is modified by the subtraction of the annually averaged daily difference between historical gauged and naturalized flows at Kingston, for the time period over which the rainfall-runoff model was parameterized. The mean correction over the time period of model calibration was 18 $m^3 s^{-1}$.

[15] Daily weather records have been simulated using the WG for 1000 year periods, for both current and future climate. These have been used as input to the catchment model, and naturalized flow quantiles have been calculated. Figures 2 and 3 show box plots of 30 year estimated natural flow quantiles for the 2050s and 2080s, respectively, under the A2 emissions scenario, with the baseline current climate estimates, and a single estimate for historical flows from 1961 to 1990, for comparison. These box plots have been calculated using the single, manually calibrated, catchment model parameterization detailed by *Wilby* [2005], by examination of the variability in the quantiles calculated for each 30 year period within the model output record. A 30 year period was chosen for comparison with the historical record. A progressive decrease with time in the level of each flow quantile can be seen. The range of predicted change in the lower flows, Q95 (the flow level exceeded on 95% of days) and Q50 (exceeded on 50% of days), also decreases with

Table 3. Bias With Respect to Historical Data and Change in Seasonal Temperature and Precipitation Predicted by UKCIP02 and PRUDENCE Climate Models for the A2 Scenario, 2080s[a]

	Mean Temperature Bias (deg C)		Mean Precipitation Bias (%)		Model Mean Temperature Change (deg C)		Model Mean Precipitation Change (%)	
Model Acronym	DJF	JJA	DJF	JJA	DJF	JJA	DJF	JJA
UKCIP02	0.83	0.36	5	−8	2.7	4.7	24	−48
ARP.HAD	0.97	−0.07	46	0	2.3	2.9	19	−19
HADR.HAD	1.22	0.94	0	−35	2.3	4.3	25	−34
HIR.ECH	1.64	−0.06	27	31	3.4	5.1	13	−30
HIR.HAD	1.93	0.92	27	−5	2.3	3.6	23	−33
RCA.ECH	2.65	0.14	60	33	3.7	5.7	34	−58
RCA.HAD	2.42	0.53	68	9	2.2	3.9	25	−46
CHRM.HAD	0.61	−0.02	9	−9	1.9	3.5	30	−42
CLM.HAD	0.59	−0.27	17	−7	2.1	4.0	31	−46
REMO.HAD	1.75	1.04	19	37	2.3	3.4	26	−40
PROMES.HAD	1.82	−1.07	27	−3	2.2	3.9	24	−29
REGCM.HAD	1.34	−0.46	30	57	2.1	3.8	31	−39
RACMO.HAD	1.56	0.58	45	5	2.2	3.6	29	−45
METNO.HAD	1.65	0.57	38	−4	2.3	3.6	26	−39

[a]DJF, December–January–February; JJA, June–July–August.

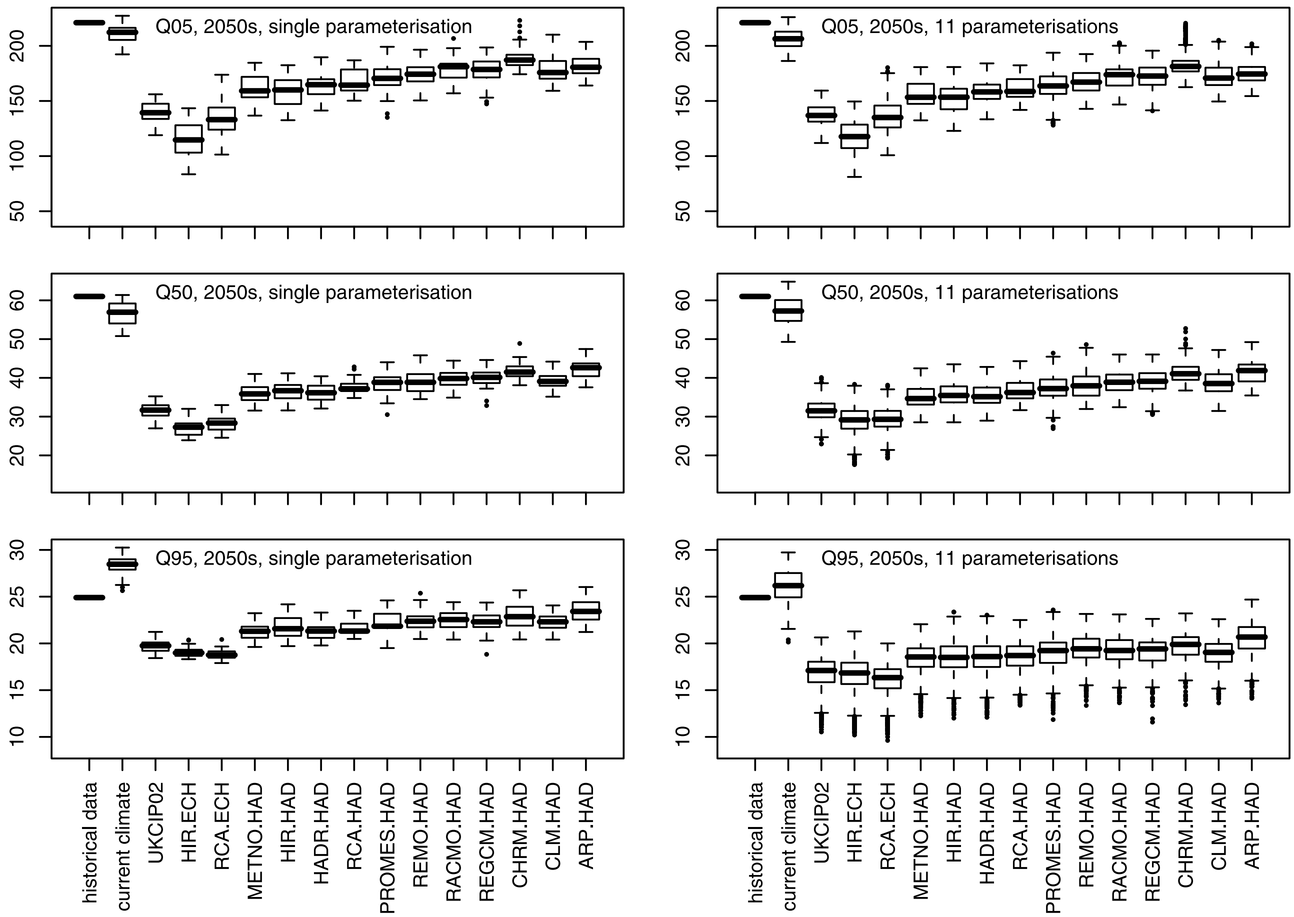

Figure 2. Estimated natural flow quantiles for the river Thames at Kingston ($m^3\ s^{-1}$), 2050s, using each of the PRUDENCE ensemble outputs and comparing a single parameterization with an ensemble of 11 parameterizations of the catchment model. (Maximum box plot whisker length is 1.5 times the interquartile range.)

time, while the range of the high-flow predictions, Q5 (exceeded on 5% of days), increases.

