PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Challenges in using probabilistic climate change information for impact assessments: an example from the water sector

Mark New, Ana Lopez, Suraje Dessai and Rob Wilby

Phil. Trans. R. Soc. A 2007 **365**, 2117-2131
doi: 10.1098/rsta.2007.2080

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A

Phil. Trans. R. Soc. A (2007) **365**, 2117–2131
doi:10.1098/rsta.2007.2080
Published online 14 June 2007

Challenges in using probabilistic climate change information for impact assessments: an example from the water sector

By Mark New[1,*], Ana Lopez[1], Suraje Dessai[2] and Rob Wilby[3]

[1]*Centre for the Environment and Tyndall Centre, Oxford University, Oxford OX1 3QY, UK*
[2]*Tyndall Centre, University of East Anglia, Norwich NR4 7TJ, UK*
[3]*Environment Agency of England and Wales, Cardiff CF24 0TP, UK*

Climate change impacts and adaptation assessments have traditionally adopted a scenario-based approach, which precludes an assessment of the relative risks of particular adaptation options. Probabilistic impact assessments, especially if based on a thorough analysis of the uncertainty in an impact forecast system, enable adoption of a risk-based assessment framework. However, probabilistic impacts information is conditional and will change over time. We explore the implications of a probabilistic end-to-end risk-based framework for climate impacts assessment, using the example of water resources in the Thames River, UK. We show that a probabilistic approach provides more informative results that enable the potential risk of impacts to be quantified, but that details of the risks are dependent on the approach used in the analysis.

Keywords: climate change; impacts; uncertainties; probabilities; water resources; ensembles

1. Introduction

Climate change impact assessments have to date relied predominantly on the scenario-based approach (Carter *et al.* 2001; Mearns *et al.* 2001). It has long been recognized that any one scenario represents a single trajectory through the cascade of uncertainty: emissions → concentrations → regional climate response → local climate response → impact (with or without feedbacks between each component of the cascade, e.g. New & Hulme 2000; IPCC 2001). The use of one or more scenarios, while useful for exploring potential climate change impacts, presents difficulties when adaptation decisions have to be made. Scenarios typically have no associated likelihood, so decision-makers faced with alternative scenarios cannot assess the relative risk of particular adaptations; the tendency may then be to choose a response to a middle of the road scenario or more conservatively, a strategy that is robust in the face of all available scenario-based information. Even a robust strategy may be difficult to implement if the decision-maker is concerned about impacts that fall outside the range suggested by the scenarios at hand.

* Author for correspondence (mark.new@ouce.ox.ac.uk).

One contribution of 13 to a Theme Issue 'Ensembles and probabilities: a new era in the prediction of climate change'.

Probability distributions of climate change impacts allow us to move to a risk-based impact and adaptation decision-making framework (Pittock *et al.* 2001). However, even for global-scale metrics such as climate sensitivity or the likelihood of exceeding a given 'dangerous' global temperature threshold, a unique probability distribution is impossible to derive due to the imprecise information available, scientific and modelling uncertainties, and different statistical estimation approaches (Hall 2007; Hall *et al.* 2007; Rougier 2007).

Although there have been previous attempts to assess local impacts within a probabilistic framework, these studies have typically scaled one or a few GCM responses by probabilities derived from a simple climate model (e.g. Jones 2000; New & Hulme 2000; Prudhomme *et al.* 2003), or have involved an assessment of the relative size of climate model and impacts model uncertainties (e.g. Aggarwal & Mall 2002; Wilby & Harris 2006; Graham *et al.* 2007), rather than a full end-to-end uncertainty analysis. Methods for addressing uncertainty in simulation models are well developed in many natural science fields, most notably hydrology (Freer *et al.* 1996; Beven 2000; Beven & Freer 2001), but relatively few climate change impact studies have drawn on these approaches (Araújo & New 2007). There have also been a number of assessments of regional scale uncertainty in climate change scenarios arising from both GCMs and regional climate models (RCMs) and/or statistical downscaling techniques (e.g. Tebaldi *et al.* 2005; Feng & Fu 2006; Frei *et al.* 2006; Haylock *et al.* 2006; Goodess *et al.* in press) and some attempts to link multiple GCM-downscaling combinations (Benestad 2004; Jasper *et al.* 2004; Pryor *et al.* 2005, 2006; Salathe 2005; Chen *et al.* 2006; Graham *et al.* 2007). But linking all these aspects of uncertainty together to address combined climate model and impacts model uncertainty in an end-to-end probabilistic framework has been fundamentally limited by a lack of sufficiently comprehensive uncertainty analyses of GCMs, which ultimately drive the impacts assessment process (Fowler *et al.* in press).

The large-ensemble GCM-modelling efforts described in this issue (Murphy *et al.* 2007) and elsewhere (Murphy *et al.* 2004; Stainforth *et al.* 2005) offer the opportunity for a 'probabilistic' approach to assess regional and local climate change impacts. Large ensemble GCM simulations, using hundreds to many tens of thousands of GCMs, potentially provide richer regional detail than multiple sampling of a few GCM patterns, as different climate forcings and initial conditions (IC) are propagated through alternative physics to a larger number of model-specific regional responses (e.g. Harris *et al.* 2006); the range in both global and regional responses from large perturbed-physics ensembles have been wider than those produced through analysis of model runs available from the global climate modelling community, the so-called 'ensembles of opportunity'. However, probabilistic climate prediction is a double-edged sword. While undoubtedly providing more information, the regional information arising from large ensemble GCM modelling remains conditional and will suffer from the same lack of uniqueness as distributions for global metrics.

In this paper, we explore the implications of this new generation of probabilistic climate information for end-to-end uncertainty analysis in impacts modelling and assessment. Our focus is at the regional to local scale, where local authorities, environmental agencies, business and other players may need to make decisions on climate change adaptation. We present the first example of how climate data from the climate*prediction*.net project can be used to generate probabilistic information that incorporates *both* climate model and impact model uncertainty, focusing on the Thames River in the UK. We first describe the experimental set-up, including the climateprediction.net data, the hydrological model (CATCHMOD) that we use and the approach to downscale the climate model outputs to the spatial scale required by CATCHMOD. We then describe the resultant probabilistic projections of future flow statistics in the Thames. We conclude the paper with a discussion of the main points arising from this research.

