WATER RESOURCES RESEARCH, VOL. 45, W08419, doi:10.1029/2008WR007499, 2009

From climate model ensembles to climate change impacts and adaptation: A case study of water resource management in the southwest of England

Ana Lopez,[1,2] Fai Fung,[1] Mark New,[1] Glenn Watts,[3] Alan Weston,[4] and Robert L. Wilby[5]

Received 2 October 2008; revised 31 March 2009; accepted 29 May 2009; published 13 August 2009.

[1] The majority of climate change impacts and adaptation studies so far have been based on at most a few deterministic realizations of future climate, usually representing different emissions scenarios. Large ensembles of climate models are increasingly available either as ensembles of opportunity or perturbed physics ensembles, providing a wealth of additional data that is potentially useful for improving adaptation strategies to climate change. Because of the novelty of this ensemble information, there is little previous experience of practical applications or of the added value of this information for impacts and adaptation decision making. This paper evaluates the value of perturbed physics ensembles of climate models for understanding and planning public water supply under climate change. We deliberately select water resource models that are already used by water supply companies and regulators on the assumption that uptake of information from large ensembles of climate models will be more likely if it does not involve significant investment in new modeling tools and methods. We illustrate the methods with a case study on the Wimbleball water resource zone in the southwest of England. This zone is sufficiently simple to demonstrate the utility of the approach but with enough complexity to allow a variety of different decisions to be made. Our research shows that the additional information contained in the climate model ensemble provides a better understanding of the possible ranges of future conditions, compared to the use of single-model scenarios. Furthermore, with careful presentation, decision makers will find the results from large ensembles of models more accessible and be able to more easily compare the merits of different management options and the timing of different adaptation. The overhead in additional time and expertise for carrying out the impacts analysis will be justified by the increased quality of the decision-making process. We remark that even though we have focused our study on a water resource system in the United Kingdom, our conclusions about the added value of climate model ensembles in guiding adaptation decisions can be generalized to other sectors and geographical regions.

Citation: Lopez, A., F. Fung, M. New, G. Watts, A. Weston, and R. L. Wilby (2009), From climate model ensembles to climate change impacts and adaptation: A case study of water resource management in the southwest of England, *Water Resour. Res.*, *45*, W08419, doi:10.1029/2008WR007499.

1. Introduction

[2] Projections of future climate change from Global Climate Models (GCMs) suggest significant impacts on the hydrological cycle [*Bates et al.*, 2008]. However, the range in the projections is in some cases very large, in particular at the regional and local scales relevant for the analysis of impacts and adaptation options in the face of climate change. These differences arise from several sources of uncertainty within GCM simulations, including radiative forcing, initial conditions, model formulation (including resolution) and model inadequacy [*Stainforth et al.*, 2007]. The analysis of multiGCM ensembles such as those available from the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC AR4)/Phase 3 of the Coupled Model Intercomparison Project (CMIP3) archives is one approach that provides information on GCM uncertainties [*Solomon et al.*, 2007]. An alternative philosophy makes use of perturbed physics ensembles (PPEs), which are specifically aimed at evaluating uncertainty in GCM formulation [*Murphy et al.*, 2007, 2004; *Stainforth et al.*, 2005]. These ensembles usually comprise a large number of runs of a state of the art climate model, where each individual run uses a version of the model with parameters representing various physical processes set to different values within their acceptable range, as defined by the experts in each

[1]Tyndall Centre for Climate Change and Climate Systems and Policy Group, School of Geography and Environment, Oxford University Centre for the Environment, Oxford, UK.
[2]Currently at London School of Economics, University of London, London, UK.
[3]Science Department, Environment Agency, Bristol, UK.
[4]Water Resources Planning, Environment Agency, Exeter, UK.
[5]Department of Geography, Loughborough University, Loughborough, UK.

0043-1397/09/2008WR007499$09.00

particular area of physics parameterisation. For each combination of parameter values an initial condition ensemble is used so that the relative contribution of formulation and initial condition uncertainty can be evaluated. PPEs provide a new approach for exploring a wide range of future climates and the potential impacts of climate change.

[3] Planning and managing reliable public water supply requires a clear understanding of future hydrological conditions; in particular, supply managers need information about the duration and magnitude of future droughts. Water companies in England and Wales have been required since 1999 to consider climate change impacts as part of their 5 yearly water supply management plans [*Environment Agency*, 2008]. While the first set of plans, developed in 1999, considered climate change by scaling river flows using simple change factors based on a single GCM [*Arnell*, 1999], the 2009 plans required a tiered set of methods that incorporated results from several different GCMs, where the range of results from the GCMs provided some sense of the uncertainty in climate change projections [*Vidal and Wade*, 2007].

[4] The availability of large ensembles of climate models, either multimodel or PPEs, represents a significant increase in climate change information. This should lead to a commensurate improvement in climate change adaptation decisions, but only if decision makers have access to analyses and decision-support techniques that can exploit this information effectively. This need is particularly urgent in water supply planning, where decisions look 25 to 50 years ahead, often creating supply infrastructure that lasts a century or more [*Environment Agency*, 2001].

[5] The purpose of this paper is to evaluate the utility of PPEs for providing improved understanding of climate change for the planning of public water supply. In particular we aim to explore techniques that are amenable to practical application by water supply managers in the next few years. To do this we use water resource modeling tools that are in common use by water supply regulators and water companies in England and Wales, illustrating our approach with a case study on the Wimbleball water resource zone in the southwest of England. This zone is sufficiently simple to demonstrate the utility of the approach, but with enough complexity to allow a variety of different decisions to be made.

[6] The PPE used in this work is part of the climateprediction.net (CPDN) project, the largest PPE experiment to date, here comprising a set of model runs obtained by perturbing 26 parameters in a version of the Hadley Centre GCM, HADCM3L. In order to compare with more standard approaches to impacts analysis, we also consider a multiGCM ensemble, collated in Phase 3 of the Coupled Model Intercomparison Project in support of the IPCC AR4 [*Solomon et al.*, 2007].

[7] In the process of translating the climate model outputs into the appropriate inputs for the hydrological and water resource models, we make some simplifying assumptions. First, we consider a PPE based on only one "parent" GCM, partially taking into account model formulation uncertainty (commonly termed model structure uncertainty) by comparing with the CMIP3 ensemble. Second, we use the individual model runs in the PPE without any prior evaluation of skill or adequacy of each run in representing the climate system (i.e., an unweighted assessment). We also use a relatively simple procedure to downscale GCM-resolution data over the South West of England to the spatial and temporal scales appropriate for the hydrological system, ignoring possible additional sources of uncertainty that could be quantified by considering other downscaling techniques.

[8] Our choice of these simple methodologies is justified because our goal is to explore in what sense, if any, information from a PPE is useful and represents an improvement compared to the use of a single/few model scenarios.

[9] Our exploratory analysis does not therefore represent a robust prediction of climate change or impacts, but illustrates the potential for using large PPEs in adaptation decision making, by exploring a range of feasible impacts and adaptation measures from a substantial sample of possible future climates.

[10] The paper is organized as follows. In section 2 we introduce the PPE climate model data and describe how we use it to simulate the catchment river flows using the hydrological model CATCHMOD. We also describe the basic features of the water resource management model LANCMOD. In section 3 we discuss the results obtained when LANCMOD is run into the future with climate change projections under a "business as usual," where demand, operating rules and supply infrastructure remain unchanged from the present-day. In section 4 we analyze how these results change under different scenarios of demand management and supply infrastructure that could be implemented to adapt or manage the effects of climate change in the future. In section 5 we summarize our results and discuss their implications for impacts studies beyond this particular case study.

2. Data and Methods

[11] The water resources system LANCMOD requires daily time series of river flows at different water abstraction points. In this section we first describe the CPDN experiment and the downscaling and bias correction techniques that we utilize to adjust GCM monthly time series to the appropriate daily input for the CATCHMOD rainfall-runoff model. We then describe CATCHMOD and the main characteristics of the simulated river flows. Finally, we describe the main features of LANCMOD when set up to simulate Wimbleball water resource zone.

2.1. CPDN and Climate Data

[12] The climate data used in our analysis has been generated by the CPDN second experiment launched in February 2006. The GCM used is the HADCM3L, a version of the U.K. Met Office Unified Model comprising a standard resolution atmospheric model coupled to an ocean model with the same resolution; this is lower resolution than the standard ocean model in HadCM3, as archived in CMIP3, which has a finer grid toward the equator. The CPDN experiment explores the effects of perturbing 26 parameters that are relevant to the way radiation, large scale clouds formation, ocean circulation, sulphate cycle, sea ice formation, the land surface and convection are simulated by the GCM (for information on the experimental set up including a description of perturbed parameterizations, forcing scenarios and data available see http://www.climateprediction.net and http://results.cpdn.org).

[13] Each simulation involves a 160-year control run with constant forcing at preindustrial concentrations and a 160-year transient run. In our work we only use the transient simulations which include two phases. In the first phase from 1920 until 2000 the experiment is forced with historical records of CO_2, volcanic and anthropogenic emission, and solar forcing. In the second phase, a range of possible future scenarios are used to force the model response between 2000 and 2080.

[14] CPDN is an ongoing experiment in which individual model simulations are carried out using idle processing capacity on personal computers volunteered by members of the general public. In our study, we concentrate on the first 246 transient simulations that were completed when this project started, and which are now part of the much larger ensemble of completed simulations. Within this subset which represents a partial sampling of the climate model parameter space, all model runs were subjected to the A1B SRES forcing scenario. One member of the ensemble has the standard (unperturbed) values of the model parameters used in earlier single model scenario simulations by the U.K. Met Office, but still uses the lower-resolution ocean. In the rest of the paper this particular model run is termed the "standard version" of HADCM3L.

[15] The CPDN experiment archives a variety of climate variables at different temporal (monthly to decadal means), and spatial scales (grid points to continental averages). Monthly time series are available, for several variables, as the global mean, the area average over 22 continental to sub continental regions similar to those defined by *Giorgi and Francisco* [2000] (these regions are defined as rectangles covering the same land area as the Giorgi regions but including the adjacent oceans, and follow the naming convention of the IPCC 4AR, http://www.ipcc.ch/ipccreports/ar4-wg1.htm), for six ocean "basins" and for eight individual grid boxes over the United Kingdom. Here, we use monthly time series of temperature, precipitation, and relative humidity for the grid box corresponding to the South West of England (48.75N, 5.625W–51.25N 1.875W).

[16] The CMIP3 data used here is derived from 21 GCMs that had completed at least one model run and provided archived precipitation and temperature for the period 1920–2080, under the A1B forcing scenario (the CMIP of models used in this analysis include BCCR-BCM2.0, CGCM3.1, CNRM-CM3, CSIRO-MK3.5, CSIRO-MK3.0, GFDL-CM2.0, GFDL-CM2.1, GISS-AOM, GISS-EH, GISS-ER, FGOALS-g1.0, INGV-ECHAM4, INM-CM3.0, IPSL-CM4, MIROC3.2 (hires), ECHO-G, EHCAM5/MPI-OM, CCSM3, PCM, MRI-CGCM2.3.2, and UKMO-HadCM3). Monthly time series of temperature and precipitation for the grid box corresponding to the Wimbleball catchment are used.

[17] For both the CPDN and CMIP3 data, the use of monthly time series ensures that the intraannual and interannual variability in each model is retained, allowing the exploration of changes in both mean climate and variability in the climate projections.

[18] We also make use of observed data for the Wimbleball catchment, namely daily time series of precipitation, potential evaporation (PE) and naturalised river flows. These data were provided by the Environment Agency of England and Wales and are used in the operation of existing operational hydrological and water resources models. Precipitation data are the area average over the Thorverton catchment derived using three groups of rain gauges for the period 1930–1984. PE is assumed to be the area average representative of the catchment, estimated from a regression with Central England temperature record for 1930–1960, and MOSES v2 data for 1961–1984. River Exe flows at Thorverton are naturalized flow sequences from 1957 to 2005.

[19] The hydrological model that we use to simulate daily discharge in the River Exe at Thorverton is CATCHMOD which requires daily time series of precipitation and potential evaporation to simulate runoff. Thus we describe in detail the downscaling of the monthly GCM time series to produce daily inputs to CATCHMOD before concentrating on the simulation of river flows.

2.1.1. Precipitation Downscaling and Bias Correction

[20] Area-average monthly precipitation for the GCM grid box over SW England is available directly from the CPDN and CMIP3 model runs. The simulated precipitation shows a seasonally varying bias over the historic period when compared to the observed rainfall (Figure 1), underestimating the mean monthly values for nearly all CPDN model runs. The same bias is observed in the CMIP3 simulations, though the effect is not as pronounced as in the CPDN ensemble. These results are consistent with the fact that strong biases are often found in climate model simulations, particularly when looking at individual grid boxes rather than larger regions or continents. As the hydrological model is parameterized to simulate river flows when driven by observed rainfall, any large biases in simulated precipitation will require the model to work outside the range for which it has been calibrated producing potentially meaningless results [*Ebi et al.*, 2007; *Wilby et al.*, 2000; *Wood et al.*, 2004, 2002].

[21] Therefore, to correct the biases in simulated precipitation and simultaneously downscale in time to generate daily precipitation, we adapt a standard methodology based on a gamma transform (or quantile-quantile mapping) that preserves the monthly precipitation distribution [*Hay et al.*, 2002; *Panofsky and Brier*, 1968; *Wood et al.*, 2004, 2002]. For the period in which observed precipitation is available (1930–1984), a gamma distribution is fitted (using maximum likelihood estimation) to the observed monthly precipitation and to the corresponding monthly precipitation from each model run. To correct for GCM bias in the period with observations the quantile for any GCM monthly precipitation value is determined, after which that model monthly value is replaced with the amount corresponding to the closest quantile in the observed distribution. At the same time, the corresponding daily data for that particular month in the observations are used, which produces a daily series that is both bias corrected and has a realistic day-to-day structure. For the rest of the analysis period (1985–2079), the model 1934–1984 distribution is used to compute the quantiles associated with each monthly value from the model in that period (1985–2079), and each model value is then replaced with the observation value closest to the mapped quantile, including the corresponding daily structure.