[16] Figures 2 and 3 also show the additional uncertainty in the estimation of the flow quantiles caused by hydrological uncertainty. Eleven new parameterizations of the catchment model have been tested, based on the two criteria employed by *Wilby* [2005], the Nash-Sutcliffe efficiency at reproducing observed flow, and an absolute flow measure. These parameterizations are Pareto-optimal, meaning that for each parameterization, no improvement in agreement with one criterion can be made without compromising the other, and have been chosen to represent the full spectrum of optimal models relating to these two criteria. Agreement between historical quantiles and those calculated from the output of these eleven models is better than that obtained for the manually calculated model shown in Figure 2, although the model output does underestimate high flows. These, however, are not the purpose of this study. The change in total uncertainty due to model parameterization increases with future time period, as the model input conditions deviate further from those used for parameterization, and it is interesting to note that while there is little difference between the estimation ranges for the high-flow quantile, Q5, the difference is greatest for the low-flow quantile, Q95, with an increase in standard deviation (not shown) between a single and multiple parameterizations by a factor of 2.4 in projections for the 2050s, and by a factor of 5.5 in projections for the 2080s. The corresponding proportional increases in standard deviation for Q50 are 1.4 and 3.3, respectively. It should be noted that a full appraisal of hydrological uncertainty should include the uncertainty from alternative hydrological models. This has not been done here, and the analysis is restricted to hydrological parameter uncertainty.

[17] The 1000 year output flow series from the catchment model have been analyzed to create, for each climate model integration, a cumulative distribution function of the annual number of days (*Na*) on which abstraction from the Thames is permitted under the winter Q50 restriction set out in the Thames corridor catchment abstraction management strategy [*Environment Agency*, 2004] described above. Figure 4 shows, as well as future predictions, a comparison of the number of days on which abstraction would have been permitted from the river Thames using historical flow records for the period 1961–1990, corresponding to the calibration period of the hydrological model, and the estimated number of days using a synthetic EARWIG weather series calibrated to current climate. The comparison between historical flows and current climate simulations show substantial agreement. Three future time periods, the 2020s, 2050s and 2080s, and the four UKCIP02 emissions scenarios corresponding to

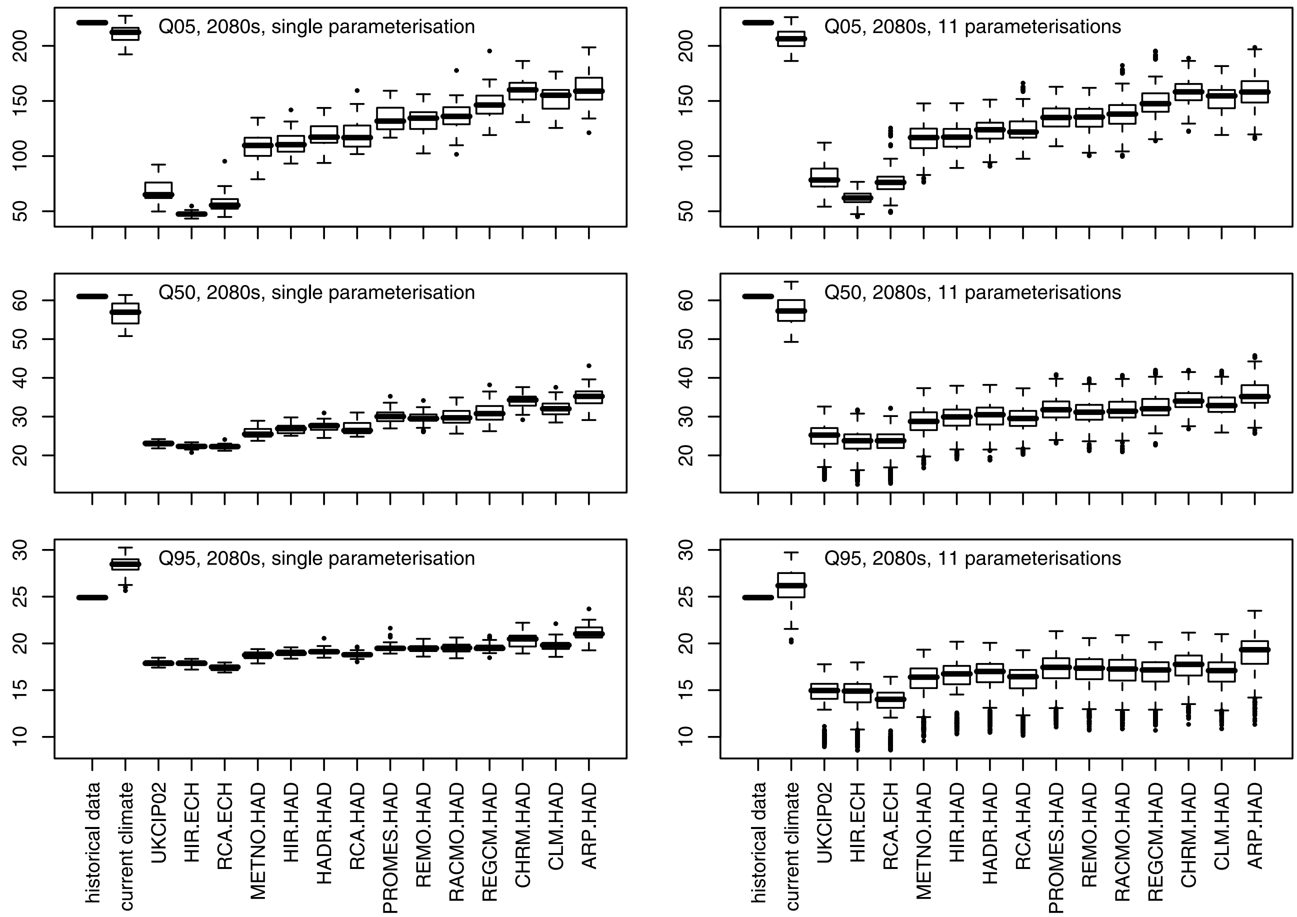

Figure 3. Estimated natural flow quantiles for the river Thames at Kingston ($m^3\ s^{-1}$), 2080s, using each of the PRUDENCE ensemble outputs and comparing a single parameterization with an ensemble of 11 parameterizations of the catchment model. (Maximum box plot whisker length is 1.5 times the interquartile range.)

the four SRES emissions scenarios, A1FI, A2, B1 and B2 [*Nakicenovic et al.*, 2000], are represented. The predictions demonstrate a significant and increasing loss of abstraction availability. The effect of different emissions scenarios on the uncertainty in abstraction availability for the 2020s is small, as indeed is the predicted difference in emissions by this time period. The influence of different emissions scenarios increases by the 2080s.

[18] The distribution functions represented in Figure 4 were calculated using EARWIG weather series derived from

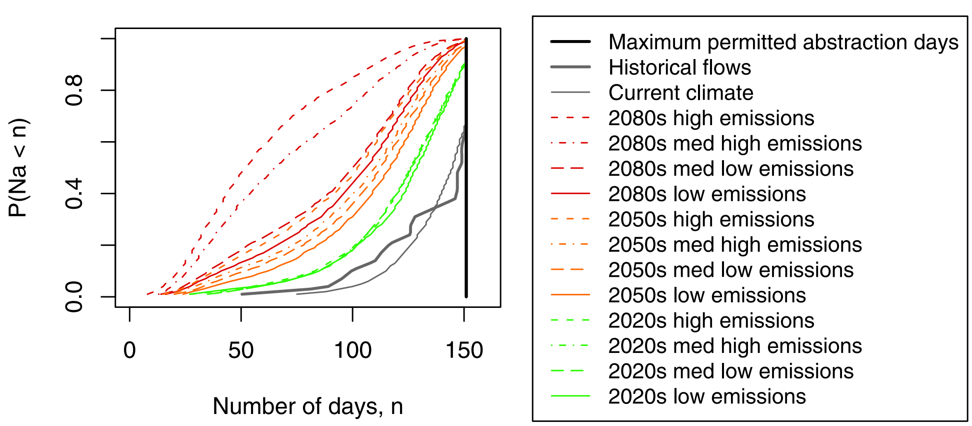

Figure 4. Cumulative distribution function of the maximum annual number of days of permitted abstraction from the river Thames (*Na*) based upon historical flows for 1961–1990; estimated using WG for current climate; and for future climate projections, using the UKCIP02 climate scenarios.