2. Methods and data

We use the initial results from the climateprediction.net experiment described in detail by Stainforth *et al.* (2005). The data from the experiment represent 2700 individual simulations with the HadSM3 climate model; each simulation comprises three 15-year periods: a calibration phase, followed by a 15 year $1\times CO_2$ 'control' simulation, and a $2\times CO_2$ simulation, in which the model moves towards an equilibrium response to $2\times CO_2$. Within this subset of the full first experiment, seven physics parameter values are perturbed and there are 449 unique combinations of perturbations. For most perturbations, there is more than one simulation, with each simulation differing only in IC. The total number of simulations in the 449 IC ensembles adds up to 2700 simulations in the 'grand ensemble'. The ensemble is therefore large, but limited in a number of ways: it comprises a sampling of only some of the uncertain physics parameters in the Hadley Centre climate model; it only samples from a single 'parent' model structure, ignoring uncertainties arising from alternative GCM model structures; it is a $2\times CO_2$ sensitivity experiment, without a full ocean model, rather than a transient experiment with a comprehensive atmosphere–ocean model such as those contributing to the last IPCC report.

Seasonal means from the last 8 years of the control and $2\times CO_2$ runs, and only for a limited number of variables, have been returned by client machines for archival in climateprediction.net data servers; we use precipitation, temperature and cloud fraction data to calculate future daily precipitation and potential evaporation to input into our hydrological model.

Many ensemble members have not reached equilibrium at the end of the $2\times CO_2$ phase. We therefore scale the $2\times CO_2$ 8-year mean responses for each variable by the ratio of global mean temperature for this period to the global mean equilibrium temperature change, estimated using the approach of Stainforth *et al.* (2005).

We use CATCHMOD to simulate daily discharge in the Thames at Teddington, London. CATCHMOD is a rainfall-runoff model used by the Environment Agency (EA) of England and Wales for water resource planning and abstraction licence allocation, and is described in detail by Wilby *et al.* (1994). It uses daily rainfall (PPT) and potential evapotranspiration (PET) data for input at sub-catchments represented in the model. This requires downscaling of the coarser resolution seasonal mean GCM data. As the archived GCM data do not support either dynamical or statistical downscaling, we use a simple *change factor* (CF) downscaling approach to produce input for CATCHMOD. For both PPT and PET, we compute a factor by which the variable will change in the future ($2\times CO_2$) compared to the present day ($1\times CO_2$) for each model run; these CFs are then used to perturb the observed daily climate data used to run CATCHMOD for present day simulations. For precipitation, the CF is the per cent change in seasonal precipitation between $1\times CO_2$ and $2\times CO_2$ periods for the GCM grid box that covers the Thames catchment. The seasonal CFs are linearly interpolated to monthly CFs and applied to the daily observed precipitation data.

The procedure for PET is more complicated as this variable is not available as direct climateprediction.net output; available model outputs of relevance to PET are temperature and cloud cover. We first estimate mean monthly PET for the present day using observed data (temperature, vapour pressure, net radiation and wind speed) with the Penman (1948) formulation. We next calculate CFs for temperature and cloud from the GCM data, which are then used to perturb the observed temperature, vapour pressure and radiation inputs to the PET calculation; the ratio of present day to perturbed Penman PET is then used as a CF to perturb the observed PET daily time series.

We note that our use of CFs forces the future time series to have the same temporal structure as the present day, and that any changes in variance simply reflect a scaling of the observed series (Diaz-Nieto & Wilby 2005). In addition, use of an 8-year average to characterize both $1\times CO_2$ and $2\times CO_2$ mean climate implies that natural variability will contribute more to the resulting CFs than in many previous impacts assessments, where usually differences of 30-year averages are considered. The influence of natural variability is reduced somewhat by averaging across IC-members, but the number of members in each IC ensemble varies from one to eight, and thus natural variability is a varying unknown for each CF.

CATCHMOD was set up with three ‘subcatchments’, each representing the area of the catchment with a similar hydrological runoff response: urban areas, clay geology and chalk geology (Wilby & Harris 2006). For each subcatchment, five parameters for CATCHMOD are determined through calibration against observed discharge. For our research, we explore the effects of uncertainty in these parameters by running CATCHMOD with 100 different combinations of parameter values, all of which produce calibration results within predefined goodness of fit limits (Wilby & Harris 2006). The underlying rationale to exploration of parameter uncertainty is similar to the climateprediction.net project; however, unlike climateprediction.net, the set of parameters values used for CATCHMOD is preselected by evaluation against observed discharge.

3. Results

(a) *Climate change information*

The simulated $1\times CO_2$ and $2\times CO_2$ precipitation and temperature at the GCM grid box covering the Thames basin are shown in figure 1. Temperature shows a similar range and distribution to the global equilibrium temperature results (Stainforth *et al.* 2005), as might be expected from a mid-latitude location. Rainfall changes in winter are almost all positive, and range up to a 50% increase compared to the control simulations. For autumn and spring, both increases and decreases in precipitation are simulated, while in summer nearly all models simulate reduced precipitation; in some instances, the reduction is as much as 80%. Cloud cover changes correlate closely to changes in precipitation (not shown).