[22] After the bias correction is performed, the long-term mean monthly precipitation of each model is very similar to the observed mean, and represents a reasonable bias-corrected estimate of precipitation over the catchment. Since

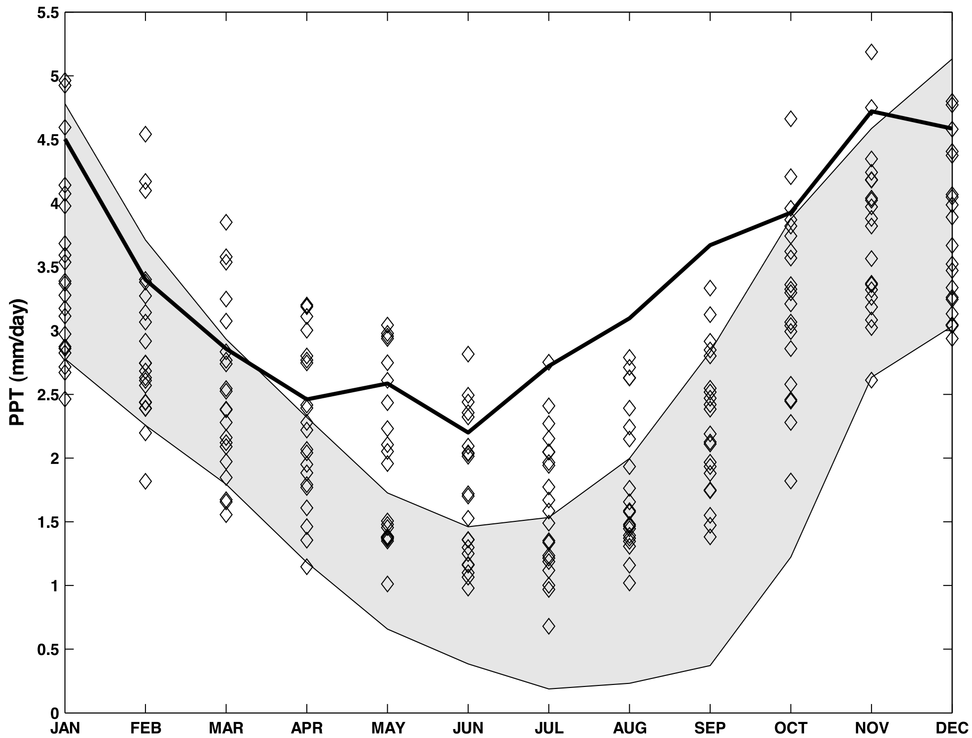

Figure 1. Mean monthly precipitation (mm/d) for the period 1930–1984. The thick line corresponds to observed monthly means, the gray shading indicates the range of precipitate simulated by the CPDN ensemble, and the black diamonds indicate the range of precipitation simulated by the CMIP3 model.

the Gamma transform method is based on mapping observed and simulated quantiles of their corresponding Gamma distributions, the methodology preserves the model intraannual variability in the sense that the sequence of wet and dry months in the raw model data is replicated in the bias corrected data, by sampling the corresponding wettest or driest quantiles in the observed distribution [*Maurer et al.*, 2007b]. Indeed, comparing the downscaled and original GCM monthly series produces correlation coefficients of 0.71–0.95 for the CPDN ensemble and 0.73–0.95 for the CMIP3 ensemble.

[23] In order to check that the bias corrected precipitation time series produce meteorological drought statistics consistent with observations between 1930 and 1984, we compare 12–36-month rainfall deficits in the bias corrected GCM series and the observed precipitation. We define the deficit for each month as the difference between the monthly precipitation and the corresponding long-term monthly mean over the period with observations (1930–1984), and calculate the cumulative deficit for any given month as the sum of the monthly deficits over the previous n months for n = 12, 24 and 36. The droughts are then defined as the largest n month deficits starting at any month. We also calculate an “observed ensemble” by randomly sampling (with replacement) each month from the observed record; this illustrates the range of possible deficits that could appear in the observed record on the assumption that monthly rainfall totals are effectively independent.

[24] Results for the 24-month cumulative deficits are shown in Figure 2, ranked by severity in each ensemble member (from largest to smallest deficit). Generally, the 24-month deficits show a range that is similar to those that would be possible from resampling the observed record. The more severe 24-month deficits in both the CPDN and CMIP3 ensembles are larger than the most severe derived from resampling the historic data, indicating that some of the models in each ensemble have more extensive drought periods. However, there is considerable overlap in the distributions, indicating that the majority of models have 24-month drought characteristics that are similar to those in the observations.

[25] Similar results were obtained for the 12-month precipitation deficits (not shown). The most extreme 12-month deficits are more severe than those from the historic data for about 10% of the CPDN ensemble, but comparable to the historic record in the case of the CMIP3 ensemble, while the less severe cumulative deficits are similar to the historic record for both ensembles. For the 36-month period (also not shown), the most extreme deficits in both model ensembles are more severe than the most severe in the historic ensemble about 25% of the time, suggesting a number of models tend to have more severe long-term deficits than those found in the observations. Nonetheless, there remains considerable (75%) overlap between model and observed deficits. When comparing CMIP3 and CPDN ensembles, 2.5% of CPDN model runs have more severe droughts than the CMIP3 ensemble in the case of the largest droughts, which might be expected from the larger sample size.

[26] In our simulations of this system, 24-month droughts are more critical than 12 month droughts: this suggests that the historic period is represented quite effectively by both ensembles. However, some of the shorter droughts are drier than the historical record for both ensembles, a fact that has

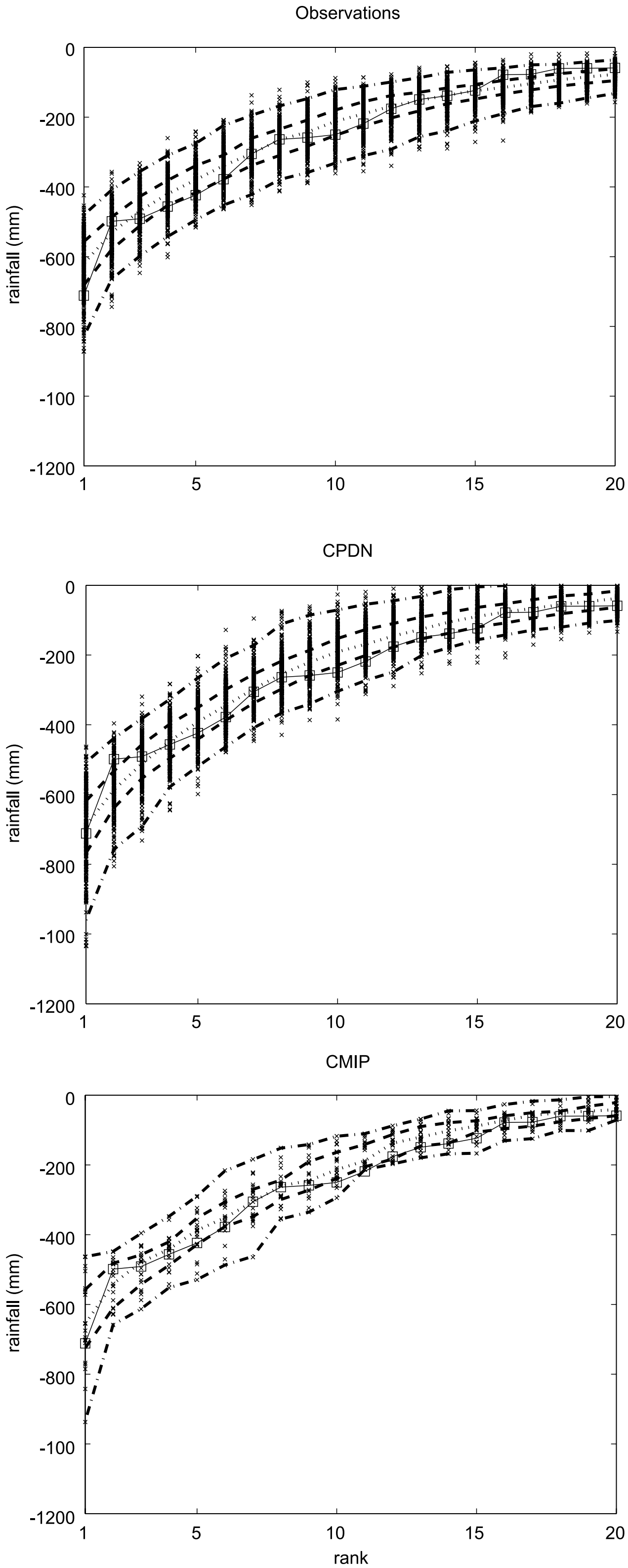

Figure 2. Consecutive 24-months precipitation deficits (in mm) as a function of severity (rank) for the (top) observational record (including 250 samples generated by randomly sampling from the observed record), the (middle) CPDN ensemble, and the (bottom) CMIP3 ensemble. The squares correspond to the observed record, the crosses correspond to individual model runs, and dashed lines represent 97.5, 75, 50, 25, and 2.5% percentiles for each ensemble.

to be taken into account when using the results of our analysis in decision making.

[27] The Gamma transform method is conservative in the sense that it is impossible to obtain daily or monthly precipitation totals that are larger or smaller than those in the observations, because of the fact that the model monthly values are effectively replaced by observed monthly values. This will affect particularly the distribution of the extreme monthly precipitation within the ensemble, since they are all mapped onto the extreme values that appear in the climatology. Previous work using ensembles of opportunity (including combinations of global and regional climate models) to analyze impacts of climate change on hydrological systems, indicates that the results depend in part on the downscaling methodology [*Fowler et al.*, 2007; *Maurer and Hidalgo*, 2008; *Salathe et al.*, 2007; L. J. Manning et al., Using probabilistic climate change information from a multi-model ensemble for water resources assessment, submitted to *Water Resources Research*, 2009]. In the case of the Gamma transform method, it was shown that it performs comparably to more sophisticated dynamical downscaling approaches [*Wood et al.*, 2004], while its computational efficiency makes it suitable for downscaling large ensembles of GCM simulations [*Maurer and Hidalgo*, 2008; *Maurer et al.*, 2007a].

[28] However, precipitation downscaling uncertainty has generally been shown to be a smaller component of the total uncertainty compared with that associated with parent climate models [see *Wilby and Harris*, 2006]. Therefore, while acknowledging that alternative downscaling techniques will produce results that differ from our approach, we concentrate here on the influence of a large PPE on impacts and adaptation, and will address the issue of alternative downscaling approaches in forthcoming work.

2.1.2. Potential Evaporation

[29] The procedure to obtain potential evaporation is more involved since this variable is not available directly from CPDN or the CMIP3 GCMs. We use temperature, and relative humidity from CPDN model runs, and observed wind speed and percentage of sunshine to estimate monthly time series of PE using the Penman formulation [*Penman*, 1948]. To simplify the calculations we assume that wind speed and percentage of sunshine do not change in the future. Furthermore, since large scale PE correlates fairly well with local PE we assume that the potential evaporation calculated with the SW England grid box data is a good representation of the catchment PE, and that monthly mean values approximate reasonably well daily values.

[30] When compared with the historical PE monthly means over 1930–1984, both the CPDN and the CMIP3 ensembles overestimate PE throughout the year, primarily because of a warm bias in the GCMs temperatures. In order to correct for these biases, the simulated PE monthly means were scaled by a factor that makes the simulated long-term monthly mean equal to the observed long-term monthly mean over 1930–1984. Thus, after bias correction, model monthly means end up superimposed with the observed ones. Note that this methodology only adjusts the long-term mean, and does not incorporate additional corrections to the variability as in the precipitation bias correction methodology. Since PE is overestimated through the year, the simulated monthly values are adjusted by factors less than

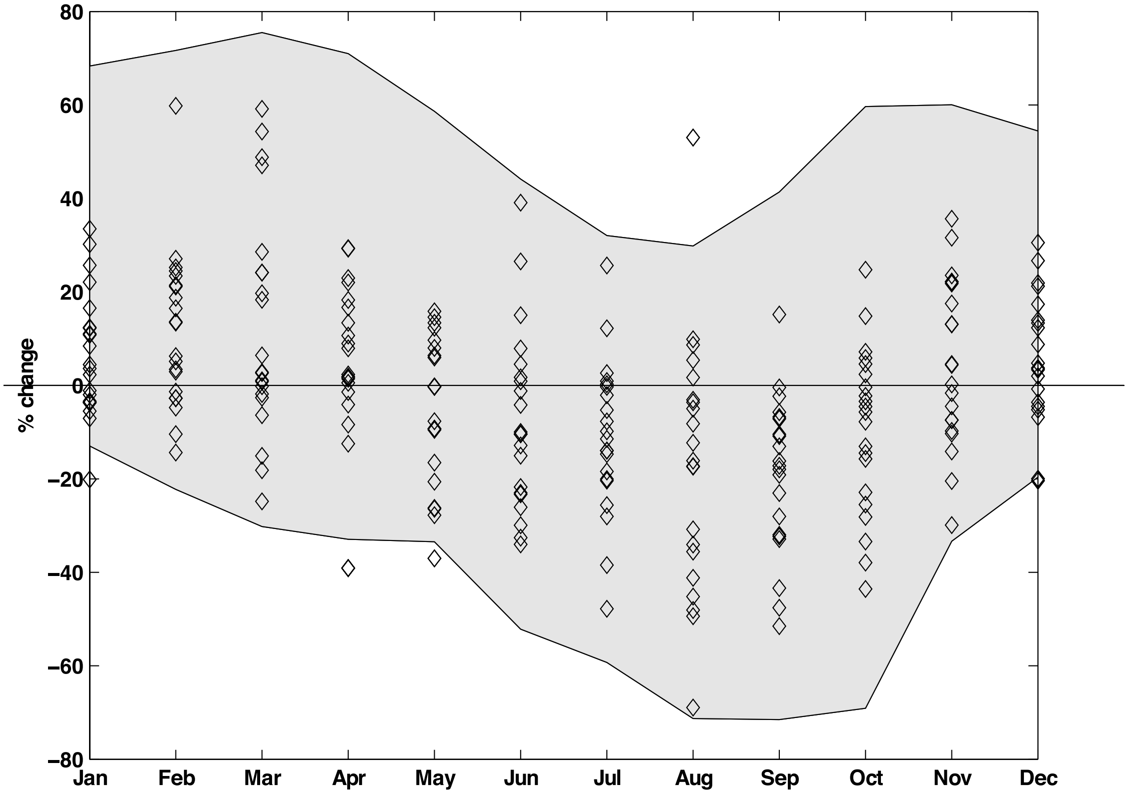

Figure 3. Percentage change in monthly mean River Exe flow at Thorverton between 2020 and 2039 and between 1961 and 1990. The gray shading indicates range of precipitation changes simulated by the CPDN ensemble, and the black diamonds indicate CMIP3 model changes.

one, therefore reducing the range of simulated PE distributions compared with the raw model data; the problem of overdispersion in downscaled PE has been reported before [*Wilby et al.*, 2006].