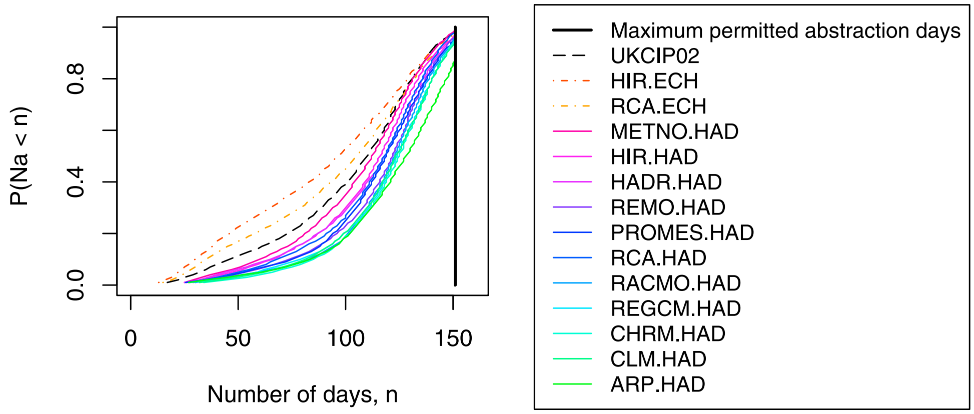

Figure 5. Cumulative distribution function of the maximum annual number of days of permitted abstraction from the Thames (*Na*), predicted for the UKCIP02 scenarios and PRUDENCE models, 2050s, medium high (SRES A2) emissions scenario.

the UKCIP02 climate model, while Figures 5 and 6 illustrate the predictions for each member of the PRUDENCE ensemble, for the time periods 2050s and 2080s, and a single emissions scenario. Note that the range of mean predicted abstraction availability for a single emissions scenario but a number of climate models (96 to 128 days, 2050s, and 48 to 118 days, 2080s) is greater than the range for a single climate model and a number of emissions scenarios (102 to 117 days, 2050s, and 52 to 106 days, 2080s), supporting the finding of other studies [e.g., *Wilby and Harris*, 2006; *Dessai and Hulme*, 2007], that the greatest source of uncertainty in climate predictions arises from the climate model.

[19] It is important to consider whether known deficiencies in the RCM modeling have an effect on these predictions of abstraction availability. While the driving GCM is extremely influential in hydrological predictions [*Wilby and Harris*, 2006], it can be seen that the predictions in Figures 5 and 6 from UKCIP02 (derived from HadCM3 and HadAM3H) are more similar to the PRUDENCE ECHAM4 models than to the other HadAM3H models. Problems with high summer water vapor pressure deficits in this RCM, causing overdrying and overheating, were reported by *Ekström et al.* [2007]. These problems were addressed by the introduction of the RCM HadRM3P (driven by the newer atmosphere-only GCM, HadAM3P, with some additional reparameterizations) which seems comparable to the HadAM3H-driven PRUDENCE RCMs for future climates in southern England. *Fowler et al.* [2007a] demonstrated agreement in temperature and precipitation changes predicted by an ensemble of 6 PRUDENCE RCMs and by the IPCC-AR4 GCMs for all seasons but summer in southeast England, where the RCM results are warmer and drier than the GCM results.

[20] As an additional check on the reasonableness of the predictions in Figures 5 and 6, a comparison has been made with two recently published studies of future water resource availability in the Thames basin [*Wilby and Harris*, 2006, hereafter WH06; *Wilby et al.*, 2006, hereafter W06]. WH06 uses Monte Carlo sampling to examine a number of the sources of uncertainty involved in prediction of future resource availability in the Thames; namely, GCM choice, downscaling method and predictor variables, and rainfall-runoff model parameterization. W06 examines water resource availability and water quality in a tributary of the Thames, the Kennet. Both studies examine outputs from a number of GCMs, including HadCM3 under the A2

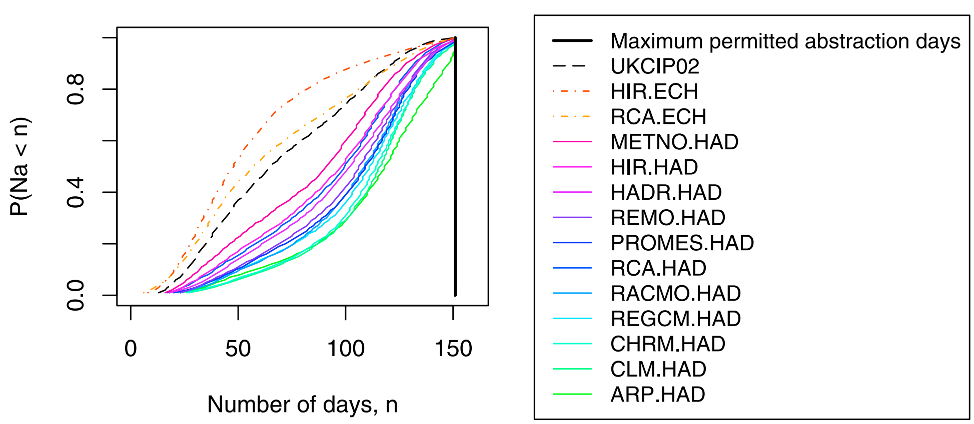

Figure 6. Cumulative distribution function of the maximum annual number of days of permitted abstraction from the Thames (*Na*), predicted for the UKCIP02 scenarios and PRUDENCE models, 2080s, medium high (SRES A2) scenario.