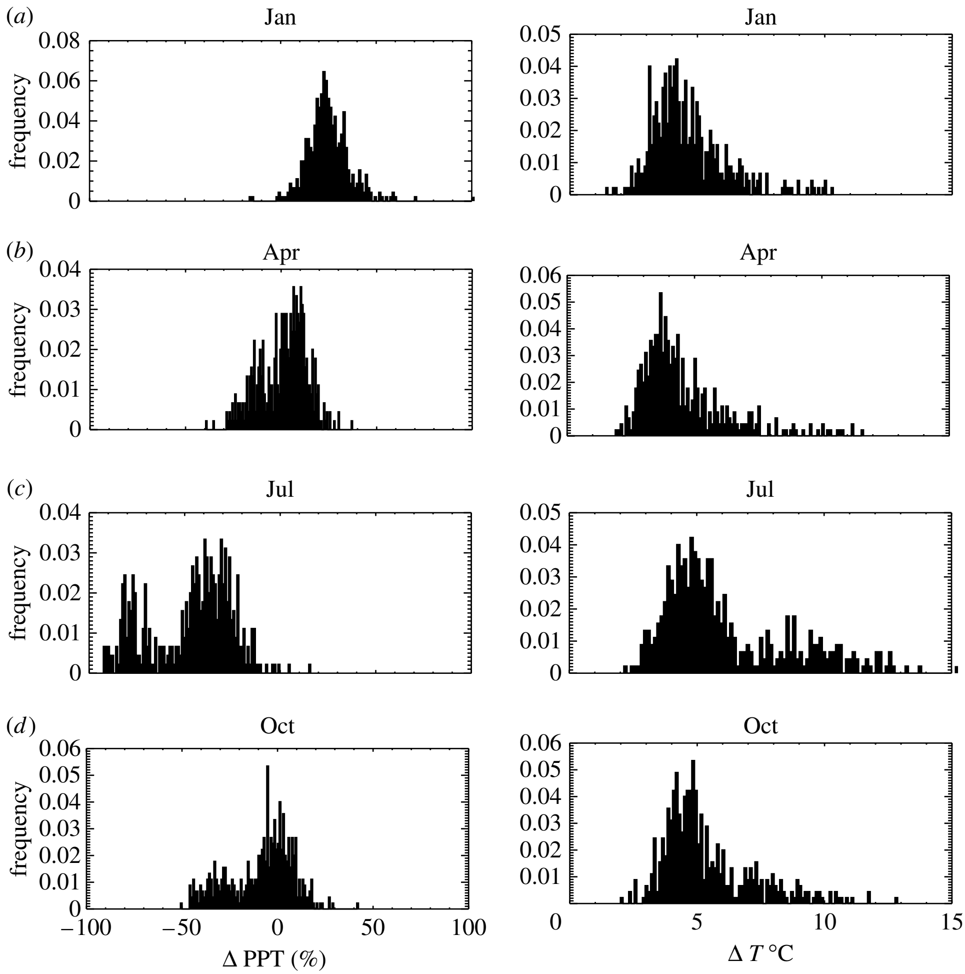

Figure 1. Simulated $1\times CO_2$ and $2\times CO_2$ climate data (precipitation change on the left and temperature change on the right) over the Thames from the climateprediction.net experiment.

In addition to the wide ranges of predicted mean changes in climate, precipitation change shows a bimodal distribution in spring, summer and autumn; this bimodality occurs over all UK and Ireland and adjacent ocean grid boxes, so appears to be a regional characteristic. The bimodality is particularly strong in summer and is not related to any individual parameter perturbation. A clearer understanding of reasons for this is difficult to come by, due to the limited set of model diagnostics that are archived. There is evidence that HadSM3 can become locked into different climate regimes over SW France, due to soil-moisture feedbacks (Clark *et al.* 2006); in some simulations, soil moisture reduces sufficiently to produce persistent surface heating. This would then affect regional circulation patterns, which may in turn affect precipitation. There also appears to be a relationship with the mean pressure gradient over the North Atlantic, since models with a high gradient under $2\times CO_2$ have lower autumn rainfall. This is consistent with an observed link between UK summer rainfall and the

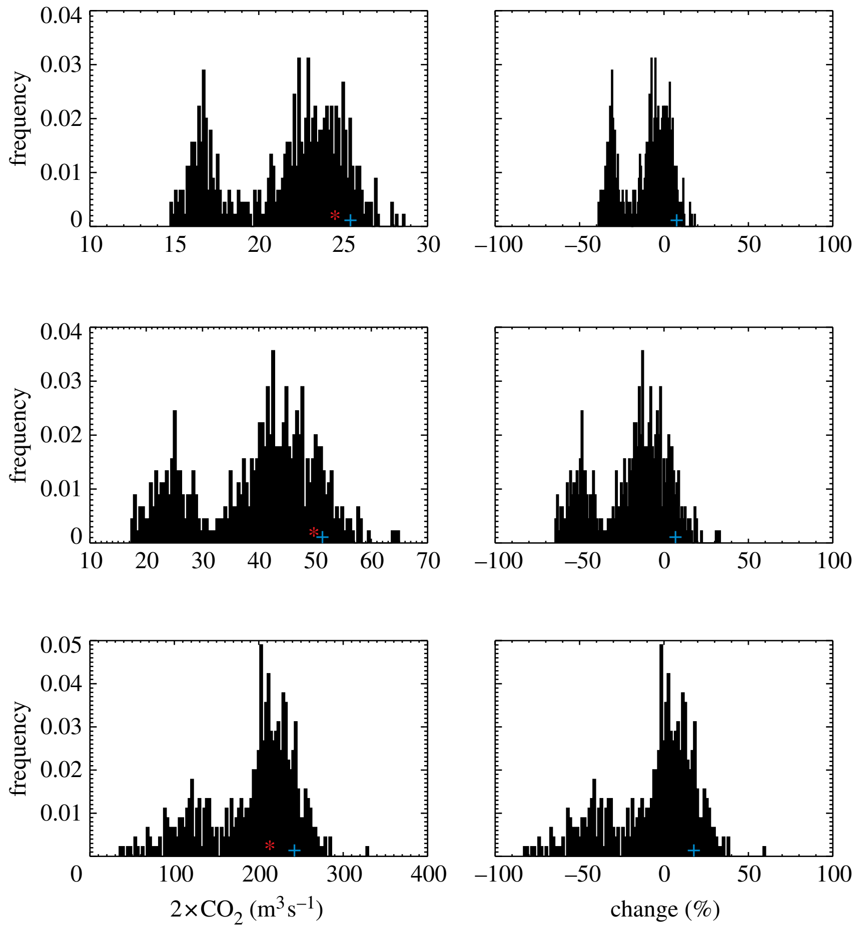

Figure 2. Changes in CATCHMOD simulated low (Q95), average (Q50) and high (Q05) flow statistics due to changes in precipitation and PET downscaled from the climateprediction.net ensemble. Q95 is the daily flow exceeded 95% of the time (low flows); Q50 is the median daily flow; Q05 is the daily flow exceeded 5% of the time (high flows). The red star shows the results when CATCHMOD is run with unperturbed present day climate data (1961–1990); blue symbol shows results for the standard version of HadSM3 climate model used in climateprediction.net.