[31] We note that bias correction is performed at the level of PE and not for the input temperature. The justification for this was that we have observed PE data at the location of interest (Thoverton, South West of England) for precipitation and potential evaporation but not for temperature.

2.2. CATCHMOD and River Flows

[32] CATCHMOD is a rainfall-runoff model used by the EA for water resource planning and abstraction license allocation, and is described in detail by *Wilby et al.* [1994]. It uses daily time series of precipitation and potential evaporation for the river catchment to simulate daily time series of runoff.

[33] For our case study CATCHMOD is set up to simulate river runoff for the Exe River at Thorverton. The effective catchment area consists of approximately 600 km^2 underlain by sandstone. Five hydrological model parameters are determined through calibration against observed discharge. As with any hydrological model, when CATCHMOD is parameterized, there is a range of values of the parameters for which the simulated flows are close enough to the observed one according to an objective function such as the Nash-Sutcliffe measure [*Wilby and Harris*, 2006]. This range of parameters translates into a range of flows all consistent with the observed flow according to a set of tolerances for this measure. Even though this uncertainty is not negligible, previous work has shown that in CATCHMOD it is small compared to the climate model uncertainty [*New et al.*, 2007; *Wilby and Harris*, 2006]. As we wish to focus on uncertainty associated with the climate model PPE, we do not explore this aspect of uncertainty in this work, and use CATCHMOD as calibrated for operational use by the Environment Agency.

[34] We pass the CPDN and CMIP3 downscaled and bias-corrected precipitation and PE time series through CATCHMOD and obtain an ensemble of flows for the Exe at Thorverton. Figure 3 shows the percentage change of mean monthly flow between 2020 and 2039 and 1961–1990. A large proportion of the CPDN ensemble members show substantial reductions in the mean flows during the summer months: 82%, 93% and 91% of the model runs have reduced flows in June, July, and August, respectively. The fraction of CMIP3 models showing a reduction in summer flows is less pronounced, being 67, 76, and 76%, respectively. The reduced summer mean flows produces similarly large reductions in low flows, illustrated in Figure 4, where flow duration curves for daily flows over the period 1961–1990 and over 2020–2039 are illustrated. The fact that the simulated low flows using observed rainfall and PE are larger than the observed flows is a consequence of the inability of CATCHMOD to simulate low flows precisely when it is calibrated to reduce errors in a wide range of flows. Consistently, low flows across CPDN and CMIP3 model ensembles are larger than observed. Moreover, the fact that the observed highest flows are about twice as large as the simulated highest flows (Figures 4a and 4c) confirms the view that CATCHMOD

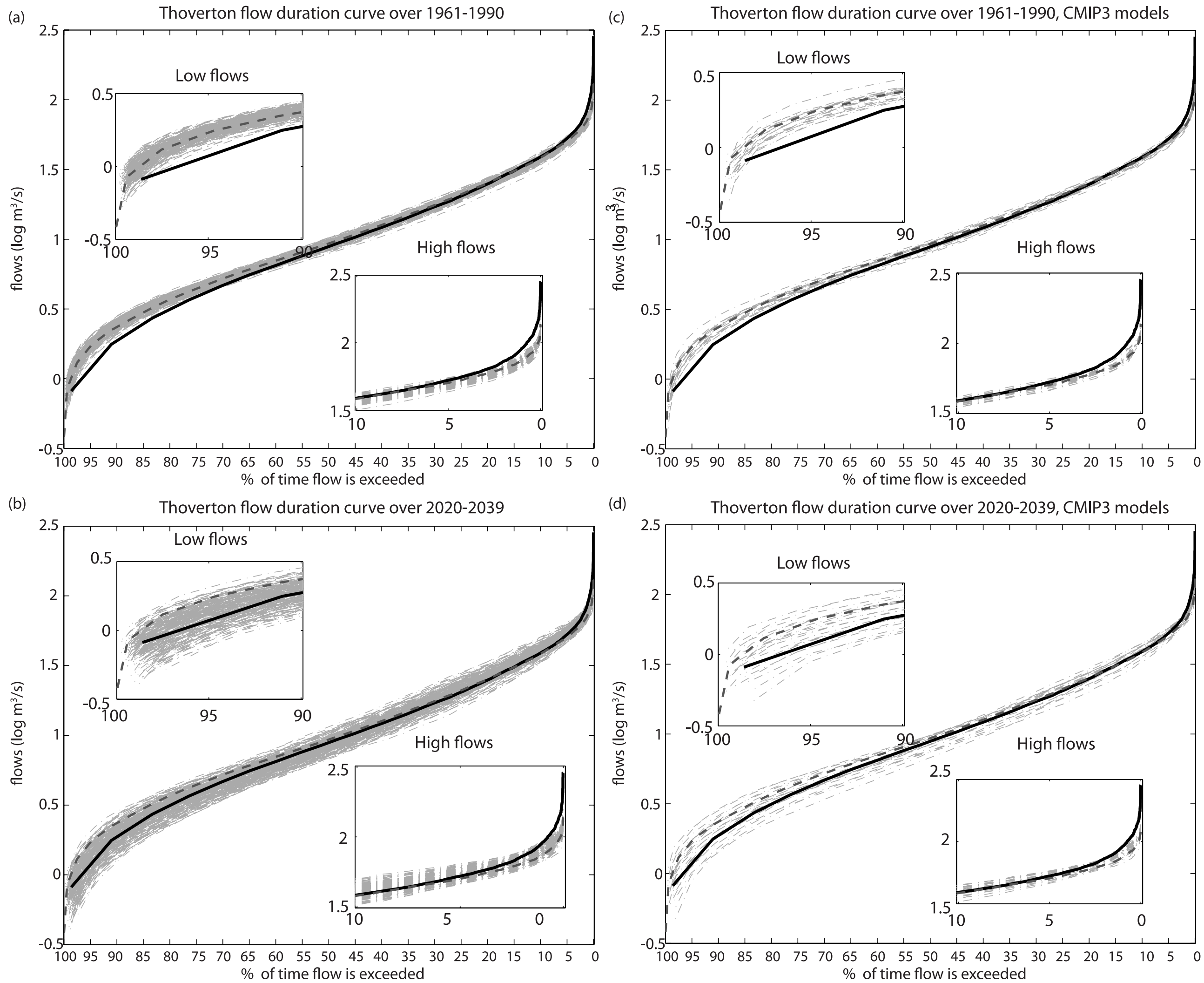

Figure 4. Flow duration curve for daily flows at Thoverton over the periods (a and c) 1961–1990 and (b and d) 2020–2039. The black solid line corresponds to observed flows, the black dashed line corresponds to simulated flows using observed precipitation and PE, and the gray dashed line correspond to CPDN model runs in Figures 4a and 4b and to CMIP3 models in Figures 4c and 4d. The black solid and dashed lines are included for reference Figures 4b and 4d.

is not a suitable tool for estimating peak flows [*Wilby et al.*, 1994].

[35] The spread in the range of simulated flows increases with time as members of the ensembles diverge in their response to the A1B forcing scenario. For instance the ensemble range in the simulated flow exceeded 90% of the time (Q90) increases by about 50% between the baseline period (Figures 4a and 4c) and 2020–2039 (Figures 4b and 4d). However, some changes are common to most of the model runs: for instance, driven by a general decrease in summer precipitation, low flows decrease relative to the baseline period as can be seen by the change in the relative position of CPDN low flows with respect to the observed low flows.

[36] Analogously, driven by a general increase in winter precipitation, high flows increase relative to the baseline period except for the highest flows (Q02 and beyond). In this case simulated peak flows are always smaller than baseline peak observed flows, perhaps because of the fact that CATCHMOD was not calibrated for high flows. Further work is required to confirm whether this is a real result or a consequence of the hydrological model or downscaling errors.

[37] In what follows we assume that the simulated flows described in this section provide information on the spread of plausible future natural flows of the river Exe at Thorverton, and consequently a range of future water availability in that catchment. In order to assess the implications of this spread of future flows for water resources, we use these simulated river flows as inputs to the water resources model described next.

2.3. Water Resource Management System

[38] The Wimbleball water resource zone is situated in SW England and supplies water to the counties of Devon and Somerset. In our simplified version of the zone simulated using LANCMOD (Figure 5), water is supplied by two reservoirs (Wimbleball and Clatworthy), the river Exe

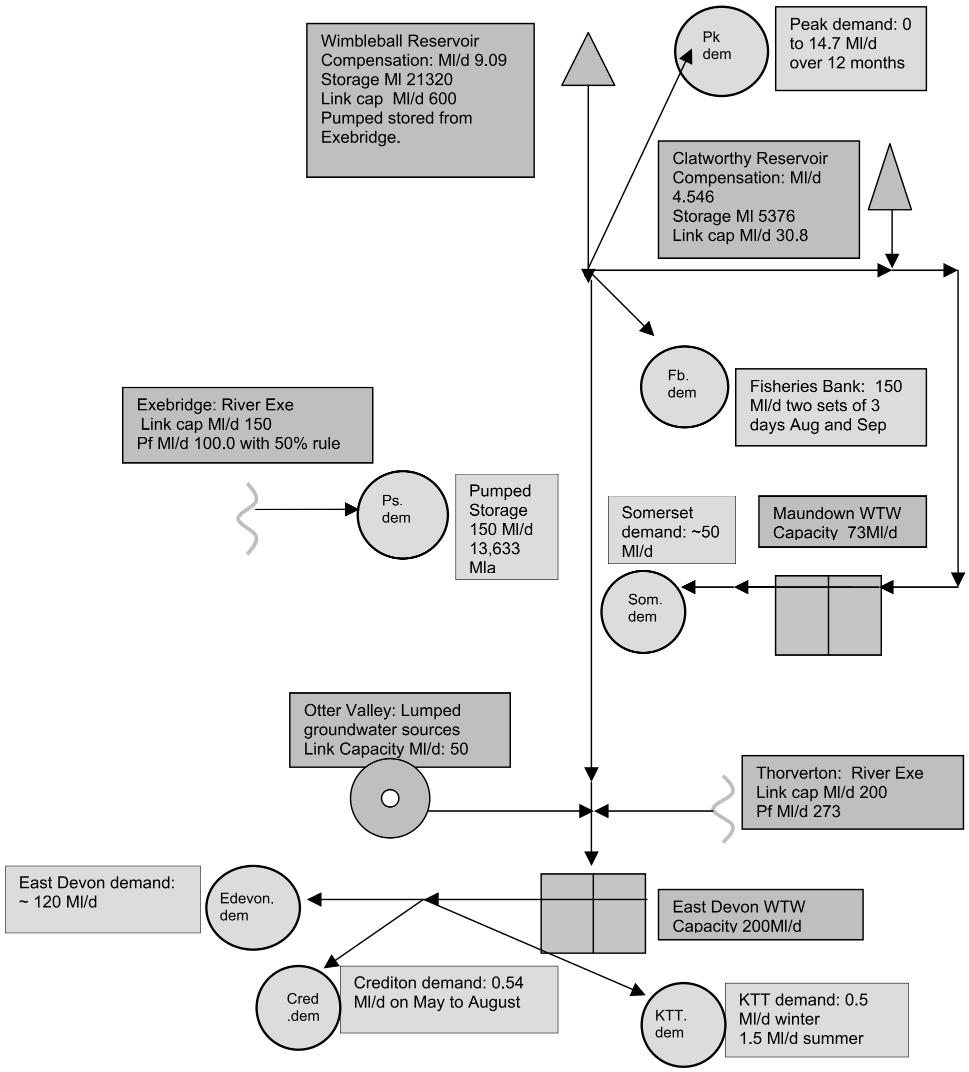

Figure 5. Schematic of Wimbleball water resource zone. Reservoirs, river abstraction points, and groundwater sources are represented by triangles, curvy lines, and punched circle, respectively. Solid circles represent different demands. WTW indicates water treatment plants, and the arrows show the direction of flow between different sources and demands.

(at two abstraction points, Exebridge and Thorverton,) and from a sandstone groundwater source. The largest demands are East Devon, which includes the city of Exeter, Somerset and "Peak." The latter two represent transfers out of the catchment to a neighboring water utility, Wessex Water. There is also a "pumped storage," which is water that is available to be transferred from the river Exe at Exebridge to Wimbleball to refill the reservoir during the winter months.