Table 4. Comparison Between Climate Model Predictions Made by Different Studies in the Thames Basin, 2080s, Medium High Scenario, Relative to Present Day, All Based on HadCM3

Study	WH06	W06	Calculated From EARWIG[a]		Calculated Directly From HadRM3H[b]
			HadRM3H	HadRM3P	
Mean winter (DJF) rainfall (%)	+30	+20	+28	+27	+24
Mean summer (JJA) rainfall (%)	−37	−36	−49	−35	−48
Mean winter (DJF) temperature (deg)	-	+2.5	+2.7	+2.4	+2.6
Mean summer (JJA) temperature (deg)	-	+3.7	+5.1	+4.6	+4.8
Mean winter (DJF) PET (%)	+12	+5	+98	+83	+58
Mean summer (JJA) PET (%)	+16	+8	+51	+46	+103

[a]Current study.
[b]*Ekström et al.* [2007].

emissions scenario, with downscaling provided by the SDSM package [*Wilby et al.*, 2002]. These studies predict mean percentage changes in Q95 for the 2080s by using HadCM3, as decreases relative to the present day of 15% (WH06) and 10% (W06),whereas the present study predicts decreases of 45% (HadRM3H) and 37% (HadRM3P). The difference between predictions of change in Q95 for the Thames and the Kennet may be partly a result of different behavior in the two catchments [*Arnell*, 2003], since the Thames, with an area of approximately 10,000 km^2, draws water from a mixture of chalk and clay horizons, while the Kennet, with an area of approximately 1000 km^2, is predominantly chalk. However, the key to the difference in the predictions of change in Q95 flow for the Thames is to be found in the comparison between climate predictions, shown in Table 4. While the predictions of change in temperature and rainfall are comparable, there is considerable difference between the predictions of change in PET, which is most significant in the summer months, as the absolute totals during the winter are comparatively small. The differences between the current study and those of W06 and WH06 relate to the different downscaling methods used: an RCM and WG with MORECS calculation, versus the regression-based method, SDSM [*Wilby et al.*, 2002]. While this may indicate a further measure of uncertainty to be found in downscaling, it should be noted that SDSM was fitted using monthly historical PET values, although it relies on daily variation of other weather variables. While performance of SDSM in reproducing current climate [*Wilby et al.*, 2006] shows reasonable skill, albeit with similar problems to EARWIG for the summer of 1976, its parameterization is heavily dependent on near-surface specific humidity, which may not be well reproduced by a climate model.

[21] Since the different downscaling methods yield such different PET change predictions, the values estimated in the current study by using RCM change factors through the medium of a WG are also compared with those of *Ekström et al.* [2007], who estimated PET values directly from the RCM output using the FAO modification of the Penman-Monteith equation. The estimates of summer PET change by *Ekström et al.* [2007] are unrealistically high, as noted in their paper, which attributed this to the tendency of this RCM to exaggerate drying, and to the enhancement of its uncertainty by the interaction of errors in other variables. By contrast, the current study uses temperature and rainfall distributional changes, and relationships with other variables derived from current climate, partially avoiding the feedback effects arising from direct use of the output of a dry climate model. This is clearly a problematic area, requiring further research beyond the scope of this study, including an improvement in climate model land surface hydrology.

[22] As an indication of the importance of correct PET calculation for hydrological impacts studies, the following sensitivity analysis has been performed. Taking a synthetic record (1000 years) of rainfall and PET representative of the Thames basin under current climate (with mean values of rainfall and PET shown in Table 5), the PET record was increased by a proportion, one season at a time, up to 100%, and changes in flow quantiles were calculated. For comparison, the rainfall record was increased in a similar way. For the ranges of PET and rainfall considered, the change in flow quantiles was directly proportional to the percentage change in PET or rainfall. Percentage change in flow quantiles are shown in Figure 7. As might be expected, increasing rainfall has a larger effect on the flow than increasing PET. It is interesting to note that apart from autumn and winter rainfall increase, the largest sensitivity is in the median flow rate. The high and low quantiles are to be expected to be more dependent on sequencing than on quantity. The effect of increasing summer PET is highest, and is approximately half that of increasing summer rainfall. The implication of this is that correct calculation of PET is essential in calculating future water resource availability.

4. Combination of Climate Predictions

[23] The discussion so far has dealt with ensemble outputs from several combinations of GCMs and RCMs. Experience in weather and climate forecasting [e.g., *Palmer*, 2000, *Min and Hense*, 2006, *Hagedorn et al.*, 2005] as well as in other sectors such as crop modeling [e.g., *Cantelaube and Terres*, 2005] indicate that better forecasting skill can be achieved by combining multiple models rather than relying upon a single model. Where observed data are available for model validation then the problem of model combination can be

Table 5. Current Mean Rainfall and Potential Evapotranspiration in the Thames Basin[a]

	Season			
	DJF	MAM	JJA	SON
Mean rainfall (mm d^{-1})	2.22	1.80	1.78	2.35
Mean PET (mm d^{-1})	0.51	1.88	2.97	1.10

[a]DJF, December–January–February; MAM, March–April–May; JJA, June–July–August; SON, September–October–November.

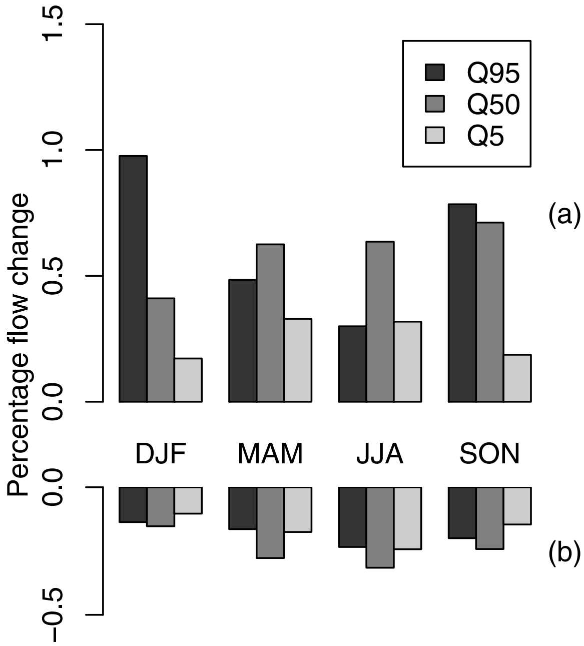

Figure 7. Percentage change in flow in the river Thames per percentage of increased (a) rainfall and (b) PET.

interpreted in the following Bayesian terms. The design and implementation of climate models is founded upon the representation of basic laws governing climate processes on Earth, but also require a number of subjective judgments on the part of modelers. For example, experts will have subjective beliefs about the appropriate parameter values for a given model or ensemble of models. This may translate into a prior distribution over the population of models. However, it is natural to test models with respect to how accurately they reproduce observed climate. This combination of subjective modeling judgments and conditioning upon observed data fits conveniently within the Bayesian framework whereby a posterior estimate of the uncertain quantities of interest is computed as an update of prior beliefs conditioned upon the observed data.

[24] However, several difficulties in application of Bayesian methods exist in the context of climate modeling [*Rougier*, 2007]. Observations of past climate are only of partial relevance to ensuring that a model adequately predicts future climate, since climate change will also depend on response to changing concentrations of atmospheric constituents. Furthermore, climate model experiments have not in the past lent themselves to maximizing the possible use of observed data, so rather simplified criteria have been used for conditioning, including various measures of model skill in reproducing current climate or trends, and of the mutual agreement of climate models in the ensemble used [*Giorgi and Mearns*, 2002; *Tebaldi et al.*, 2005; *Greene et al.*, 2006; *Furrer et al.*, 2007]. Most methods for combining projections, including the one adopted here, are based on implicit or explicit assumptions of independence between the different models, so that the errors in different models tend to cancel out. However, GCMs, even from different centers have been shown not be independent [*Tebaldi and Knutti*, 2007; *Jun et al.*, 2008a, 2008b] being based on the same theoretical or sometimes empirical assumptions and run at a similar resolution. RCMs will suffer from similar dependencies, and in addition, RCMs driven by the same GCM-derived boundary conditions will show dependency because of those boundary conditions. While these limitations are now well recognized, *Tebaldi and Knutti* [2007, p. 2060] remark that "no formal approach at quantifying this dependence has been worked out yet." Therefore, here, as in previous studies, we start with a method based upon a model independence assumption, but then go on to explore, more comprehensively than in previous studies, sensitivity of the predictive variable to model weights.