North Atlantic Oscillation in the preceding winter (Wilby 2001). The available model diagnostics do not allow us to ascertain whether the pressure gradient-rainfall relationship is linked in any way to soil-moisture feedbacks.

(b) Simulated flow

We first describe how the downscaled climateprediction.net data described above propagate through the 'standard' CATCHMOD version (i.e. the version with a single set of parameter values, as used by the EA). For each of the 449 simulations with CATCHMOD, we calculate $1\times CO_2$ and $2\times CO_2$ flow percentiles as follows:

— Q05, the daily flow exceeded 5% of the time, which represents high flows,
— Q50, the median daily flow, and

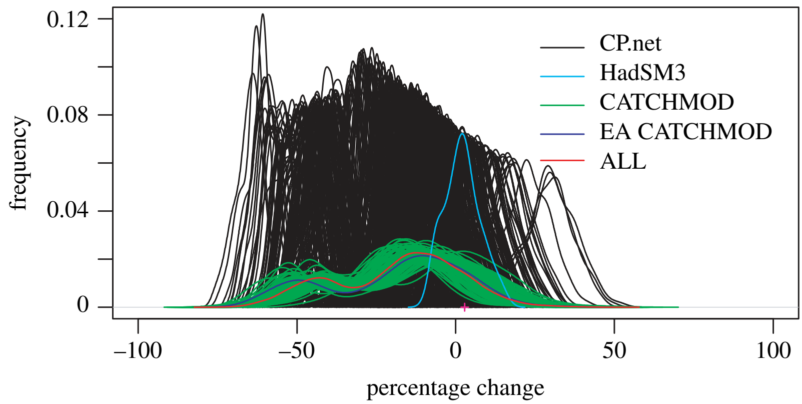

Figure 3. Changes in CATCHMOD simulated Q50 when uncertainties in CATCHMOD parameters are combined with the climateprediction.net ensemble. Each black curve is a smoothed frequency histogram obtained by combining one climateprediction.net IC ensemble with 100 CATCHMOD model versions. Green curves show the response of each CATCHMOD version combined with all climateprediction.net results. The red curve is the frequency distribution from all possible climateprediction.net–CATCHMOD combinations. For reference, the results from (i) the standard HadSM3 model with all CATCHMOD versions (light blue) and (ii) EA CATCHMOD with all climateprediction.net ICs (dark blue) are also shown. The red cross shows the result of the singular combination of the standard HadSM3 and EA CATCHMOD.

— Q95, a low-flow index corresponding to the daily flow exceeded on 95% of days, commonly used for resources assessment in catchment abstraction management plans by the EA.

The distribution of these percentiles across the 449 climateprediction.net ICs are shown in figure 2. For low and median flows, most realizations produce a decrease in the future. Of particular note is that most simulations result in reduced flows when compared with the standard atmospheric model (blue cross in figure 2), which was used, albeit coupled to a full ocean model, to generate the current set of UK climate change scenarios (Hulme *et al.* 2002). This illustrates a potential limitation of a scenario-based approach to impacts assessment; in this case, a single projection using the standard model provides a rather high estimate of future water resource availability when compared with other parameter combinations.

The bimodal distribution in precipitation produces either a second mode (Q95 and Q50) or negative skew (Q05) in the flow statistics. For high flows, while the proportion of simulations showing increases and decreases are roughly equal, the skewed distribution means that there are a relatively large number of cases where high flows are reduced by more than 40%, but all increases (bar one) are less than +40%.

We next consider the changes in simulated flow arising from both climateprediction.net and CATCHMOD parameter uncertainty. Here, we calculate flow statistics for 44 900 simulations with CATCHMOD, each simulation a unique combination of one of the 449 climateprediction.net IC outputs and one of the 100 CATCHMOD parameterizations (figure 3). If the standard HadSM3 model projections are run through all versions of CATCHMOD (light blue curve in figure 3), the range of responses in Q50 is −15 to +20%; similar ranges, with a

Table 1. Frequency of future monthly flows in climateprediction.net–CATCHMOD ensemble *below* low-flow thresholds identified in the present-day (1961–1990) simulations: (i) the lowest flows between 1961–1990 (LMMF 61–90) and (ii) the 10th percentile of monthly mean flow (MMF10).

month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
LMMF 61–90	0.039	0.067	0.033	0.034	0.039	0.056	0.071	0.075	0.311	0.195	0.039	0.065
MMF10 61–90	0.231	0.153	0.130	0.139	0.151	0.217	0.310	0.330	0.380	0.283	0.203	0.306

different central value, arise from combining any one climateprediction.net IC with the 100 CATCHMOD versions (black curves in figure 3). A similar result arises for Q05 and Q95 (not shown). Thus, the wide spread of climateprediction.net outputs dominate the spread in simulated changes, with different versions of CATCHMOD modulating the climateprediction.net signal. Nonetheless, if one compares the range of changes in Q05, Q50 and Q95 when using only the standard CATCHMOD to those using the full ensemble, CATCHMOD parameter uncertainty adds an additional 23% to the range for Q50, 16% for Q05 and 35% for Q95; thus low flows are most sensitive to hydrological model uncertainty.

(c) *Implications for water resource planning*

The simulated flows described above provide important information on the spread of plausible future natural flow levels in the Thames, and therefore an indication of possible future change in raw water availability. To illustrate this, we identify the lowest mean monthly flow (LMMF) and the 10th percentile for mean monthly flow (MMF10) in the 1961–1990 period; we then calculate the frequency with which monthly flow in the $2\times CO_2$ in the full climateprediction.net–CATCHMOD ensemble does not reach these levels (table 1). For reference, flows lower than LMMF would have a present-day frequency no higher than 0.033 (return period of 30 years); flows lower than MMF10 have a present-day frequency of 0.10 (10-year return period).