[39] The main reservoir within the catchment is Wimbleball. It was completed in 1979 on Exmoor, and impounds water from the river Haddeo, a tributary of the Exe, with net storage of 21320Ml and a surface area of 150 ha. It supplies Exeter and parts of East Devon by releasing water into the river Exe to support abstraction at Tiverton and Exeter.

[40] The LANCMOD representation of the Wimbleball system requires as inputs daily time series of river Exe flows at Thorverton and Exebridge, and daily inflows into Wimbleball and Clatworthy. Thorverton flow is simulated by CATCHMOD as described in section 2.2. The other three flow time series are calculate by scaling the Thorverton flows by a factor that ensures that the long-term mean flows coincide with those calculated using the flow estimation software, Low Flows 2000 [*Young et al.*, 2003]. This simplification has been demonstrated by the Environment Agency to be effective in modeling current operation of the Wimbleball system. It is important to note that we assume that the scaling factors will remain the same under climate change. Thus the effects of changes in the reservoir inflows due to climate change are taken into account by representing them as scaled versions of flows of Exe at Thoverton.

[41] LANCMOD is set up to simulate the functioning of the system when different demand profiles, and control rules for river abstractions and functioning of the reservoirs are stipulated. The current values used by LANCMOD as demand profiles (Table 1) are based on present values of water consumption within the catchment. These are mean monthly demands that do not include interannual and intra-month variability, concealing the fact that, for instance during summer hot spells peak demands can be much

Table 1. Annual Demand Profiles for Wimbleball Water Resource System[a]

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
East Devon	114	116	116	115	120	130	138	138	115	111	111	114
Somerset	47.35	48	48.2	49.15	50.3	53.55	56.75	59.55	48.2	46.85	46.1	46
Peak demand	6.3	6.3	6.3	14.7	14.7	14.7	14.7	14.7	14.7	0	6.3	6.3
Pumped storage	150	150	150	0	0	0	0	0	0	0	150	150
KTT	0.5	0.5	0.5	0.5	1.5	1.5	1.5	1.5	1.5	0.5	0.5	0.5
Crediton	0	0	0	0	0.54	0.54	0.54	0.54	0	0	0	0

[a]The values are in Ml/d.

higher. Indeed, our simplified water resource model is not suitable for assessing the impact of peak demands since more information about the structure of the demand and much more detail about the system's constraints would be required. Solutions to guarantee peak demands usually include improving local storage (service reservoirs), larger pumps, pressure reduction, etc., which LANCMOD, in common with other simplified models, cannot simulate. However, models such as LANCMOD are widely used because they represent an efficient way to explore how climate change will affect water supply. They constitute one step in a tiered approach to supply system design where detailed design and day-to-day operational modeling would require complex models that could not easily simulate the impacts of a hundred years of climate data.

[42] The pumped storage is modeled as a demand, and represents the water pumped from the Exe at Exebridge that is available to be transferred to Wimbleball reservoir to ensure the refilling of the reservoir during the winter. It is set up to transfer 150 Ml/d from November to March when required by the control rules governing the reservoir, up to an annual maximum of 13633Ml. The fisheries bank only takes up 150 Ml/d on 2–4 August and 2–4 September.

[43] In section 3 we will concentrate on East Devon and the two demands representing the water transferred to Wessex Water (Peak demand and Somerset), since these are the most significant ones. Notice that here Peak demand is just the name of a demand point and does not represent a peak demand in the operational sense.

[44] To illustrate how the functioning of Wimbleball reservoir is simulated, Figure 6 shows reservoir storage levels for each month of the year over the period 1930–2005, simulated using CATCHMOD river flows between 1930 and 1957, and observed naturalized flows between 1957 and 2005. We also show selected percentiles from the distribution of the monthly historical reservoir levels, along with the only time-dependent operating control rule for the reservoir. The other control rules (not shown) are constants at 21320 Ml, 1100 Ml and 0 Ml (100%, 5% and 0% of capacity). The operation of the reservoir by LANCMOD is as follows: (1) When the storage level is between the maximum capacity (21320 Ml) and the time varying control rule, water is released into the system at a rate of 350 Ml/d and there is no transfer from the pumped storage. (2) When the reservoir storage falls under the time varying control level, the transfer of water from the pumped storage is triggered at a rate of 150 Ml/d up to a maximum of 13633 Ml per annum, but only between November and March. (3) The same rule applies if the storage level falls under 1100 Ml. (4) For any storage level there is a compensation flow released back into the river of 9.1 Ml/d.

[45] Clearly this system is demand driven, as its priority is to satisfy the different demands. One consequence of these rules is that water from the reservoir is always released into the system at a maximum rate of 350 Ml/d, limited only by the volume of water left in the reservoir, even when its level is so low that it would in practice be inappropriate to operate in this way. In other words, in our simulation the only role of the control rules is to determine when water is pumped into Wimbleball. The volume supplied by the pumped storage is limited by what can be extracted at Exbridge, where the flow has to exceed 400 Ml/d (4.62 m^3/s) to be able to supply 150 Ml/d (1.73 m^3/s) to the pumped storage. When the flows fall under 340, 280, 220, 160, 130, and 120 Ml/d, water is supplied to the pumped storage at a rate of 120, 90, 60, 30, 15, and 10 Ml/d, respectively.

3. Water Resources Model Under Business as Usual Demand Scenario

[46] In this section we describe the response of the water resource system when run into the future with current reservoir capacities and demand profiles, and under the same operational rules; this is termed the "business as usual" (BAU) scenario.

[47] We will present our results as monthly averages for the different variables of interest, as the daily structure of precipitation that we used to generate the river flows was borrowed from the observations. In doing the bias correction and simultaneously the temporal downscaling, we have assigned to each model month a daily structure taken from the observed month with the closest quantile in its Gamma distribution. Therefore any property that depends closely on daily structure will be partly reflecting characteristics of the observed climate and not necessarily of model data. On the other hand, we expect monthly averages to be more faithful indicators of model behavior.

3.1. Wimbleball Reservoir

[48] A natural question to ask is whether the management of the reservoir would need changing under future climatic conditions. To this end, we assume that our ensemble of climate model runs provides a sample of what might happen at any chosen time slice in the future. Figure 7 is similar to Figure 6 except we now show the distribution of Wimbleball reservoir levels for a single year in the future, as simulated using the 246 members of the CPDN ensemble. A similar diagram could be constructed for any year in the future. The different percentiles in Figure 7 represent the storage level exceeded by a given fraction of models for each month in 2040. Figure 7 suggests that if one uses the climate model ensemble as practitioners usually look at historical records, the existing operating control rule designed following

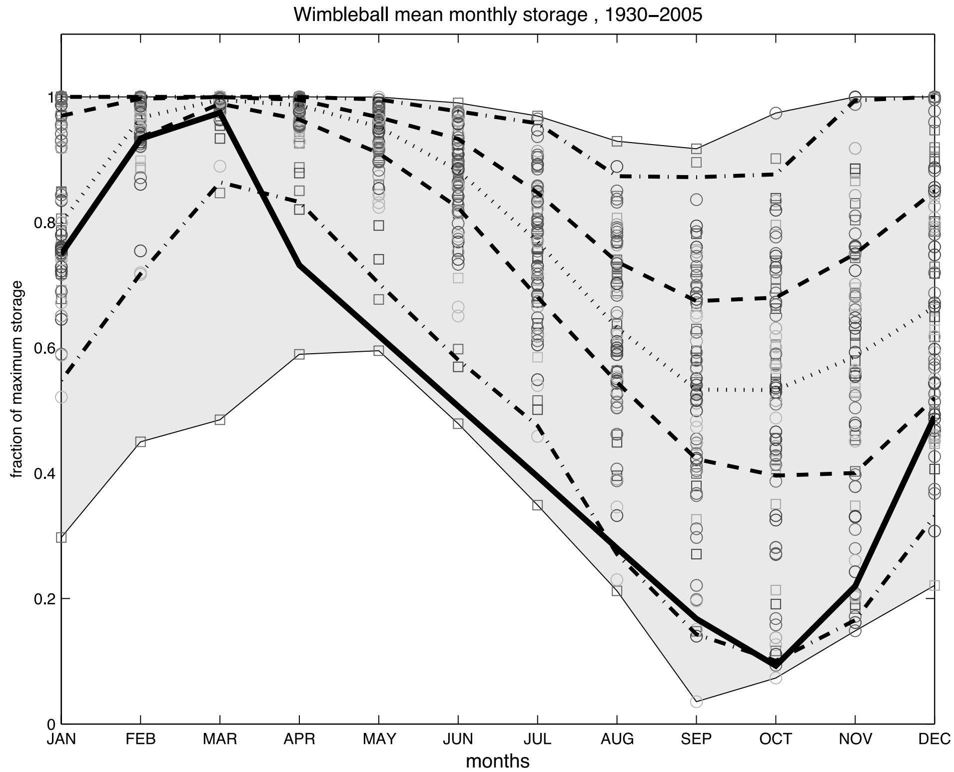

Figure 6. Wimbleball mean monthly storage for historical data represented as fraction of maximum storage. Squares are storage levels estimated using simulated flows between 1930 and 1957. Circles are levels using observed river flows between 1957 and 2005. The black solid line indicates the control rule described in the text. Dashed lines represent (from bottom to top) the reservoir level exceeded 2.5, 25, 50, 75, and 97.5% of the time, respectively.

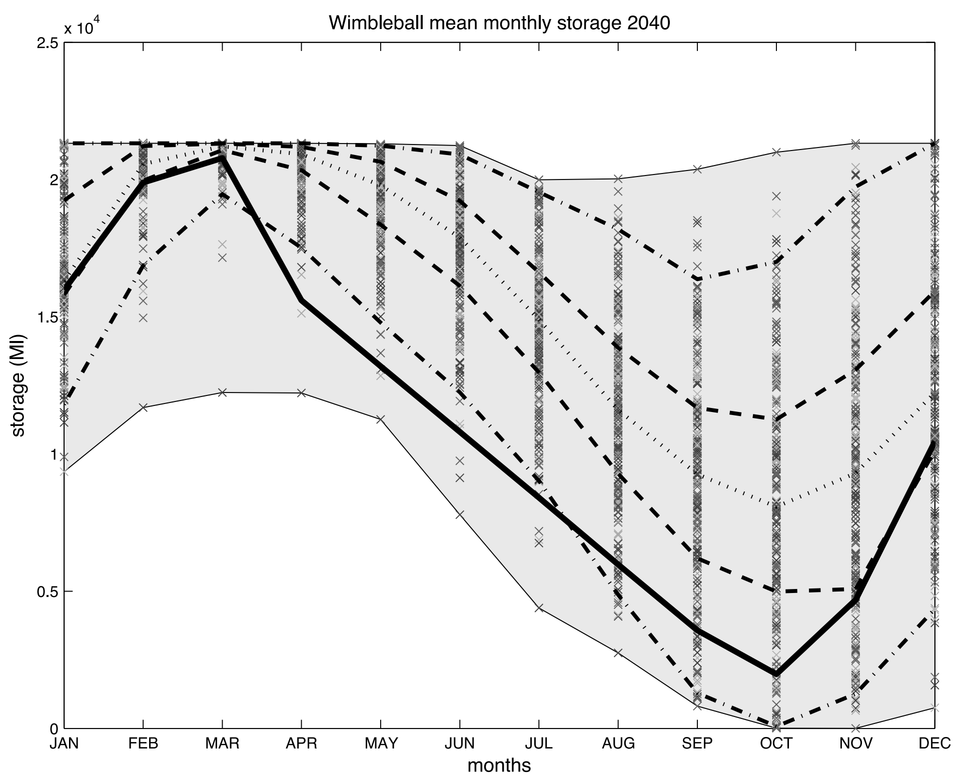

Figure 7. Wimbleball fraction of maximum storage level for 2040 using CPDN ensemble (crosses). The black solid line indicates the control rule described in the text. Dashed lines represent (from bottom to top) the reservoir level simulated by 2.5, 25, 50, 75, and 97.5% of the model runs, respectively.