[25] Two relevant methods have been published, both providing Bayesian estimates of climate change based on GCM predictions, both conditioned on performance in predicting mean temperature over part-continental regions specified by *Giorgi and Francisco* [2000]. The method of *Greene et al.* [2006] weights models on their performance in matching historical trends of mean regional temperature throughout the 20th century and applies this weighting throughout future integrations. However, this method requires transient climate model output, and is unsuitable for application to stationary RCM time slice output. By contrast, the method of *Tebaldi et al.* [2005] and *Smith et al.* [2007], motivated by the study of *Giorgi and Mearns* [2002], applies to an ensemble of different realizations of a current and future time slice. This method weights models by two criteria: their ability to reproduce current mean regional temperature (or alternatively rainfall), and their agreement with the ensemble consensus estimate of the future temperature mean.

[26] The method of *Tebaldi et al.* [2005] has been adapted by *Fowler et al.* [2007a, 2007b], to combine RCM predictions for a division of the United Kingdom into nine smaller, climatologically coherent regions [*Wigley et al.*, 1984], with areas of the order of 10,000 km^2, and has been applied here for the Thames basin (10,000 km^2). As well as producing probability density functions for the change in rainfall and temperature, a set of jointly estimated parameters can be interpreted as weights for the members of the ensemble of climate models, which depend on each model's success in describing the current climate and to a lesser degree on the model's agreement with the consensus estimate of change.

[27] The Bayesian method combining projections from multiple climate models is fully described by *Tebaldi et al.* [2005]. Rather succinctly, it is assumed that n climate models are used to predict the change in a desired (scalar) climate attribute, such as mean regional temperature or precipitation, over a given season of the year. Each model i simulates the attribute under current (X_i) and future (Y_i) climate. We also observe the value of the attribute in the real world, X_0. The following statistical model is assumed for the data:

$$X_i = \mu + \eta_i, i = 0, \ldots, n$$
$$Y_i = \nu + \beta(X_i - \mu) + \xi_i/\sqrt{\theta}, i = 1, \ldots, n,$$

where μ is the real but uncertain climate attribute under current conditions, around which both observations and modeled quantities are distributed with a random error; ν is the corresponding uncertain mean of the climate attribute

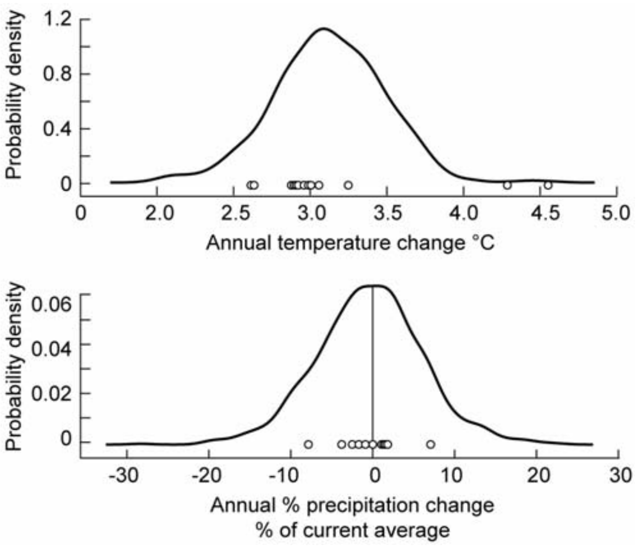

Figure 8. Probability distribution functions of the changes in annual mean temperature and precipitation in the Thames catchment. Circles show mean change predicted by individual models.

under future conditions; β represents correlation between current and future climate simulation for a given model, and is assumed here as common to all models; θ is a variance scaling parameter between current and future climate simulations, assumed here the same for all models, allowing for future projections to be less accurate than current simulations; and $\eta_i \sim N(0, \lambda_i^{-1})$, $\xi_i \sim N(0, \lambda_i^{-1})$ are Gaussian error terms, whose variances depend specifically on each of the different models.

[28] All the parameters in this formulation are considered random variables with prior distributions chosen to be as uninformative as possible. As a result of the Bayesian updating, the posterior mean of ν, the future value of the climate attribute, is, using standard techniques [*O'Hagan and Forster*, 2004], approximately $\Sigma\lambda_i[Y_i - \beta(X_i - \mu)]/\Sigma\lambda_i$, a weighted sum of the model predictions. The weights are given by the values λ_i, the inverse of the model-specific error variances, which are variables estimated in the Bayesian scheme, and have their own posterior probability distributions; an approximation to the posterior mean value of each λ_i is given by

$$E(\lambda_i|\{X_0, \cdots X_n, Y_1, \cdots Y_n\}) \approx \frac{a+1}{b + \frac{1}{2}\left\{(X_i - \tilde{\mu})^2 + \theta[Y_i - \tilde{\nu} - \beta(X_i - \tilde{\mu})]^2\right\}}, \quad (1)$$

where a and b are parameters of the Gamma prior distributions, and $\tilde{\mu}$ and $\tilde{\nu}$ are the posterior means of the current (μ) and future (ν) climate attribute under consideration [*Tebaldi et al.*, 2005]. *Tebaldi et al.* [2005] estimated the parameters of this model using a Markov chain Monte Carlo algorithm to predict mean temperature changes over 22 large regions covering the globe, from an ensemble of 9 GCM integrations. The method was extended by *Fowler et al.* [2007a] to apply to changes in either mean temperature or mean rainfall over the northwest region of the United Kingdom, based on an ensemble of six RCM model integrations from PRUDENCE. *Fowler et al.* [2007b] used the same method to produce probability distributions of change for all nine UK regions using 13 RCM integrations.

[29] The model combination studies discussed above, together with some alternative proposals in the literature have used ensembles of climate projections to produce probability distributions of global, regional or gridded mean temperature [*Tebaldi et al.*, 2005; *Greene et al.*, 2006; *Furrer et al.*, 2007; *Jun et al.*, 2008a, 2008b]. Such methods could be adapted to provide distributions of change in other climate variables. However, these approaches are less readily usable to provide input for WGs, which use a large number of climate variables to provide scenarios of local future climate i.e., for the WG described above: monthly or seasonal change factors for rainfall mean, variance, skewness and 24 h lagged autocorrelation, percentage of dry days, and temperature mean and variance. One approach to resolving this problem would be to generate (joint) predictive distributions for all of the variables used to adjust the WG to future climate predictions This is a challenging task that has not to date been achieved. An alternative approach is to fit the WG to each of the RCM outputs separately, and weight each WG output on the basis of the performance of the RCM [*Fowler et al.*, 2007a, 2007b]. This approach relies upon the assumption that the predictive variable used for weighting is also an appropriate predictive variable for the impact of interest. In the context of water resources assessment, the use of rainfall and temperature as predictive variables for weighting is appropriate.

[30] The method of *Fowler et al.* [2007b] was applied to the Thames catchment, for both mean annual temperature change, and mean annual rainfall change, using as input the 13 PRUDENCE RCM outputs listed in Table 3, for the 2080s, under the assumption of A2 emissions scenario. Figure 8 shows the probability distribution functions of these changes, together with the actual mean changes for each of the contributing RCMs. A median increase of mean temperature of some 3°C is predicted, while the median change in mean rainfall is close to zero. Distributions of weighting variables λ_i have been also calculated, weighting both on mean annual temperature change, and on mean annual rainfall change. The mean values for these distributions, together with the 90% posterior probability intervals, are given in Table 6.