The lowest monthly flows in 1961–1990 occur in 1976 (January–August) and 1974 (September–December); 1976 is well known as a year with the most extensive drought conditions over southern England in recent years, and severe water shortages over most areas of the UK (Jones *et al.* 2006). It is used by some water utilities as a worst-case scenario for future resource planning, especially in southern England. In the Thames, summer 1976 flows are thought to be the lowest since 1865, at just 20% of the 1961–1990 average discharge (Jones *et al.* 2006). In the $2\times CO_2$ ensemble, the frequency of flows lower than LLMF ranges between 3% (similar to today) and 30%, depending on the month, with the highest frequencies occurring in late summer and autumn (a reflection of the reduced summer rainfall across most of the climateprediction.net ensemble). For MMF10, the frequency ranges from 13% in March (similar to present day) to over 30% for late summer and autumn, more than a threefold increase.

While these results provide information about the change in frequency of stressful water resource situations, the use of probabilistic climate data in a planning context requires more than consideration of GCM and hydrological model uncertainties addressed here. Future demand is also subject to considerable ambiguity, mainly because of uncertainty about changes in regional population, housing stock and industrial demands, but also due to changed (probably increased) *per capita* water use in a warmer climate. For example, current projections indicate that about 200 000 new households will form each year out to 2026, of which 60% will be in the south of England (EA 2007). This implies a 10–15% increase in reservoir capacity to meet rising water demand, at a cost of £3 billion. It is also envisaged that measures will be taken to improve water efficiency of new homes as well as the current housing stock. From April 2007, all publicly funded housing will have to be built to the Level 3 standard of the Code for Sustainable Homes, which means no more than 105 l per person per day, compared with the current UK average of 150 l (EA 2007). Making existing homes more water efficient could help meet approximately 40% of the future demand arising from new communities in the region, but considerable uncertainty exists as to the extent to which this can be achieved.

A thorough assessment of the implications of the probabilistic climate data for the Thames would therefore require simulations with a model representing the full water-resource system for the river, with the flexibility to include uncertainty in future demand and possible new abstraction and storage schemes.

We illustrate the type of information that can potentially be provided in such a water resource assessment using the trigger storage levels for reservoirs and flows set by the Environment Agency for the discharge in the River Thames at Teddington. These operating rules set out the demand management measurements that follow from progressively lower reservoir storage levels and river flows in the lower Thames. Under critical water storage conditions that vary through the year, the four trigger levels in the Thames are 800, 600, 400 and 300 Ml d^{-1}. Thus in January, when there remains a good chance of further rainfall to replenish reservoirs before demand peaks in late summer, reservoir storage must drop below 63 000 Ml for the Level 4 threshold of 300 Ml to be reached; in August, when there is little likelihood of replenishment, Level 4 is reached at a much higher storage level of 125 000 Ml. These thresholds invoke water saving publicity campaigns (Level 1), sprinkler bans and voluntary restrictions of inessential water use (Level 2), banning inessential water use and reduced pressure in the distribution system (Level 3) and finally major cuts of supply on a rota basis and use of standpipes (Level 4), respectively.

We first consider the changes in frequency of these thresholds being reached when the outputs from the simulations are used in their ‘unprocessed’ frequency distribution (i.e. without any post-processing to account, for example, for the uneven sampling of climateprediction.net parameter space). For July, the flow thresholds were not met in the 1961–1990 simulations some 1.5% of the time for the Level 4 trigger and 3.8% of the time for the more lenient Level 1 threshold (table 2; figure 4). When the frequency output from the $2\times CO_2$ ensemble is analysed the Level 1 and Level 4 targets are not met 6 and 16% of the time, respectively; this represents a quadrupling of the likelihood of triggering demand management measures relative to 1961–1990. The highest frequency of future failure is in August, at the end of the summer dry period, where the Level 1 target is not met 22% of the time, or an average of once every 4 years, and the Level 4 target is not met 8.5% of the time, once every 12 years. In January, the frequency of any of the thresholds being met is very low in both 1961–1990 and $2\times CO_2$ simulations owing to the generally higher flows in winter. Note that these will not correspond directly to the frequency of *implementation* of demand management

Table 2. Frequency with which EA water-demand management flow thresholds at Teddington are reached under present-day and $2\times CO_2$ climates (figure 4). ((i) 1961–1990, present-day simulated flows; (ii) unprocessed: using the $2\times CO_2$ climateprediction.net–CATCHMOD outputs directly; (iii) uniform, equal likelihood across the range of the $2\times CO_2$ climateprediction.net–CATCHMOD outputs; (iv) normal, assuming a Gaussian distribution across the range of the $2\times CO_2$ climateprediction.net–CATCHMOD outputs.)

		flow target (Ml d^{-1})			
		300 (Level 4)	400 (Level 3)	600 (Level 2)	800 (Level 1)
January	1961–1990	0.0002	0.0002	0.0005	0.0005
	unprocessed	0.0003	0.0008	0.0034	0.0100
	uniform	0.0090	0.0120	0.0170	0.0230
	Gaussian	0.0016	0.0017	0.0019	0.0021
July	1961–1990	0.0153	0.0278	0.0346	0.0381
	unprocessed	0.0612	0.0821	0.1198	0.1613
	uniform	0.0310	0.0420	0.0640	0.0850
	Gaussian	0.0025	0.0030	0.0045	0.0064
August	1961–1990	0.0274	0.0371	0.0533	0.0688
	unprocessed	0.0850	0.1102	0.1607	0.2237
	uniform	0.0310	0.0420	0.0630	0.0850
	Gaussian	0.0025	0.0030	0.0044	0.0064

measures, which are only triggered if the flow reaches a given threshold *and* the reservoir storage also below a critical threshold; our hydrological model does not simulate reservoir storage, so these joint probabilities cannot be calculated.