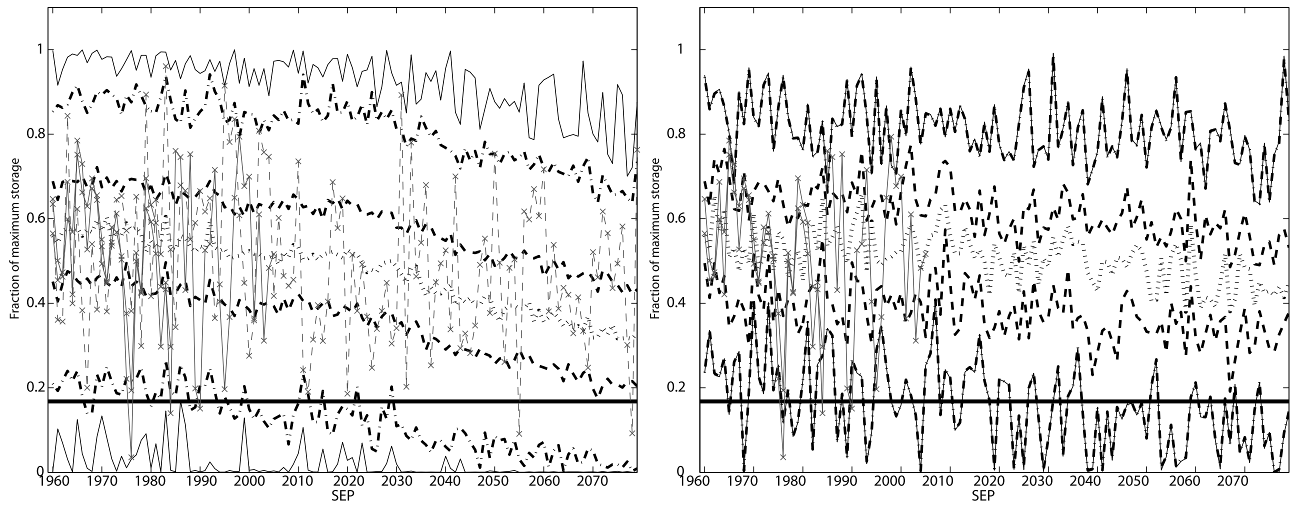

Figure 8. Mean monthly fraction of maximum storage level for September between 1960 and 2079 for the (left) CPDN ensemble and the (right) CMIP3 ensemble. The black lines represent (from top to bottom) maximum values (solid line), 97.5% (dotted-dashed line), 75% (dashed line), 50% (dotted line), 25% (dashed line), 2.5% (dotted-dashed line), and minimum values (solid line) across the climate model ensembles. The thick solid line corresponds to the control rule described in the text. Light gray and dark gray crosses connected by solid lines indicate storage levels simulated by LANCMOD using observed flows and simulated historical flows, respectively. Crosses connected by dashed lines in Figure 8 (left) correspond to storage levels for the CPDN model run with standard values of the physical parameters and illustrate the variability of a single model run.

historical records might need some adjustment in the future. To illustrate this we look at the driest period of the year, August to October. In the case of reservoir capacity simulated with the historical records, in half of the years reservoir levels are above 55% capacity, and in three quarters of the years levels are above 40% capacity. In the case of the CPDN ensemble simulations in the 2040s, the reservoir level exceeded by half the models is lower, at 40% of capacity while that exceeded by three quarters of the models drops to 25% of capacity. Similar values are obtained for the 50% and 25% proportions of the CMIP3 ensemble, although the lowest simulated capacities are similar to the 25th percentile of the CPDN ensemble (not shown), consistent with the fact that CPDN is drier than CMIP3.

[49] More interestingly, we can explore how the reservoir storage level across the ensemble changes over time. Figure 8 represents the storage level for a single month as a function of time between 1960 and 2079; we show September as one crucial month toward the end of the summer, when the reservoir level often becomes particularly low. Figure 8 shows that the storage levels associated with each CPDN percentile decrease slowly from the present to about the 2020s, and more rapidly later on. For instance the 50th percentile goes from nearly 60% of full capacity to about 50% by the 2020s, and ends up at nearly 30% by the end of the simulation period. In the case of the CMIP3 ensemble the 50th percentile shows a less pronounced drying behavior toward the end of the simulation period. For the lower percentiles, it is hard to identify the trends because of the noise associated with the smaller size of this ensemble. Therefore while the CPDN ensemble shows a clear drying trend, this is not as evident in the CMIP3 ensemble. As we discuss later, in this work we did not perform any validation of the CPDN ensemble. However, if the result of Figure 8 holds after such a validation is carried out, they would suggest that the CMIP3 ensemble does not include possible future realizations of the driest climates, and it might give a false confidence that low flows would not be a problem in the future.

[50] Our results suggest that a large climate model ensemble allows for the quantification of the evolving risk of the reservoir reaching critically low levels under current operating rules and for that particular climate model ensemble.

3.2. East Devon and Somerset Demands

[51] LANCMOD is designed so that different priorities can be assigned for the order in which demands are supplied from different sources. For instance Peak demand is only supplied from Wimbleball reservoir, while Somerset demand is satisfied first by Wimbleball and, if not enough water is available from this reservoir; thereafter by Clatworthy. In the case of East Devon, the demand is supplied first from the groundwater source, then from river Exe at Thorverton, and finally from Wimbleball. This priority order is important when analyzing how demands are satisfied under different scenarios. For instance East Devon has priority over the other demands on the groundwater source. Since in the present configuration groundwater supply is fixed to be 50 Ml/d, nearly half of East Devon annual mean (120 Ml/d) is guaranteed by this source, imposing a lower bound on the possible deficit even under climate change.

[52] When LANCMOD is run using historical flows between 1930 and 2005, the only time that East Devon demand cannot be satisfied is September 1976, representing a ~1% risk of failure. On the other hand, when run using the CPDN ensemble, between zero and 3 models fail for any given year during the baseline period, 1960–1989 (Figure 9),

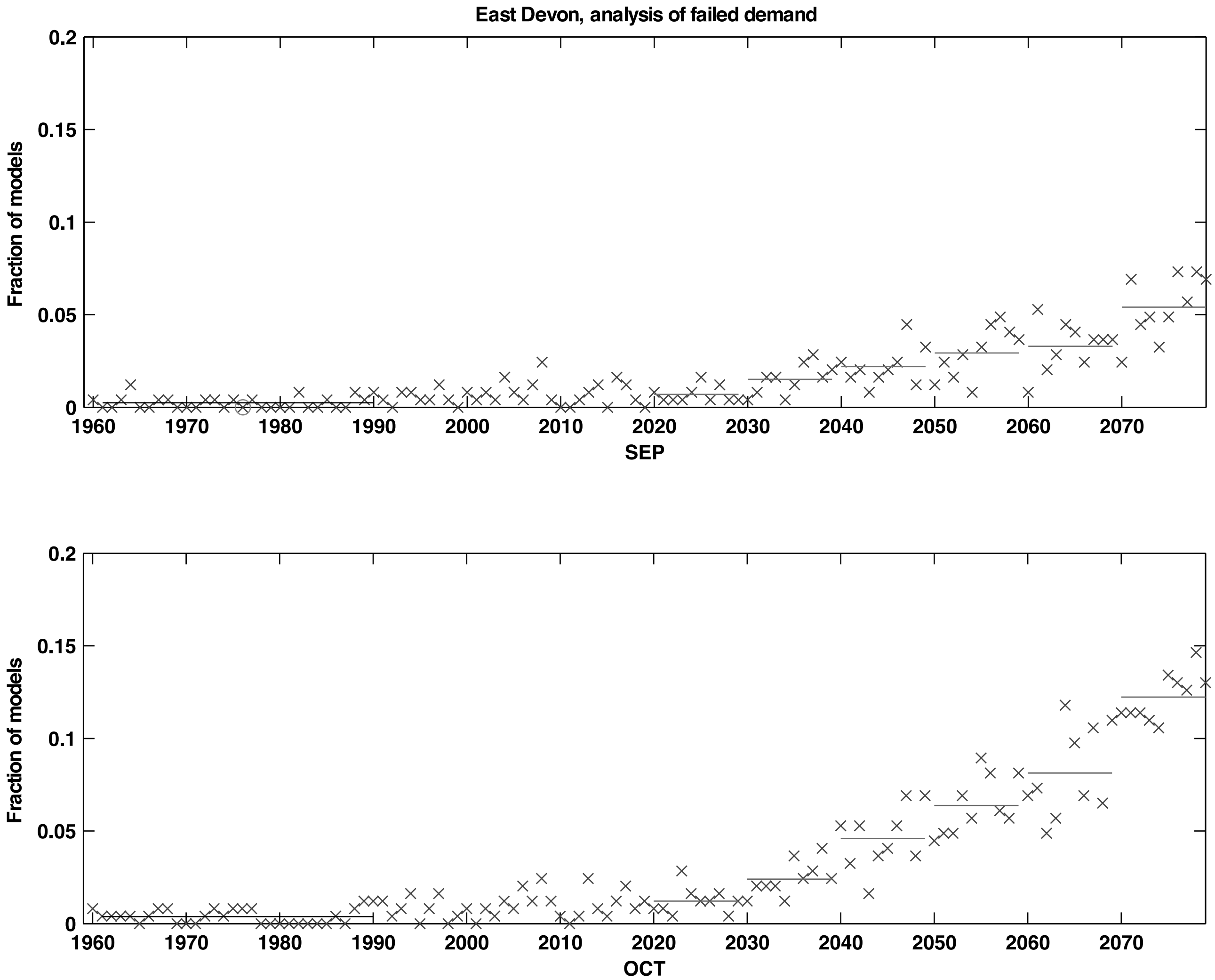

Figure 9. East Devon demand. The fraction of models that fail to supply (top) September and (bottom) October average monthly demand each year. The horizontal lines indicated the mean over 1960–1989 on the left and the corresponding decades after 2020. The circle on September 1976 signals the only time demand could not be supplied within the historical record.

representing a risk of failure not larger than 1.2% across the ensemble. This is consistent with reservoir levels shown in Figure 8, where fewer than 2.5% of the models have storage levels below the control rule line in September. Of course, the fact that the storage is less than the control rule value does not imply automatically that a demand will not be satisfied, since LANCMOD is set up to try to satisfy all the demands as its first priority. Therefore, before failing, LANCMOD will try to exhaust all possibilities including depleting the reservoirs and looking for alternative sources of water within the system. In particular, in the case of East Devon, it will be supplied first by the groundwater source and Exe at Thorverton, and ultimately by Wimbleball. Recall that in our simplified version of the system, the groundwater source guarantees a constant supply of 50 Ml/d unaffected by climate change. This correlates with the fact that even though about 2.5% of the CPDN models have very low reservoir level in September 1960–1990, only a small fraction of them actually fails to satisfy East Devon demand.

[53] Looking to the future, the ensemble shows that the fraction of CPDN simulations failing to supply East Devon demand in September for any decade beyond 2030 is at least three times that in the baseline period ($\sim$0.6%), and reaches about 5% of the models in the 2070s (Figure 9). Figure 9 (bottom) shows analogous information for October. A higher fraction of models fail in this month at any time compared to September, suggesting that the critical period for satisfying demand will shift toward the autumn, according to our climate model ensemble. The pattern of failures in November (not shown) is similar to the one in September, and there are even fewer failures from December through the rest of the winter.

[54] At Somerset BAU demand is roughly half East Devon and can also be satisfied by Clatworthy. Therefore, even though the fraction of CPDN models failing in the baseline period is similar to East Devon (about 0.2% on average), the increase in the future is lower, with only about 2% of simulations failing by the 2070s (Figure 10). For Somerset, there is also an increment in the fraction of

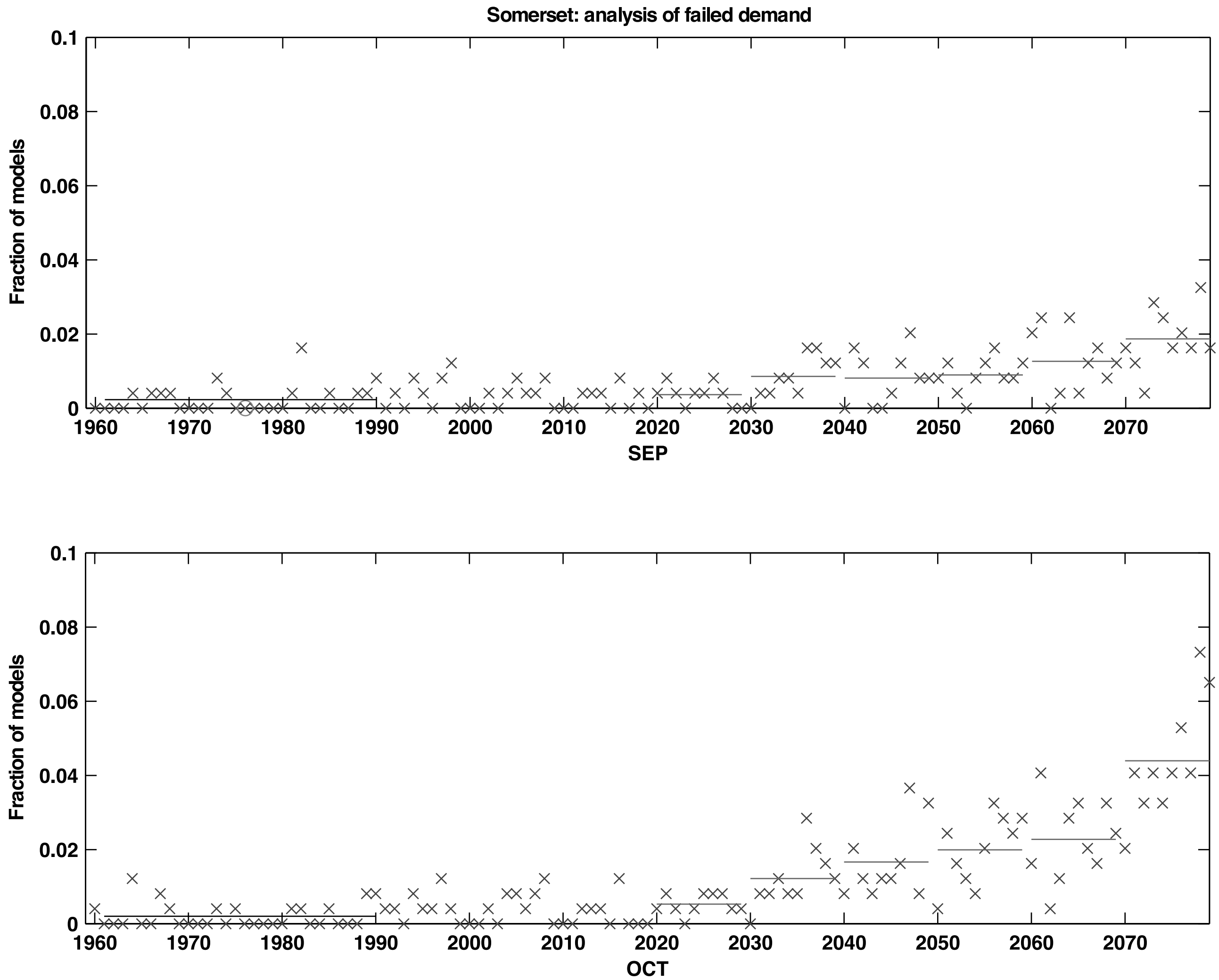

Figure 10. Same as Figure 9 for Somerset demand.

models failing in October (Figure 10 (bottom)) and similarly in November (not shown) compared to September.