[31] Figure 9 shows the cumulative distribution functions for abstraction availability in the Thames, using unweighted averaging of climate predictions, and also using weightings derived from the mean values in Table 6, with weighting on the basis of both temperature and precipitation. Either choice in predictive variable used for weighting is justified in that the winter abstraction availability is well correlated with both variables. The weighting scheme of *Fowler et al.* [2007b] in this case generates results that are close to a uniform weighting scheme, although where output derived from the contributing climate models show large variations, the weighting scheme may produce very different results. Thus, at first sight, it may seem unnecessary to use a complex weighting scheme to combine climate projections. However, while in this instance the weights are not far from uniform the estimate of associated uncertainty is of utmost relevance

Table 6. Mean Values for Weighting Functions, λ_i[a]

	Temperature	Precipitation
ARP.HAD	0.07 (0.01, 0.21)	0.13 (0.03, 0.30)
HADR.HAD	0.12 (0.02, 0.28)	0.02 (0.00, 0.05)
HIR.ECH	0.1 (0.02, 0.21)	0.08 (0.01, 0.18)
HIR.HAD	0.07 (0.01, 0.15)	0.09 (0.02, 0.22)
RCA.ECH	0.04 (0.01, 0.12)	0.02 (0.00, 0.06)
RCA.HAD	0.04 (0.01, 0.10)	0.06 (0.01, 0.15)
CHRM.HAD	0.06 (0.01, 0.15)	0.03 (0.01, 0.09)
CLM.HAD	0.04 (0.01, 0.11)	0.07 (0.01, 0.18)
REMO.HAD	0.07 (0.01, 0.18)	0.12 (0.02, 0.27)
PROMES.HAD	0.05 (0.01, 0.14)	0.08 (0.01, 0.20)
REGCM.HAD	0.08 (0.01, 0.19)	0.06 (0.01, 0.14)
RACMO.HAD	0.12 (0.02, 0.28)	0.12 (0.02, 0.27)
METNO.HAD	0.13 (0.03, 0.29)	0.14 (0.03, 0.30)

[a]The 90% probability intervals are given in parentheses.

to water resource planning decisions. Nonetheless, it should be noted in this context that care must be taken in interpreting the uncertainty ranges depicted by this analysis. It is recognized that the uncertainties arising from model choice have a dominant role in shaping the range of possible future outcomes [see *Tebaldi and Knutti*, 2007], and that uncertainties found here are limited by the narrow choice of climate models used, which are recognized as being unlikely to be independent. A broader range of climate models may approximate more closely the full range of uncertainty.

[32] In order to explore the sensitivity of the predicted water resource to the weighting of climate model outputs we have constructed the following sensitivity model. The expected values from the Bayesian analysis are written as a vector of model weights λ_i: $i = 1, \ldots, n$, calculated according to the posterior analysis including equation (1). Using the procedure of *Fowler et al.* [2007b] we have estimated the weights $\tilde{\lambda}_i$ with the 90% posterior probability intervals in Table 6, now denoted $[\tilde{\lambda}_i - \Delta_i^-, \tilde{\lambda}_i + \Delta_i^+]$. In our sensitivity test we consider the set of weights:

$$U(\alpha, \boldsymbol{\lambda}) = \left\{ \boldsymbol{\lambda} : \tilde{\lambda}_i - \alpha\Delta_i^- \leq \lambda_i \leq \tilde{\lambda}_i + \alpha\Delta_i^+, i = 1, \ldots, n, \sum_{i=1}^{n} \lambda_i = 1, \quad 0 \leq \lambda_i, \forall i \right\}, \ \alpha \geq 0.$$

In other words, we explore the effect of different weighting combinations, where the scale on which the weights are permitted to vary depends on the relative size of the Bayesian posterior probability interval. This uncertainty model does not make any assumptions about the distribution of the weights, beyond the limited information available from the analysis in the form of these posterior probability intervals, and the constraint that the model weights are nonnegative and sum to unity. The effect of different levels of uncertainty on the predicted abstraction availability is shown in Figure 10. For each value of α we plot the greatest and least values of the predictive probability of interest (the exceedance probability of the number of days of abstraction).

[33] The sensitivity analysis demonstrated here provides decision makers with a means of testing the robustness of water resource management decisions to the uncertainties in climate model weighting, by analyzing the range of performance of each option at increasing values of the parameter α [*Hipel and Ben-Haim*, 1999]. Different management options may be expected to perform in different ways at different levels of uncertainty. A robust option is thought of as one whose performance is acceptable even at high levels of uncertainty. The approach is analogous to the methods of *Dessai and Hulme* [2007] and *Groves and Lempert* [2007] in that it deals with a set of possible climate futures, but here the set-based approach is used to demonstrate the uncertainties in an underlying probabilistic model. Robust decision making can be supported by testing a range of alternative water resource management strategies in order to identify those that perform acceptably well under the uncertain range of possible futures. In this case, the uncertainties are encapsulated in the climate model weighting uncertainty, which also effectively explores other (but not all) aspects of uncertainty, including the unknown dependence between the different climate models in the ensemble.

5. Conclusions

[34] A practical approach for using ensembles of climate model outputs in a water resources assessment has been presented. While the example has been one of water resources assessment the techniques used are applicable

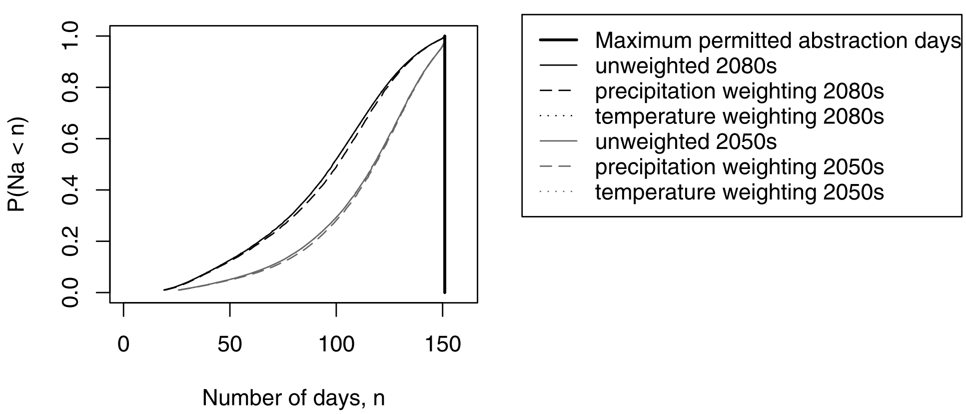

Figure 9. Combined prediction of abstraction availability: comparison of unweighted and weighted combinations, A2 scenario.