(*d*) *Alternative sampling strategies*

The examples presented above represent an illustrative sensitivity study, where the outcomes are conditional on a number of factors arising from the experimental strategy, including: the choice of climate model, hydrological model, climate and hydrological model parameters to be perturbed, sampling across these parameters, climate variables available, and downscaling methodology. A different experimental set-up would produce different results (Rougier 2007), though we cannot say how different they would be. For example, the bimodal distribution in rainfall change may contain real information about the behaviour of the climate system or it may be an artefact of the GCM structure, the limited number of GCM parameters assessed or of GCM parameter combinations that, with more extensive evaluation, are considered to produce unrealistic climate system behaviour. Various post-processing methods to account for some of the artefacts of the experimental set-up are possible. An emulator can be used to estimate the full response surface across the parameter space, as will be done for the 2008 UK climate change scenarios (Murphy *et al.* 2007); similarly, evaluation of the climateprediction.net ensemble against observations may down weight or exclude particular areas of parameter space (Murphy *et al.* 2004).

Given that the distribution of climate impacts will depend on experimental set-up and post-processing, we explore the effect of two simple alternatives to the direct use of climateprediction.net–CATCHMOD data to estimate frequency of low-flow

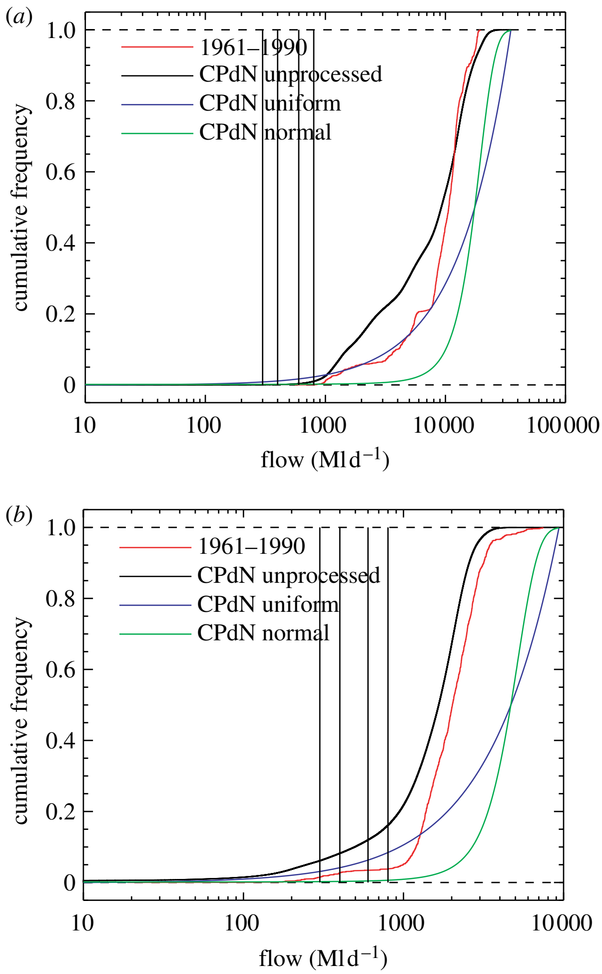

Figure 4. Cumulative frequencies of (*a*) January and (*b*) July monthly discharge for the Thames at Teddington, in the context of environmental flow targets (300, 400, 600 and 800 Ml d^{-1}) set by the Environment Agency for different reservoir capacities. Red shows the frequency for the present day flows (1961–1990). The remaining curves show the frequency from the climateprediction.net–CATCHMOD under different sampling strategies: black, sampling of unprocessed output; blue, assuming a uniform distribution over the range of outputs; green, assuming a Gaussian distribution centred on the middle of the range.

thresholds. These illustrate the point that different likelihoods of impacts will arise dependent on the methodology chosen. The first approach uses uniform sampling across the range of the ensemble, making no assumptions about the distribution within the range; all outcomes within the range of predicted flow statistics are equally probable. The second analysis assumes that the distribution is Gaussian across the range of the unprocessed data; here we set the middle of the range to the mean, and the range is assumed to correspond to 6 standard deviations of the Gaussian.

Results (table 2; figure 4) show that with uniform sampling the likelihood of any demand management threshold being reached is lower in the key summer months of July and August when compared with using unprocessed output. This is because the distribution of the unprocessed $2\times CO_2$ ensemble flows is strongly skewed towards reduced flows (figure 2); uniform sampling reduces the likelihood in this more-populated negative part of the range. For the same reason, Gaussian sampling also reduces the likelihoods of the thresholds in summer. These likelihoods are similar or smaller than the present day (1961–1990) ones, whereas the unprocessed data yield up to a quadrupled likelihood; for example, unprocessed data suggest a Level 3 likelihood in August of 0.11, while the Gaussian or uniform sampling data suggest a likelihood not much greater than today. A water utility may make quite different infrastructure decisions when faced with a Level 3 situation occurring more than once in every 10 years compared to only once every 25 years.

Clearly, if the flow thresholds of interest were nearer the middle of the range (or closer to the end of the range) of simulated flows, the relative frequencies would change; however, they would remain different, in some cases markedly different, in a way that is dependent on the post-processing strategy. For the January flow targets, uniform sampling does, in fact, produce a higher frequency of failure than the unprocessed distribution (albeit a low 0.9 and 2% for 300 and 800 Ml d^{-1}). The few very low flows in January produce a long tail to the distribution of the $2\times CO_2$ ensemble flows; in such a situation, uniform sampling results in a cumulative frequency in the tails of the distribution that is greater than the raw data.

4. Discussion

Our analysis has *illustrated* the potentially rich information that can be obtained by using large perturbed-physics ensemble outputs in a climate change impact assessment. The approach can clearly provide more information than a scenario-based impact assessment. This is illustrated in figure 2, where a scenario approach might produce one or several points on the horizontal axis, whereas with probabilistic information, a frequency distribution or probability distribution can be estimated, and the risks of an adverse impact can be calculated and used to make a risk-based judgement. But figure 4 also shows that different approaches to analysing probabilistic information may lead to a different risk-based decision.