4. Adaptation and Management Options Under Climate Change

[55] Up to this point we have described the simulated changes in water availability under a business as usual scenario, assuming that demand patterns and reservoir maximum storage will not change in the future. We will discuss next how the system responds under different scenarios for supply and demand management options.

[56] Various changes can be made to the water resource model set up to explore how different adaptation strategies can decrease the risk of supply failure in the future, making the system more resilient to climate change. Options include reducing demands to comply with water saving policies, increasing the volume of water available in the reservoirs, reducing the transfer of water outside the catchment, increasing the pumping rate to the reservoir, and changing the control rules that govern the river abstractions reducing the flows maintained for the environment. The above is simply a list of the changes that could be simulated given our water resource model.

[57] In a real situation the path chosen to adapt to possible impacts of climate change will depend on many factors that include, but are not limited to, the climate information. In particular issues such as the cost of the different options, their impacts on the environment, public response, technical feasibility, as well as demographic and water use changes, will play important roles in the decision.

[58] In this case study we will concentrate on four different scenarios that can be simulated by making minor changes to way LANCMOD is set up. Two scenarios are based on consumption reduction, one on increasing supply and one on combined increased supply and reduced demand. Our goal here is to analyze how the climate model information can be used to help inform these management options. An integrated assessment of possible adaptation strategies, taking into account the full range of socioeconomic factors and their uncertainties, will be the focus of future work.

4.1. Demand Reduction Options

[59] For the purposes of this study, we have assumed that the baseline household demand in the area is 150 l/h/d, the current average for England and Wales. The recent United Kingdom's government water strategy [*Department for*

Environment, Food, and Rural Affairs, 2008] aims for this to reduce to 130l/h/d, suggesting that a reasonable scenario for demand reduction would be about 15% less than the current values. In this scenario, we are implicitly assuming that nonhousehold and other demands also fall by the same proportion.

[60] Since the two most significant demands through the year are East Devon and Somerset combined with Peak demand, we devise two different scenarios involving them. The other demands are either very small (less than 2% of East Devon or Somerset), or operate only a few days during the year (Fisheries bank), having relatively little impact on the whole system, therefore we leave these unaltered.

[61] In the first demand management scenario, labeled ED_{red}, we assume that only the annual East Devon demand profile is reduced by 15%. East Devon is the largest demand in the system with a yearly figure of 120 Ml/d, and ED_{red} results in a reduction of 18 Ml/d. This illustrative scenario could be seen as representing, for example, a demand management program targeted on part of the resource zone, and addressing long-term water consumption as opposed to peak demands.

[62] In the second scenario, labeled ALL_{red}, we reduce by 15% all main demands: East Devon, Peak and Somerset, amounting to an annual average reduction of about 28 Ml/d. Both scenarios assume that a water saving strategy is put in place, and other factors, such as population changes, remain constant.

[63] From the water resource management point of view, it is important not only to know whether there will be a failure, but also the frequency of occurrence of any number of consecutive monthly or annual failures. For instance, we might be interested in knowing how the occurrence of one monthly failure will change in the 2030s compared with the baseline across the climate model ensemble. Or the analogous change for failures that occur in 2 or more consecutives months, and how these changes depend on the demand scenario proposed.

[64] In Tables 2a and 2b we show the results for the number of models that fail to satisfy Somerset and East Devon demand during a single month, and from 2 to 6 consecutive months, over seven decades in the future. We do not tabulate failure frequency for periods longer than 6 months, as there is only a single failure of this severity in the final decade.

[65] If the BAU management scenario is adhered to, the fraction of CPDN models that fail to supply both Somerset and East Devon increases markedly in the future. For example single month failures occur 1–4 times as frequently by the 2030s, and 9–10 times more often in the 2070s. The ED_{red} scenario significantly reduces the number of CPDN models failing to supply demand in the future, by up to two thirds compared to the BAU in 2070s. Compared to the present-day BAU, ED_{red} still produces more failures but the increase is much lower than under BAU. The ALL_{red} scenario does not affect East Devon significantly, but is more effective at Somerset, as this scenario involves a reduction in Somerset demand on top of the East Devon reduction.

[66] As would be expected, the fraction of CPDN models failing in the historical baseline period for the ED_{red} scenario is equal or smaller than for the BAU scenario. If all demands are reduced (ALL_{red}) the situation is further improved specially for Somerset, with zero consecutive failures in the baseline period in this case. In the future, under any scenario, the number of CPDN models failing increases, consistent with the reduction in summer river flows simulated by the ensemble; even though winter flows increase in some ensemble members, they are clearly not sufficient to refill the reservoirs (see section 4.2). Compared to the BAU scenario, both ED_{red} and ALL_{red} result in far fewer consecutive monthly failures, especially after the 2040s. ALL_{red} scenario is again particularly effective for Somerset, completely eliminating isolated monthly failures and working very well until the 2060s for 2 or 3 consecutive monthly failures.

[67] A comparison of ED_{red} and ALL_{red} scenarios for East Devon suggests that the management of different demands interacts nonlinearly within this system, since even though reducing all demands does not affect particularly the number of isolated monthly failures in East Devon, it does have an effect for two consecutive monthly failures.

[68] In the case of the CMIP3 ensemble, the fraction of failures for the BAU scenario is larger in the base period for East Devon but the increase in the number of failures is smaller consistent with the fact that this ensemble shows lower risk of emptying the Wimbleball reservoir toward the end of the simulation period. As expected, we can observe similar patterns of changes under different demands options for both, the CPDN and CMIP3 ensembles, especially in the case of East Devon where a small trend in the number of failures is present.

[69] We also evaluate the changes in multiyear system failure. Tables 3a and 3b shows the number of ensemble members failing to meet demand in single and multiple consecutive years, over 30-year periods. Failures for more than 6 consecutive years are very rare and appear mostly in the last three decades of the simulation period. Here we define an annual failure whenever the annual volume of water supplied does not coincide with the annual volume required, independently of whether that occurs in just one or more months within the year. Therefore, Tables 3a and 3b do not provide information about when and how the demand could not be satisfied within any particular year. Nevertheless, they do provide information about how demand management options can remediate the fact that under simulated future climate and BAU demand and supply, the fraction of models failing 2 or more consecutive years increases in time.

[70] Under the BAU scenario only one CPDN ensemble member fails to meet Somerset demand in any 2 consecutive years in 1960–1989, but this increases to 10 and 27 members for 2020–2049 and 2050–2079, respectively; the increase in consecutive year failures is even greater at East Devon. When East Devon demand is reduced by 15% (ED_{red}), the number of models failing reduces by three times at Somerset and at least six times at East Devon for 2020–2049. Implementing the additional 15% reduction in Somerset and Peak demand (ALL_{red}) completely eliminates two annual consecutive failures in the future, apart from final period for East Devon. Thus, for this system, relatively small year-round demand reductions could eliminate the need for more drastic measures in critical dry consecutive years.

[71] The number of failures in 2 or more consecutive years is practically nonexistent for the CMIP3 ensemble.

Table 2a. Number of Models That Fail to Satisfy Somerset Demand for Between 1 and 6 Consecutive Months on a Decadal Basis[a]

Demand/Supply Scenario	Consecutive Monthly Failures														
	1					2					3				
	BAU	ED_{red}	ALL_{red}	L_{res}	$L+All_{red}$	BAU	ED_{red}	ALL_{red}	L_{res}	$L+All_{red}$	BAU	ED_{red}	ALL_{red}	L_{re}	$L+All_{red}$
1960–1989	3(3)	2(2)	0	3(3)	0	3(3)	1(1)	0	2(2)	0	2(2)	2(2)	0	2(2)	0
2020–2029	2(2)	2(2)	0	2(2)	0	5(6)	2(2)	0	4(6)	0	7(7) [1]	1(1)[1]	0	5(5) [1]	0
2030–2039	2(2) [1]	2(2)	0	3(3)	0	21(23)	10(10)	0	17(19)	0	6(6)	3(3)	0	6(6) [1]	0
2040–2049	10(11)	4(4)	0	8(8)	0	24(27) [2]	7(7) [1]	0	22(23) [2]	0	8(8)	7(7)	0	6(6)	0
2050–2059	14(14) [3]	4(4) [1]	0	13(14) [2]	0	28(31)	9(9)	0	18(19)	0	13(14)	5(5)	1(1)	9(9)	1(1)
2060–2069	17(19)	5(5)	0	14(16)	0	35(40) [1]	16(16) [1]	2(2)	32(35) [1]	2(2)	13(15)	5(5)	0	8(10)	0
2070–2079	29(33)	9(10)	0	22(25) [1]	0	51(53) [3]	16(17)	2(2)	31(31) [2]	1(1)	20(24) [1]	12(16) [1]	1(1)	21(22) [1]	1(1)

Demand/Supply Scenario	Consecutive Monthly Failures														
	4					5					6				
	BAU	ED_{red}	ALL_{red}	L_{res}	$L+All_{red}$	BAU	ED_{red}	ALL_{red}	L_{res}	$L+All_{red}$	BAU	ED_{red}	ALL_{red}	L_{re}	$L+All_{red}$
1960–1989	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2020–2029	2(2) [1]	2(2)	0	2(2)	0	0	0	0	0	0	0	0	0	0	0
2030–2039	5(5) [1]	6(6)	0	6(6)	0	4(4)	2(2)	0	3(3)	0	0	0	0	0	0
2040–2049	5(5)	4(4)	0	5(5)	0	2(2)	1(1)	0	1(1)	0	0	0	0	0	0
2050–2059	7(8)	5(6)	0	6(7)	0	2(2)	1(1)	0	2(2)	0	0	0	0	0	0
2060–2069	7(7)	6(6)	0	6(6)	0	1(1)	0(0)	0	1(1)	0	0	0	0	0	0
2070–2079	14(19)	9(9)	0	13(16)	0	3(3)	4(4)	1(1)	4(4)	0	1(1)	1(1)	0	1(1)	1(1)

[a]In the baseline period (1960–1989), this is the mean number of models per decade. Within each cell the first value is the number of CPDN models, the second value (in parentheses) is the total number of failures for the CPDN ensemble, and the third value (in brackets) is the total number of CMIP3 models (only indicated when different from zero).

Table 2b. Number of Models That Fail to Satisfy East Devon Demand for Between 1 and 6 Consecutive Months on a Decadal Basis[a]

Demand/Supply Scenario	Consecutive Monthly Failures														
	1					2					3				
	BAU	ED_{red}	ALL_{red}	L_{res}	$L+All_{red}$	BAU	ED_{red}	ALL_{red}	L_{res}	$L+All_{red}$	BAU	ED_{red}	ALL_{red}	L_{re}	$L+All_{red}$
1960–1989	10(10) [2]	2(2) [1]	2(2)	5(5) [1]	0	5(5) [1]	1(1)	0	5(5)	0	1(1)	0	0	1(1)	0
2020–2029	19(21) [2]	4(4)	3(3)	14(14)	3(3)	13(13) [1]	3(3)	1(1)	9(9) [1]	0	8(9)	0	0	5(5)	0
2030–2039	31(35)	4(4) [1]	5(5) [1]	22(24)	5(5)	28(31) [3]	12(12)	9(9)	18(22) [2]	5(5)	15(15)	1(1)	0	8(8)	0
2040–2049	41(49) [3]	12(12)	11(11)	29(30) [3]	8(8)	53(63) [3]	13(13)	10(11)	38(44)	3(4)	19(19)	5(5)	2(2)	14(15)	2(2)
2050–2059	63(79) [2]	19(21)	18(19)	41(53)	6(6)	60(78) [4]	14(15)	8(8)	40(48) [4]	5(6)	26(26)	5(5)	4(5)	12(13)	3(4)
2060–2069	82(108) [5]	14(14)	13(14)	45(50) [2]	5(5)	70(99) [1]	23(24)	15(15)	47(55) [1]	14(14)	27(31)	6(6)	5(5)	23(25)	3(3)
2070–2079	93(137) [6]	32(37) [1]	30(33) [1]	57(77) [4]	18(18) [1]	110(168) [4]	28(30) [2]	15(16) [1]	75(100) [3]	10(10)	31(34)	5(6)	5(6)	22(23)	5(6)

Demand/Supply Scenario	Consecutive Monthly Failures														
	4					5					6				
	BAU	ED_{red}	ALL_{red}	L_{res}	$L+All_{red}$	BAU	ED_{red}	ALL_{red}	L_{res}	$L+All_{red}$	BAU	ED_{red}	ALL_{red}	L_{re}	$L+All_{red}$
1960–1989	0(0)	0	0	0	0	0(0)	0	0	0	0	0(0)	0	0	0	0
2020–2029	0(0) [1]	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2030–2039	1(1)	0	0	1(1)	0	0	0	0	0	0	0	0	0	0	0
2040–2049	3(3)	0	0	1(1)	0	0	0	0	0	0	0	0	0	0	0
2050–2059	5(6)	1(1)	0	3(4)	0	0	0	0	0	0	0	0	0	0	0
2060–2069	4(4)	0	0	4(4)	1(1)	1(1)	1(1)	1(1)	1(1)	0	1(1)	0	0	0	0
2070–2079	9(10)	1(1)	1(1)	5(6)	0	0	0	0	0	0	0	0	0	0	0

[a]Numbers in the baseline period (1960–1989) are for the mean number of models per decade. Within each cell the first value is the number of CPDN models, the second value (in parentheses) is the total number of failures for the CPDN ensemble, and the third value (in brackets) is the total number of CMIP3 models (only indicated when different from zero).