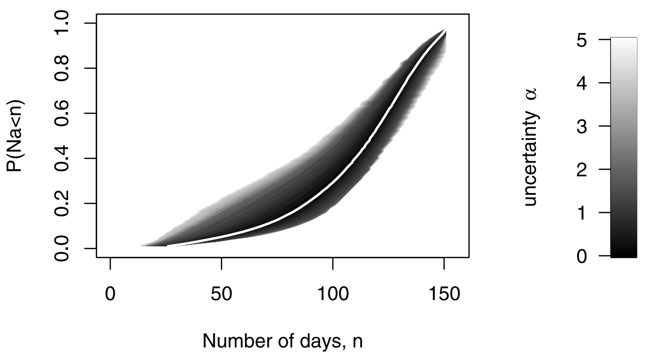

Figure 10. Effect of uncertainty in Bayesian model weights on predicted abstraction availability, 2050s.

to a host of climate impacts assessments that require high-resolution time series downscaled from climate model outputs.

[35] An ensemble of RCM outputs has been used to generate an estimate of the uncertainties in regional climate change, and a stochastic weather generator has been used to produce daily weather series for each of the RCM model outputs. These series have been input to a rainfall-runoff model of the Thames catchment. Comparison between this and other studies of the Thames shows that different downscaling methods, even from the same GCM, can produce very different results in terms of predicted changes in flow. It has been shown that the estimation of potential evapotranspiration for future climates is very significant in the assessment of water resource availability in the Thames catchment.

[36] A Bayesian approach has been used to generate catchment-specific weighting factors for each of the climate model outputs and hence a weighted combination of predictions. An extended sensitivity analysis has been used to explore the potential range of variation in the predictive quantity of interest, making use of the variance estimates obtained from the Bayesian model. The results demonstrate the prospect of severe reduction in water availability in the Thames catchment in a wide range of combinations of climate model projections. This provides the basis for decision making that is robust to model uncertainties. The approach to combining ensemble outputs of RCMs with weather generators and impacts models is generally applicable and is expected to become widely used with the increasing availability of ensemble model output, for example from the forthcoming UKCP09 climate scenarios.

[37] **Acknowledgments.** The PRUDENCE data were obtained from the project Web site http://prudence.dmi.dk. The authors wish to thank Richard Davis for suggesting the application, Rob Wilby for making available a version of CATCHMOD and for extensive discussions, Stephen Blenkinsop for preparing the PRUDENCE data, Ali Ford for adapting EARWIG to use the PRUDENCE outputs, and Marie Ekström for making available data on PET change. The authors would also like to thank the associate editor and four anonymous reviewers for their comments, which have improved the paper. The research was supported by the UK Engineering and Physical Sciences Research Council in project GR/S18052 "CRANIUM: Climate change risk analysis, new impact and uncertainty methods." H.J.F. was supported by NERC postdoctoral fellowship award (2006–2009) NE/D009588/1.

References

Allen, R. G., M. Smith, L. S. Pereira, and A. Perrier (1994), An update for the calculation of reference transpiration, *ICID Bull.*, *43*, 35–92.

Arnell, N. W. (2003), Relative effects of multi-decadal climatic variability and changes in the mean and variability of climate due to global warming: Future streamflows in Britain, *J. Hydrol.*, *270*, 195–213, doi:10.1016/S0022-1694(02)00288-3.

Blenkinsop, S., and H. J. Fowler (2007), Changes in drought frequency, severity and duration for the British Isles projected by the PRUDENCE regional climate models, *J. Hydrol.*, *342*, 50–71, doi:10.1016/j.jhydrol.2007.05.003.

Blenkinsop, S., P. D. Jones, S. R. Dorling, and T. J. Osborn (2008), Observed and modelled influence of atmospheric circulation on central England temperature extremes, *Int. J. Climatol.*, *29*, 1642–1660, doi:10.1002/joc.1807.

Cantelaube, P., and J.-M. Terres (2005), Seasonal weather forecasts for crop yield modeling in Europe, *Tellus, Ser. A*, *56*, 476–487.

Christensen, J. H., T. R. Carter, M. Rummukainen, and G. Amanatidis (2007), Evaluating the performance and utility of regional climate models in climate change research: Reducing uncertainties in climate change projections—The PRUDENCE approach, *Clim. Change*, *81*, suppl. 1, 1–6, doi:10.1007/s10584-006-9211-6.

Christensen, N. S., and D. P. Lettenmaier (2007), A multimodel ensemble approach to assessment of climate change impacts on the hydrology and water resources of the Colorado River Basin, *Hydrol. Earth Syst. Sci.*, *11*, 1417–1434.

Collins, M., B. B. B. Booth, G. R. Harris, J. M. Murphy, D. M. H. Sexton, and M. J. Webb (2006), Towards quantifying uncertainty in transient climate change, *Clim. Dyn.*, *27*, 127–147, doi:10.1007/s00382-006-0121-0.

Cowpertwait, P. S. P. (1991), Further developments of the Neyman-Scott clustered point process for modeling rainfall, *Water Resour. Res.*, *27*, 1431–1438, doi:10.1029/91WR00479.

Davis, R. J. (2001), The effects of climate change on river flows in the Thames region, *Hydrol. Hydrom. Rep. 00/04*, Environ. Agency, Reading, U. K.

Dessai, S., and M. Hulme (2007), Assessing the robustness of adaptation decisions to climate change uncertainties: A case study on water resources management in the east of England, *Global Environ. Change*, *17*, 59–72, doi:10.1016/j.gloenvcha.2006.11.005.

Ekström, M., P. D. Jones, H. J. Fowler, G. Lenderink, T. A. Buishand, and D. Conway (2007), Regional climate model data used within the SWURVE project. 1: Projected changes in seasonal patterns and estimation of PE, *Hydrol. Earth Syst. Sci.*, *11*, 1069–1083.

Environment Agency (2004), Thames corridor abstraction management strategy, Environ. Agency, Reading, U. K.

Faulkner, D. (1999), *Flood Estimation Handbook*, vol. 2, *Rainfall Frequency Estimation*, Inst. of Hydrol., Wallingford, U. K.

Fowler, H. J., S. Blenkinsop, and C. Tebaldi (2007a), Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling, *Int. J. Climatol.*, *27*, 1547–1578, doi:10.1002/joc.1556.

Fowler, H. J., C. Tebaldi, and S. Blenkinsop (2007b), Probabilistic estimates of climate change impacts on flows in the River Eden, Cumbria, paper presented at 10th National Hydrology Symposium: Sustainable Hydrology for the 21st Century, Br. Hydrol. Soc., Exeter, U. K.

Furrer, R., R. Knutti, S. R. Sain, D. W. Nychka, and G. A. Meehl (2007), Spatial patterns of probabilistic temperature change projections from a multivariate Bayesian analysis, *Geophys. Res. Lett.*, *34*, L06711, doi:10.1029/2006GL027754.

Giorgi, F., and R. Francisco (2000), Evaluating uncertainties in the prediction of regional climate change, *Geophys. Res. Lett.*, *27*, 1295–1298, doi:10.1029/1999GL011016.

Giorgi, F., and L. O. Mearns (2002), Calculation of average, uncertainty range and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging" (REA) method, *J. Clim.*, *15*, 1141–1158, doi:10.1175/1520-0442(2002)015<1141:COAURA>2.0.CO;2.

Greene, A. M., L. Goddard, and U. Lall (2006), Probabilistic multimodel regional temperature change projections, *J. Clim.*, *19*, 4326–4343, doi:10.1175/JCLI3864.1.

Groves, D. G., and R. J. Lempert (2007), A new analytic method for finding policy-relevant scenarios, *Global Environ. Change*, *17*, 73–85, doi:10.1016/j.gloenvcha.2006.11.006.