Moving from such an illustrative example to a more complete analysis would require a number of additional elements in the methodology we have used. These include, but are not limited to: (i) use of the transient climateprediction.net ensemble which assesses a wider range of physics perturbations and simulates the transient response to past and future GHG forcing with a coupled ocean–atmosphere model, (ii) incorporation of more sophisticated downscaling methodologies, (iii) consideration of GCM, downscaling and hydrological model structural uncertainties, (iv) estimation of the true response surface(s) for impacts across the parameter ranges in the hierarchy of models used in the end-to-end impacts forecast system, (v) a more sophisticated approach to assessing (and weighting) the skill of individual model combinations in the forecast system, (vi) use of a water resource systems model that enables the assessment of the interplay of demand and supply under different socioeconomic and water infrastructure scenarios, and finally (vii) the development of a methodology that links all these components.

The development of an approach that comprehensively addresses these issues in an end-to-end probabilistic assessment is non-trivial and may be beyond the resources of many organizations. The next set of UK Climate Change Scenarios will provide an 'off-the-shelf' set of probabilistic climate information for many users, but with the proviso the information is dependent on a specific methodology. Further, a full probabilistic impact assessment will require considerable work to estimate probabilities across the entire 'uncertainty cascade'. Organizations without sufficient resources to undertake a full assessment may still be interested in information arising from perturbed-parameter modelling. For example, simply looking at the ranges of predicted outputs, even though their reliability may be questioned, enables an analysis of exposure to them and the risk of not taking the right decisions (Stainforth *et al.* 2007). If potential exposure is deemed serious—and this raises socio-political considerations as individual judgements will need to be made in relation to the accepted level of risk—then a more comprehensive probabilistic assessment might be justified.

However, even with a more comprehensive methodology, the resulting outputs remain conditional: they are the research team's current impacts likelihoods, given the available data and resources (Dessai & Hulme 2004; Rougier 2007). With more data, more resources or an alternative experimental design, the likelihoods will not be the same, though they may or may not be similar.

The challenge therefore is to make use of the richer information that large-ensemble impacts forecasts provide, but to avoid the temptation to consider the results to be fixed, that is, to be 'the probability' of a particular impact. The impacts assessment and, if required, assessment of adaptation options need to be robust in the face of wide uncertainties and the inevitability of estimates of the uncertainty changing over time (Popper *et al.* 2005; Lempert *et al.* 2006). Blind use of a single generation of probabilistic impact information raises the possibility of maladaptation.

The design of methodologies for using large-ensemble climate modelling data in impacts assessment is a developing field, in terms of (i) post-processing of global climate model data and downscaling (Murphy *et al.* 2007; Fowler *et al.* in press), (ii) linking the climate data through impacts to create an end-to-end 'probabilistic forecast system', and (iii) development of approaches for making decisions with probabilistic impacts information. We have shown what an end-to-end impacts assessment might look like, but considerable further work is required to ensure that uncertainty at all steps of the assessment are quantified (such as in the downscaling). Future work is aimed at improving the end-to-end methodology, exploring the relative advantages of simple and sophisticated approaches to probabilistic impacts modelling, and, through the use of real-water resource planning models, developing methodologies for assessing adaptation options and making adaptation decisions.

This study was supported by Environment Agency Science Project SC050045 and by the Tyndall Centre. The views expressed in the paper reflect those of the authors and are not necessarily indicative of the position held by the Environment Agency.

References

Aggarwal, P. K. & Mall, R. K. 2002 Climate change and rice yields in diverse agro-environments of India. II. Effect of uncertainties in scenarios and crop models on impact assessment. *Clim. Change* **52**, 331–343. (doi:10.1023/A:1013714506779)

Araújo, M. B. 2007 Ensemble forecasting of species distributions. *Trends Ecol. Evol.* **22**, 42–47. (doi:10.1016/j.tree.2006.09.010)

Benestad, R. E. 2004 Tentative probabilistic temperature scenarios for northern Europe. *Tellus Series A-Dyn. Meteorol. Oceanogr.* **56**, 89–101.

Beven, K. 2000 On model uncertainty, risk and decision making. *Hydrol. Process.* **14**, 2605–2606. (doi:10.1002/1099-1085(20001015)14:14<2605::AID-HYP400>3.0.CO;2-W)

Beven, K. & Freer, J. 2001 Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *J. Hydrol.* **249**, 11–29. (doi:10.1016/S0022-1694(01)00421-8)

Carter, T. R. *et al.* 2001 Developing and applying scenarios. In *Climate change 2001: impacts, adaptation, and vulnerability* (ed. IPCC), pp. 145–190. Cambridge, UK: Cambridge University Press.

Chen, D. L. *et al.* 2006 Using statistical downscaling to quantify the GCM-related uncertainty in regional climate change scenarios: a case study of Swedish precipitation. *Adv. Atmos. Sci.* **23**, 54–60. (doi:10.1007/s00376-006-0006-5)

Clark, R. T., Brown, S. J. & Murphy, J. M. 2006 Modeling northern hemisphere summer heat extreme changes and their uncertainties using a physics ensemble of climate sensitivity experiments. *J. Clim.* **19**, 4418–4435. (doi:10.1175/JCLI3877.1)

Dessai, S. & Hulme, M. 2004 Does climate adaptation policy need probabilities? *Clim. Policy* **4**, 107–128.

Diaz-Nieto, J. & Wilby, R. L. 2005 A comparison of statistical downscaling and climate change factor methods: impacts on low flows in the River Thames, United Kingdom. *Clim. Change* **69**, 245–268. (doi:10.1007/s10584-005-1157-6)

EA 2007 Climate change and water management: a UK perspective. In *Time to Adapt—Climate Change and the European Water Dimension, Berlin, 12–14 February 2007.*

Feng, J. M. & Fu, C. B. 2006 Inter-comparison of 10-year precipitation simulated by several RCMs for Asia. *Adv. Atmos. Sci.* **23**, 531–542. (doi:10.1007/s00376-006-0531-2)

Fowler, H. J. *et al.* In press. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *Int. J. Climatol.*

Freer, J., Beven, K. & Ambroise, B. 1996 Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach. *Water Resour. Res.* **32**, 2161–2174. (doi:10.1029/96WR03723)

Frei, C., Schöll, R., Fukutome, S., Schmidli, J. & Vidale, P. L. 2006 Future change of precipitation extremes in Europe: intercomparison of scenarios from regional climate models. *J. Geophys. Res. Atmos.* **111**, D06 105. (doi:10.1029/2005JD005965)

Goodess, C. M. *et al.* In press. An intercomparison of statistical downscaling methods for Europe and European regions—assessing their performance with respect to extreme weather events and the implications for climate change applications. *Clim. Change.*

Graham, L., Andréasson, J. & Carlsson, B. 2007 Assessing climate change impacts on hydrology from an ensemble of regional climate models, model scales and linking methods: a case study on the Lule River basin. *Clim. Change* **81**, 293–307.