Table 3a. Somerset Demand[a]

Demand Scenario	Consecutive Yearly Failures														
	1					2					3				
	BAU	ED_{red}	ALL_{red}	L_{res}	$L+All_{red}$	BAU	ED_{red}	ALL_{red}	L_{res}	$L+All_{red}$	BAU	ED_{red}	ALL_{red}	L_{re}	$L+All_{red}$
1960–1989	24(25) [1]	15(15) [1]	0 [1]	22(22) [1]	0	1(1)	1(1)	0	1(1)	0	0	0	0	0	0
1990–2019	33(34) [2]	17(17) [3]	1(1)	29(30) [2]	0	2(2) [1]	1(1)	0	2(2) [1]	0	0	0	0	0	0
2020–2049	64(84) [7]	36(42) [4]	0	57(69) [6]	0	10(10)	3(3)	0	10(10)	0	1	1(1)	0	1(1)	0
2050–2079	122(208) [7]	71(94) [3]	7(7)	103(161) [6]	6(6)	27(30)	10(11)	0	22(24)	0	5(5)	1(1)	0	4(4)	0

Demand Scenario	Consecutive Yearly Failures														
	4					5					6				
	BAU	ED_{red}	ALL_{red}	L_{res}	$L+All_{red}$	BAU	ED_{red}	ALL_{red}	L_{res}	$L+All_{red}$	BAU	ED_{red}	ALL_{red}	L_{re}	$L+All_{red}$
1960–1989	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1990–2019	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2020–2049	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2050–2079	2(2)	0	0	1(1)	0	0	0	0	0	0	0	0	0	0	0

[a]Number of models that fail to satisfy Somerset demand during 1–6 consecutive years for the 30-year periods shown. Within each cell the first value is the number of CPDN models, the second value (in parentheses) is the total number of failures for the CPDN ensemble, and the third value (in brackets) is the total number of CMIP3 models (only indicated when different from zero).

Table 3b. East Devon Demand[a]

Demand Scenario	Consecutive Yearly Failures														
	1					2					3				
	BAU	ED_{red}	ALL_{red}	L_{res}	$L+All_{red}$	BAU	ED_{red}	ALL_{red}	L_{res}	$L+All_{red}$	BAU	ED_{red}	ALL_{red}	L_{re}	$L+All_{red}$
1960–1989	47(49) [8]	9(9) [2]	8(8) [2]	30(30) [5]	2(2) [1]	1(1)	0	0	1(1)	0	0	0	0	0	0
1990–2019	81(97) [4]	19(19)	11(11)	61(73) [4]	9(9)	4(4)	0	0	3(3)	0	0	0	0	0	0
2020–2049	126(206) [8]	49(54) [1]	39(42) [1]	101(144) [6]	26(27)	20(21) [1]	0	0	12(12)	0	1(1)	0	0	1(1)	0
2050–2079	190(500) [12]	89(138) [3]	78(111) [2]	149(315) [10]	53(64) [1]	60(91) [1]	10(10)	5(5)	44(53)	4(4)	16(19)	0	0	7(7)	0

Demand Scenario	Consecutive Yearly Failures														
	4					5					6				
	BAU	ED_{red}	ALL_{red}	L_{res}	$L+All_{red}$	BAU	ED_{red}	ALL_{red}	L_{res}	$L+All_{red}$	BAU	ED_{red}	ALL_{red}	L_{re}	$L+All_{red}$
1960–1989	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1990–2019	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2020–2049	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2050–2079	2(2)	0	0	1(1)	0	0	0	0	0	0	0	0	0	0	0

[a]Number of models that fail to satisfy East Devon demand during 1–6 consecutive years for the 30-year periods shown. Within each cell the first value is the number of CPDN models, the second value (in parentheses) is the total number of failures for the CPDN ensemble, and the third value (in brackets) is the total number of CMIP3 models (only indicated when different from zero).

This is consistent with the results shown in Figure 2; the most extreme droughts simulated by the CMIP3 models are not as severe as those simulated by the driest CPDN models. In the case of isolated 1-year failures, there is a sixfold increase for BAU and Somerset (comparable with CPDN), and a 1.5-fold increase for East Devon (smaller than CPDN). Once again, the response of the CMIP3 ensemble to different management options is similar to the CPDN ensemble for isolated yearly failures.

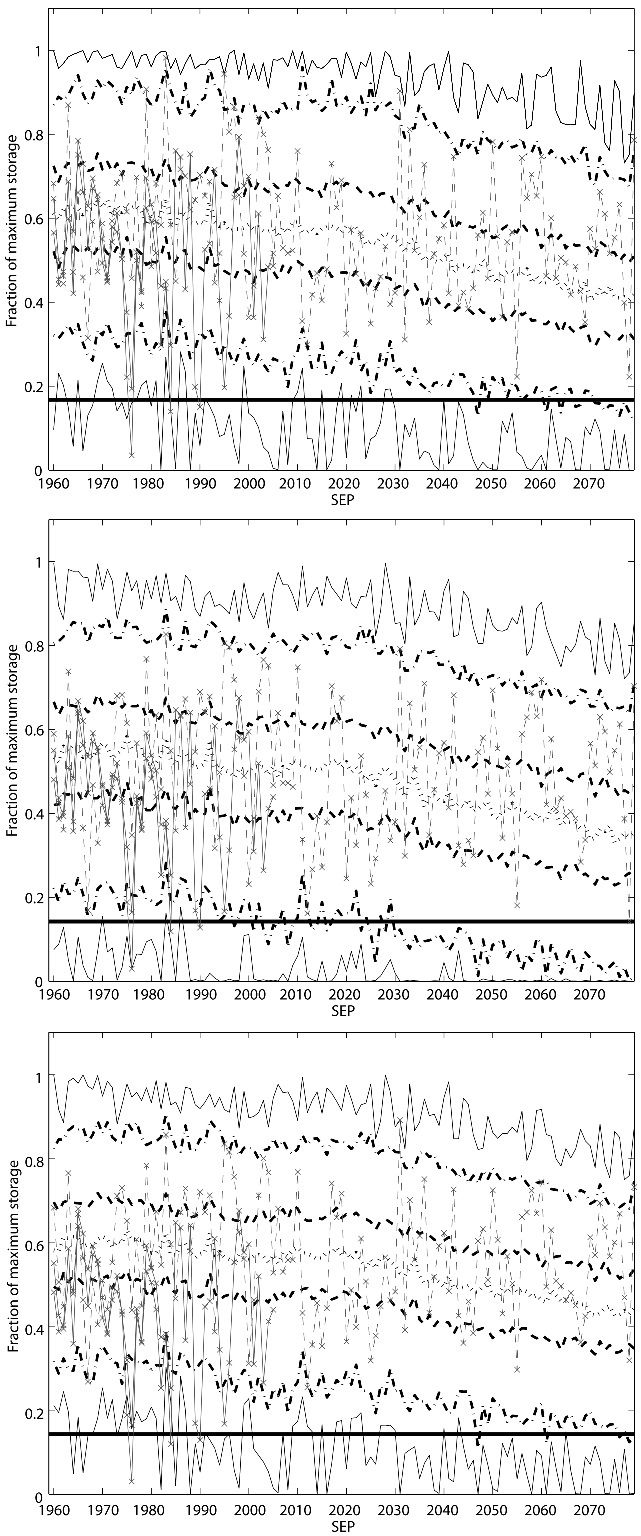

Figure 11. CPDN ensemble Wimbleball September storage levels as a function of time for (top) ALL_{red}, (middle) L_{res}, and (bottom) L+ ALL_{red} demand and supply management scenarios. The black lines represent (from top to bottom) maximum values (solid line), 97.5% (dotted-dashed line), 75% (dashed line), 50% (dotted line), 25% (dashed line), 2.5% (dotted-dashed line), and minimum values (solid line) across the climate model ensemble. The thick solid line corresponds to the control rule. Light gray and dark gray crosses connected by solid lines indicate storage levels simulated by LANCMOD using observed flows and simulated historical flows, respectively, and are included here for reference. Crosses connected by solid lines correspond to storage levels for the CPDN model run with standard values of the physical parameters.

4.2. Supply Management Options

[72] An alternative to reducing demand is to add another source of water. One way to do this in LANCMOD is to increase the size of the reservoir. Although this may not be feasible in practice for Wimbleball, increasing reservoir size is often an option, as it is relatively uncontroversial and often cost-effective. It is also an easy way to represent an additional source of water within the current model set up. Increasing the depth of the reservoir by 1m augments the storage from 21320 Ml to 25075Ml, an increase of 18% in the volume of water stored. Since we do not change any other parameter in the model, such as link capacities or control rules, the limitations in the amount of water that can be released into the system will still be controlled by these factors. However, the fact that the reservoir can store more water during the periods of high flows changes the behavior of the reservoir in relation to other scenarios. This is illustrated in Figure 11 which shows Wimbleball reservoir storage in September as a function of time for the ALL_{red} demand reduction scenario, the increased reservoir level scenario (L_{res}), and these two scenarios implemented in combination (L+ ALL_{red}).

[73] If only demand reduction is implemented, much of the effect of drier summers can be alleviated: many more CPDN models exceed any given threshold compared to the BAU scenario. For example, the storage level exceeded by half the models in the 2070s shifts from 30% under BAU to 40% under ALL_{red}. Furthermore, the risk of occurrence of very low reservoir levels across the ensemble, as indicated by the 2.5 percentile, is delayed from the 2030s under BAU until the 2070s under ALL_{red}.

[74] When the reservoir capacity is increased without reducing demand, we see that the behavior of the 50th percentile is similar to the ALL_{red} case, improving the chances of having the reservoir half full as compared with the BAU scenario. However, the risk of very low reservoir levels as indicated by the 2.5th percentile does not change significantly compared to the BAU scenario, suggesting that even though there is more storage capacity within the system, the change in river flows across the ensemble are not enough to make use of this greater capacity in the driest years. If capacity is increased and demand reduced (L+ALL_{red}) there is little improvement over ALL_{red} when looking at the behavior of the percentiles as a function of time. Obviously in absolute values, the fractions of storage level

represent a larger storage capacity for L_{res} and L+ALL_{red} (fractions of 25075Ml) than ALL_{red} (fractions of 21320Ml).

[75] Similar patterns of changes under the different supply demand scenarios are observed for the CMIP3 ensemble (not shown). In particular the lowest reservoir levels are nearly all the time above the control rule for ALL_{red} and L+ALL_{red}, and oscillate around the control rule similarly to BAU (Figure 8) for $L_{res.}$ However, as for the BAU scenario pictured in Figure 8, the data is noisier and trends within each management scenario are not as easily identifiable as for the CPDN ensemble.

[76] How the supply to different demands is affected by these two last scenarios can be assessed in Tables 2a, 2b, 3a, and 3b. In the case of Somerset demand we observe that scenarios ALL_{red} and L+ ALL_{red} are equally effective in reducing the number of consecutive monthly and annual failures, and much more effective than ED_{red} and L_{res} in isolation. ED_{red} and L_{res} are roughly equally effective until the 2040s, but the former becomes more effective thereafter, when it is clear that increasing the size of the reservoir does not improve Somerset situation significantly as compared to the BAU scenario for the CPDN ensemble. This is consistent with the previous observation that increasing the size of the reservoir does not change significantly the risk of very low reservoir level across the ensemble after the 2030s. In the case of the CMIP3 ensemble, a similar response is registered for isolated yearly failures. However, in the case of monthly failures the numbers are too small to arrive at a clear conclusion, since because of the size of the CMIP3 ensemble, any variation representing less than 5% of the ensemble (corresponding to 1 model) is meaningless. Notice that values have been rounded in the baseline period, i.e., if only 1 model fails over the 30-year period, that would give 0.33~0 listed in Tables 2 and 3.

[77] For East Devon the situation is different. Once again increasing the size of the reservoir only is not as effective in reducing failures as ALL_{red}. However, if this is combined with a reduction in all demands (L+ ALL_{red}), the number of single month failures for the CPDN ensemble reduces after the 2040s, compared with the demand reduction only scenarios. Similarly, for two or more consecutive monthly failures L+ ALL_{red} performs equally or better than ED_{red} and ALL_{red} at any time. This suggests that increasing the reservoir capacity combined with demand reduction measures can work effectively at reducing the number of consecutive monthly failures. This conclusion also holds for CMIP3 ensemble.

[78] A similar response emerges for failures in consecutive years. Increasing the reservoir size only is not very effective, suggesting than even though a large proportion of the ensemble projects wetter winters, these do not completely compensate for the drier summers. However, when a larger reservoir size is combined with demand reductions, the number of isolated annual failures for CPDN goes from 49 at the baseline for BAU, to 27 in the period 2020–2049, with a similar response for the CPDN ensemble.

[79] For 2 or more consecutive failures the demand reduction scenarios and L+ ALL_{red} are equally effective up to the middle of the 21st century, suggesting that the main limitation to reduce the number of consecutive annual failures is water availability and not storage capacity, at least until the early 2050s. Consecutive yearly failures are extremely rare in the CMIP3 ensemble; therefore no conclusion can be reached in this case with respect to the effectiveness of the different demand/supply management options.

5. Discussion

[80] In this work we have described an approach to use a large perturbed physics GCM ensemble to provide potentially useful information for the study of impacts and adaptation to climate change in a water resources system. We base our analysis on the only publicly available perturbed physics ensemble (CPDN). We also include a comparison with the results obtained using a smaller ensemble of opportunity: the CMIP3 ensemble.

[81] The hydrological model (CATCHMOD), and the water resource system model (LANCMOD) employed, are operational decision-support tools used by the Environment Agency of England and Wales. In this way, we ensure that our case study is a working example of direct relevance for current environmental planning in the United Kingdom.

[82] In the process of translating the climate model outputs into the appropriate inputs for the hydrological and water resource models, we have made some simplifying assumptions.