Hagedorn, R., F. J. Doblas-Reyes, and T. N. Palmer (2005), The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept, *Tellus, Ser. A*, *57*, 219–233.

Hagemann, S., and D. Jacob (2007), Gradient in the climate change signal of European discharge predicted by a multi-model ensemble, *Clim. Change*, *81*, 309–327, doi:10.1007/s10584-006-9225-0.

Hall, J. (2007), Probabilistic climate scenarios may misrepresent uncertainty and lead to bad adaptation decisions, *Hydrol. Processes*, *21*, 1127–1129, doi:10.1002/hyp.6573.

Hipel, K. W., and Y. Ben-Haim (1999), Decision making in an uncertain world: Information-gap modeling in water resources management, *IEEE Trans. Syst. Man Cybern. C*, *29*, 506–517, doi:10.1109/5326.798765.

Hough, M. N., and R. J. A. Jones (1997), The United Kingdom Meteorological Office rainfall and evaporation calculation system: MORECS version 2.0—An overview, *Hydrol. Earth Syst. Sci.*, *1*, 227–239.

Hulme, M., et al. (2002), Climate change scenarios for the United Kingdom: The UKCIP02 scientific report, Tyndall Cent. for Clim. Change Res., Sch. of Environ. Sci., Univ. of East Anglia, Norwich, U. K.

Jun, M., R. Knutti, and D. W. Nychka (2008a), Spatial analysis to quantify numerical model bias and dependence: How many climate models are there?, *J. Am. Stat. Assoc.*, *103*, 934–947.

Jun, M., R. Knutti, and D. W. Nychka (2008b), Local eigenvalue analysis of CMIP climate model errors, *Tellus, Ser. A*, *60*, 992–1000.

Kilsby, C. G., P. D. Jones, A. Burton, A. C. Ford, H. J. Fowler, C. Harpham, P. James, A. Smith, and R. L. Wilby (2007), A daily weather generator for use in climate change studies, *Environ. Modell. Software*, *22*, 1705–1719, doi:10.1016/j.envsoft.2007.02.005.

Min, S.-K., and A. Hense (2006), A Bayesian approach to climate model evaluation and multi-model averaging with an application to global mean surface temperatures from IPCC AR4 coupled climate models, *Geophys. Res. Lett.*, *33*, L08708, doi:10.1029/2006GL025779.

Moberg, A., and P. D. Jones (2004), Regional climate model simulations of daily maximum and minimum near-surface temperatures across Europe compared with observed station data 1961–1990, *Clim. Dyn.*, *23*, 695–715, doi:10.1007/s00382-004-0464-3.

Murphy, J. M., B. B. B. Booth, M. Collins, G. R. Harris, D. M. H. Sexton, and M. J. Webb (2007), A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles, *Philos. Trans. R. Soc. A*, *365*, 1993–2028, doi:10.1098/rsta.2007.2077.

Murphy, J. M., et al. (2009), UK climate projections science report: Climate change projections, Hadley Cent., Met Off., Exeter, U. K.

Nakicenovic, N., et al. (2000), *IPCC Special Report on Emissions Scenarios*, 599 pp., Cambridge Univ. Press, Cambridge, N. Y.

Nawaz, N. R., and A. J. Adeloye (2006), Monte Carlo assessment of sampling uncertainty of climate change impacts on water resources yield in Yorkshire, England, *Clim. Change*, *78*, 257–292, doi:10.1007/s10584-005-9043-9.

New, M., A. Lopez, S. Dessai, and R. Wilby (2007), Challenges in using probabilistic climate change information for impact assessments: An example from the water sector, *Philos. Trans. R. Soc. A*, *365*, 2117–2131, doi:10.1098/rsta.2007.2080.

O'Hagan, A., and J. Forster (2004), *Kendall's Advanced Theory of Statistics*, vol. 2b, *Bayesian Inference*, Edward Arnold, London.

Palmer, T. N. (2000), Predicting uncertainty in forecasts of weather and climate, *Rep. Prog. Phys.*, *63*, 71–116, doi:10.1088/0034-4885/63/2/201.

Perry, M., and D. Hollis (2005), The generation of monthly gridded datasets for a range of climatic variables over the UK, *Int. J. Climatol.*, *25*, 1041–1054, doi:10.1002/joc.1161.

Prudhomme, C., N. Reynard, and S. Crooks (2002), Downscaling of global climate models for flood frequency analysis: Where are we now?, *Hydrol. Processes*, *16*, 1137–1150, doi:10.1002/hyp.1054.

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

Smith, R. L., C. Tebaldi, D. Nychka, and L. Mearns (2007), Bayesian modeling of uncertainty in ensembles of climate models, *J. Am. Stat. Assoc.*, *104*, 97–116.

Tebaldi, C., and R. Knutti (2007), The use of the multi-model ensemble in probabilistic climate projections, *Philos. Trans. R. Soc. A*, *365*, 2053, doi:10.1098/rsta.2007.2076.

Tebaldi, C., D. Smith, D. Nychka, and L. Mearns (2005), Quantifying uncertainty in projections of regional climate change: A Bayesian approach to the analysis of multi-model ensembles, *J. Clim.*, *18*, 1524–1540, doi:10.1175/JCLI3363.1.

Wigley, T. M. L., J. M. Lough, and P. D. Jones (1984), Spatial patterns of precipitation in England and Wales and a revised, homogenous England and Wales precipitation series, *J. Climatol.*, *4*, 1–25, doi:10.1002/joc.3370040102.

Wilby, R. L. (2005), Uncertainty in water resource model parameters used for climate change impact assessment, *Hydrol. Processes*, *19*, 3201–3219, doi:10.1002/hyp.5819.

Wilby, R. L., and I. Harris (2006), A framework for assessing uncertainties in climate change impacts: Low flow scenarios for the River Thames, UK, *Water Resour. Res.*, *42*, W02419, doi:10.1029/2005WR004065.

Wilby, R. L., B. Greenfield, and C. Glenny (1994), A coupled synoptic-hydrological model for climate change impact assessment, *J. Hydrol.*, *153*, 265–290, doi:10.1016/0022-1694(94)90195-3.

Wilby, R. L., C. W. Dawson, and W. M. Barrow (2002), SDSM—A decision support tool for the assessment of regional climate change impacts, *Environ. Modell. Software*, *17*, 147–159, doi:10.1016/S1364-8152(01)00060-3.

Wilby, R. L., P. G. Whitehead, A. J. Wade, D. Butterfield, R. J. Davis, and G. Watts (2006), Integrated modelling of climate change impacts on water resources and quality in a lowland catchment: River Kennet, UK, *J. Hydrol.*, *330*, 204–220, doi:10.1016/j.jhydrol.2006.04.033.

Wilks, D. S., and R. L. Wilby (1999), The weather generation game: A review of stochastic weather models, *Prog. Phys. Geogr.*, *23*, 329–357.

H. J. Fowler, J. W. Hall, C. G. Kilsby, and L. J. Manning, School of Engineering and Geosciences, Newcastle University, Newcastle upon Tyne NE1 7RU, UK. (lucy.manning@newcastle.ac.uk)

C. Tebaldi, Climate Central, One Palmer Square, Suite 330, Princeton, NJ 08542, USA.