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

Hall, J., Fu, G. & Lawry, J. 2007 Imprecise probabilities of climate change: aggregation of fuzzy scenarios and model uncertainties. *Clim. Change* **81**, 265–281. (doi:10.1007/s10584-006-9175-6)

Harris, G. R., Sexton, D. M. H., Booth, B. B. B., Collins, M., Murphy, J. M. & Webb, M. J. 2006 Frequency distributions of transient regional climate change from perturbed physics ensembles of general circulation model simulations. *Clim. Dyn.* **27**, 357–375. (doi:10.1007/s00382-006-0142-8)

Haylock, M. R., Cawley, G. C., Harpham, C., Wilby, R. L. & Goodess, C. M. 2006 Downscaling heavy precipitation over the United Kingdom: a comparison of dynamical and statistical methods and their future scenarios. *Int. J. Climatol.* **26**, 1397–1415. (doi:10.1002/joc.1318)

Hulme, M. *et al.* 2002 Climate change scenarios for the United Kingdom: the UKCIP02 scientific report. Tyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia, Norwich, UK.

IPCC 2001 Climate change 2001: synthesis report—summary for policymakers.

Jasper, K. *et al.* 2004 Differential impacts of climate change on the hydrology of two alpine river basins. *Clim. Res.* **26**, 113–129.

Jones, R. N. 2000 Analysing the risk of climate change using an irrigation demand model. *Clim. Res.* **14**, 89–100.

Jones, P. D., Calanca, P., Gyalistras, D. & Fuhrer, J. 2006 Extended riverflow reconstructions for England and Wales, 1865–2002. *Int. J. Climatol.* **26**, 219–231. (doi:10.1002/joc.1252)

Lempert, R. J., David, G. G., Steven, W. P. & Steve, C. B. 2006 A general, analytic method for generating robust strategies and narrative scenarios. *Manage. Sci.* **52**, 514–528. (doi:10.1287/mnsc.1050.0472)

Mearns, L. O. *et al.* 2001 Climate scenario development. In *Climate change 2001: the scientific basis* (eds J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell & C. A. Johnson), pp. 739–768. Cambridge, UK: Cambridge University Press.

Murphy, J. M., Sexton, D. M., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M. & Stainforth, D. A. 2004 Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature* **430**, 768–772. (doi:10.1038/nature02771)

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

New, M. & Hulme, M. 2000 Representing uncertainty in climate change scenarios: a Monte–Carlo approach. *Integr. Assess.* **1**, 203–213. (doi:10.1023/A:1019144202120)

Penman, H. L. 1948 Natural evaporation from open water, bare soil and grass. *Proc. R. Soc. A* **193**, 120–145. (doi:10.1098/rspa.1948.0037)

Pittock, A. B., Jones, R. N. & Mitchell, C. 2001 Probabilities will help us plan for climate change—without estimates, engineers and planners will have to delay decisions or take a gamble. *Nature* **413**, 249–249. (doi:10.1038/35095194)

Popper, S. W., Lempert, R. J. & Bankes, S. C. 2005 Shaping the future. *Sci. Am.* **292**, 66–71.

Prudhomme, C., Jakob, D. & Svensson, C. 2003 Uncertainty and climate change impact on the flood regime of small UK catchments. *J. Hydrol.* **277**, 1–23. (doi:10.1016/S0022-1694(03)00065-9)

Pryor, S. C., Schoof, J. T. & Barthelmie, R. J. 2005 Climate change impacts on wind speeds and wind energy density in northern Europe: empirical downscaling of multiple AOGCMs. *Clim. Res.* **29**, 183–198.

Pryor, S. C., Schoof, J. T. & Barthelmie, R. J. 2006 Winds of change?: projections of near-surface winds under climate change scenarios. *Geophys. Res. Lett.* **33**, L11 702. (doi:10.1029/2006GL026000)

Rougier, J. 2007 Probabilistic inference for future climate using an ensemble of climate model evaluations. *Clim. Change* **81**, 247–264. (doi:10.1007/s10584-006-9156-9)

Salathe, E. P. 2005 Downscaling simulations of future global climate with application to hydrologic modelling. *Int. J. Climatol.* **25**, 419–436.

Stainforth, D. A. *et al.* 2005 Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature* **433**, 403–406. (doi:10.1038/nature03301)

Stainforth, D. A., Downing, T. E., Washington, R., Lopez, A. & New, M. 2007 Issues in the interpretation of climate model ensembles to inform decisions. *Phil. Trans. R. Soc. A* **365**, 2163–2177. (doi:10.1098/rsta.2007.2073)

Tebaldi, C., Smith, R. L., Nychka, D. & Mearns, L. O. 2005 Quantifying uncertainty in projections of regional climate change: a Bayesian approach to the analysis of multimodel ensembles. *J. Clim.* **18**, 1524–1540. (doi:10.1175/JCLI3363.1)

Wilby, R. L. 2001 Seasonal forecasting of UK river flows using preceding North Atlantic pressure patterns. *J. Chart. Inst. Water Environ. Manage.* **15**, 56–63.

Wilby, R. L. & Harris, I. 2006 A framework for assessing uncertainties in climate change impacts: low-flow scenarios for the River Thames, UK. *Water Resour. Res.* **42**, W02 419. (doi:10.1029/2005WR004065)

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