[83] First, we use all PPE and CMIP3 ensemble members "as is," without any previous evaluation of their relative skill in simulating the climate system. Various methodologies have been proposed to weight different model runs within a perturbed physics ensemble [*Murphy et al.*, 2007] and different GCMs within an ensemble of opportunity [*Lopez et al.*, 2006; *Tebaldi and Knutti*, 2007; *Tebaldi et al.*, 2005], or to constraint climate predictions using observations of past climate change [*Stott and Forest*, 2007]. These approaches assume that models can be weighted according to a number of possible metrics in order to obtain a meaningful probabilistic projection for the climate variables of interest. The metrics can be either global, regional/local or a combination of both. Global metrics assume that models should perform adequately at the global scale. These include for instance the climate prediction index [*Murphy et al.*, 2004]. Regional/local metrics quantify how different variables perform at the scale relevant for the climate change impact being analyzed and weight the models accordingly; an example of this being the impacts relevant climate prediction index proposed by *Wilby and Harris* [2006]. At the other end of the spectrum, some authors [*Stainforth et al.*, 2007] argue that using current observations to calibrate or weight models to produce forecast probabilities of climate change is incorrect, and misleading to the users of climate science. The underlying reason is that climate models are simulating a nonstationary system and past observations cannot possibly sample the full state space. In our work we ignore these issues and take the very practical approach of considering all stable model runs as members of our sample.

[84] Second, we have assumed that the model data for the grid box over the South West of England is a reliable input for the downscaling method used to derive local daily precipitation and PE inputs to the CATCHMOD model. Further, although the quantile-quantile transform method to bias correct/downscale precipitation has been shown to be comparable to more sophisticated methods, more work is needed to develop bias correction and downscaling methodologies that ensure the correct time and spatial correlations in

the downscaled data [*Hay et al.*, 2002; *Venema et al.*, 2006; *Wood et al.*, 2004].

[85] Our choice of these relatively simple methodologies to process the climate model data is justified by the fact that in this work we are mostly interested in analyzing how the climate model ensembles can be used in impacts/adaptation studies, and not in predicting accurately how any particular climate variable will behave in the future.

[86] Once the river flows simulated by CATCHMOD are fed into the water resource management model we observe that the reservoir operating rules that work properly under historical conditions might need some revision in the future provided that we interpret the information across the ensemble in the same way practitioners use historical information, i.e., fraction of models having a given reservoir level at any time slice in the future as fraction of time that the given reservoir level was observed in the historical record.

[87] Even more valuable perhaps is the fact that the CPDN experiment stores time series of relevant climate variables for each ensemble member. This provides time-dependent information that is internally consistent for each perturbed physics model, a feature that is extremely important for water resources management. Since this time-dependent information has been generated by a fully dynamical climate model, we assume that it provides a range of future possible paths consistent with the state of the art climate science. Therefore we can then look at relative changes in time in the reservoir storage across the model ensemble for instance, or compute changes in the frequency of occurrence of different events of interest simulated using our PPE. This is particularly relevant for multimonth and multiyear dry periods.

[88] Our results show that for this perturbed physics ensemble, the fraction of ensemble members failing to satisfy demand in the period 1960–1989 is similar to the failure frequency when using observed river flow data for 1930–2005. However, the frequency of failure increases steadily in the future under a business as usual demand and supply management scenario.

[89] To illustrate how different management options can affect this result, we analyze the response of the system under four alternative scenarios, a 15% demand reduction at East Devon, a 15% demand reduction in all major demands, an 18% increase in capacity for Wimbleball reservoir, and a final one combining the last two options. We note in passing that many other strategies could have been explored within the modeling framework, such as, increased groundwater exploitation, or changes to the consented conditions for pumped storage (i.e., volume and season of take) to better utilize the available resource during the wetter winters.

[90] The effectiveness of the different measures depends on the component of the demand that is analyzed and the planning horizon of interest. For instance, both demand management options can be quite effective in reducing the number of failures across the ensemble, particularly the ALL_{red} scenario for Somerset. In the case of East Devon, when demand management is supplemented with a larger storage capacity, the frequency of failures is even further reduced toward the end of the 21st century. For that planning horizon the CPDN ensemble indicates that larger storage capacity only is not enough to significantly reduce the frequency of failures, largely because of the lack of water available to be stored. However, a larger storage capacity combined with demand reduction ($L + ALL_{red}$) seems to be a more appropriate choice to largely reduce the possibilities of failure to supply East Devon demand.

[91] Similar responses to the different supply/demand management options are observed in the CMIP3 ensemble. However, in this case, the ensemble does not exhibit the same range of responses as the CPDN ensemble, and consequently the risk of very low reservoir levels appears reduced. Obviously if a PPE had been available for each member of the CMIP3 ensemble, the range would certainly have included the drier simulations of the CPDN models, since HADCM3 is also part of the CMIP3 ensemble. Our results suggests that in this particular case, the use of just the CMIP3 ensemble may lead to false optimism about the future, by failing to represent the possible range of modeling outcomes. It is also clear, though, that if a PPE were available for all the members of the CMIP3 ensemble these drier results might be balanced by a larger number of wetter model runs.

[92] As we have already discussed, the scenarios proposed here are simple options that can be easily implemented with LANCMOD, and allow us to describe how the system responds and whether there are any physical limitations imposed by, for instance, lack of water available to be stored in the system. In a more realistic situation the path chosen to adapt to possible impacts of climate change will depend on many other factors that include the climate information, but are not limited to that. An integrated assessment of possible adaptation strategies, taking into account issues such as the cost of the different options, their impacts on the environment, public response, technical feasibility, population and water use changes, will be the focus of future work.

[93] This study demonstrates the value added by the use of large climate model ensembles as opposed to a small number of scenarios used in impacts studies to date.

[94] It is clear that the two climate model ensembles analyzed in this work have very different characteristics. The PPE aims to quantify the uncertainty in future projections because of uncertainty in the physical parameters of one GCM. In particular CPDN aims to quantify the model uncertainty of the HADCM3 climate model. The ensemble of opportunity, CMIP3, is an ensemble of different GCMs with different model structures, and possibly different parameterizations, but does not capture the model uncertainty of each model individually.

[95] Our results indicate that an ensemble of opportunity cannot represent the range of modeling uncertainty associated with an individual climate model, even if it is only HADCM3. The additional projections provided by model uncertainty yield a wider range of possible future conditions and allow fuller exploration of adaptation options. This is particularly helpful in a reservoir system where long sequences of low flows are important. We expect that multiple projections obtained by exploring the different uncertainties inherent to climate modeling, would give a better understanding of possible long sequences and will allow considering the benefit of different management options in a way that a few deterministic scenarios would not. The overhead in additional time and expertise for carrying out the impacts analysis will be justified by the increased robustness of decisions.

[96] It is clear that adaptation to climate change is often context specific. Different sectors will have different climate

information needs and more or less sophisticated approaches to use this information for the impacts' analysis. In this sense, the United Kingdom's water sector is perhaps one of the better prepared to undertake the challenge of using large ensembles of climate models since it already posses hydrological and water resource management models potentially adaptable to the novel climate data. However, even though our case study is specific to the water sector, the key conclusions regarding the value added by the use of large climate model ensembles in impacts studies are expected to be valid for other sectors.

[97] **Acknowledgments.** We thank Nick Faull, Tolu Aina, and Milo Thurston from the CPDN team for their support in obtaining the data. This study was supported by Environment Agency Science Project SC050045 and by the Tyndall Centre for Climate Change. We thank Harriet Orr for her support as the Environment Agency project manager. 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

Arnell, N. W. (1999), Effect of global warming on water resources: Implications of multi-decadal variability and changes in year-to-year variability, report, Environ. Agency, Bristol, U. K.

Bates, B. C., Z. W. Kundzewicz, S. Wu, and J. P. Palutikof (2008), Climate change and water, technical report, 210 pp., Intergov. Panel on Clim. Change, Geneva, Switzerland.

Department for Environment, Food, and Rural Affairs (2008), Future water: The government's water strategy for England, report, London.

Ebi, K. L., R. Woodruff, A. V. Hildebrand, and C. Corvalan (2007), Climate change-related health impacts in the Hindu Kush-Himalayas, *EcoHealth*, *4*, 264–270, doi:10.1007/s10393-007-0119-z.

Environment Agency (2001), Water resources for the future: A strategy for England and Wales, report, 96 pp., London.

Environment Agency (2008), Water resources planning guideline, report, 181 pp., London.

Fowler, H. J., S. Blenkinsop, and C. Tebaldi (2007), 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.

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.

Hay, L. E., M. P. Clark, R. L. Wilby, W. J. Gutowski Jr., G. H. Leavesley, Z. Pan, R. W. Arritt, and E. S. Takle (2002), Use of regional climate model output for hydrologic simulations, *J. Hydrometeorol.*, *3*, 571–590, doi:10.1175/1525-7541(2002)003<0571:UORCMO>2.0.CO;2.

Lopez, A., C. Tebaldi, M. New, D. A. Stainforth, M. A. Allen, and J. A. Kettleborough (2006), Two approaches to quantifying uncertainty in climate change., *J. Clim.*, *19*, 4785–4796, doi:10.1175/JCLI3895.1.

Maurer, E. P., and H. G. Hidalgo (2008), Utility of daily vs. monthly large-scale climated data: And intercomparison of two statistical downscaling methods, *Hydrol. Earth Syst. Sci.*, *12*, 551–563.

Maurer, E. P., L. Brekke, T. Pruitt, and P. B. Duffy (2007a), Fine resolution climate change projections enhance regional climate change impact studies, *Eos Trans. AGU*, *88*(47), 504, doi:10.1029/2007EO470006.

Maurer, E. P., I. T. Stewart, C. Bonfils, P. B. Duffy, and D. Cayan (2007b), Detection, attribution, and sensitivity of trends towards earlier stream flow in the Sierra Nevada, *J. Geophys. Res.*, *112*, D11118, doi:10.1029/2006JD008088.

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

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. London Ser. A*, *365*, 1993–2028.

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

Panofsky, H. A., and G. W. Brier (1968), *Some Applications of Statistics to Meteorology*, Pa. State Univ. Press, University Park, Pa.

Penman, H. (1948), Natural evaporation from open water, bare soil and grass, *Proc. R. Soc. London Ser. A*, *193*, 120–146, doi:10.1098/rspa.1948.0037.

Salathe, E. P., P. W. Mote, and M. W. Wiley (2007), Review of scenario selection and downscaling methods for the assessment of climate change impacts on hydrology on the United States Pacific Northwest, *Int. J. Climatol.*, *27*, 1611–1621, doi:10.1002/joc.1540.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marqis, and K. Avery (2007), *IPCC 2007: Climate Change 2007: The Physical Science Basis. Contribution of Working Group 1 to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge Univ. Press, Cambridge, U. K.

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., M. R. Allen, E. R. Tredger, and L. A. Smith (2007), Confidence, uncertainty and decision-support relevance in climate predictions, *Philos. Trans. R. Soc. London Ser. A*, *365*, 2145–2161, doi:10.1098/rsta.2007.2074.

Stott, P. A., and C. E. Forest (2007), Ensemble climate predictions using climate models and observational constraints, *Philos. Trans. R. Soc. London Ser. A*, *365*, 2029–2052, doi:10.1098/rsta.2007.2075.

Tebaldi, C., and R. Knutti (2007), The use of multi-model ensemble in probabilistic climate projections, *Philos. Trans. R. Soc. London Ser. A*, *365*, 2053–2075, doi:10.1098/rsta.2007.2076.

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

Venema, V., S. Bachner, H. W. Rust, and C. Simmer (2006), Statistical characteristics of surrogate data based on geophysical measurements, *Nonlinear Processes Geophys.*, *13*, 449–466.

Vidal, J. P., and S. Wade (2007), A framework for developing high-resolution multi-model climate projections: 21st century scenarios for the UK, *Int. J. Climatol.*, *28*, 843–858.

Wilby, R., 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., 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., L. E. Hay, W. J. Gutowski Jr., R. W. Arritt, E. S. Takle, Z. Pan, G. H. Leavesley, and M. P. Clark (2000), Hydrological responses to dynamically and statistically downscaled climate model output, *Geophys. Res. Lett.*, *27*, 1199–1202, doi:10.1029/1999GL006078.

Wilby, R. L., P. G. Whitehead, A. J. Wade, D. Butterfield, R. J. Davies, 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.

Wood, A. W., E. P. Maurer, A. Kumar, and D. P. Lettenmaier (2002), Long-range experimental hydrological forecasting for the eastern United States, *J. Geophys. Res.*, *107*(D20), 4429, doi:10.1029/2001JD000659.

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier (2004), Hydrological implications of dynamical and statistical approaches to downscaling climate model outputs, *Clim. Change*, *62*, 189–216, doi:10.1023/B:CLIM.0000013685.99609.9e.

Young, A. R., R. Grew, and M. G. R. Holmes (2003), Low flows 2000: A national water resources assessment and decision support tool, *Water Sci. Technol.*, *48*, 119–126.

F. Fung and M. New, Tyndall Centre for Climate Change and Climate Systems and Policy Group, School of Geography and Environment, Oxford University Centre for the Environment, South Parks Road, Oxford OX1 3QY, UK.

A. Lopez, London School of Economics, University of London, Houghton Street, London WC2A 2AE, UK. (ana.lopez@ouce.ox.ac.uk)

G. Watts, Science Department, Environment Agency, Government Buildings, Burghill Road, Bristol BS10 6BF, UK.

A. Weston, Water Resources Planning, Environment Agency, Manley House, Kestrel Way, Exeter EX2 7LQ, UK.

R. L. Wilby, Department of Geography, Loughborough University, Loughborough LE11 3TU, UK.