i. Produce a large Monte Carlo sample (10^6 members) of the parameter space of surface and atmospheric processes in HadSM3, using our emulator (Section 3.2.3) to estimate multiannual mean global fields of the set of the recent climate variables identified as observational constraints in Section 3.2.9, and of the equilibrium response to doubled CO_2 for the set of variables for which future projections are required (Table 1.1), at UK land and marine points in our global climate model (downscaling is handled later in step (vi)). Uncertainties in emulated model output, observational errors and discrepancy are accounted for by sampling from their specified distributions, obtained respectively from calibration of the emulator against climate model simulations, estimates of observational errors statistics derived either from the use of alternative datasets or (where available) formal published estimates (Section 3.2.9), and the use of HadSM3 to predict the results of an ensemble of alternative climate models (see Section 3.2.8). At this stage, an interim weight is calculated for each Monte Carlo sample member, based on the recent climate observables but neglecting the Braganza *et al.* (2003) indices of historical temperature change.

ii. Sub-sample 25,000 of the 10^6 members. This is necessary because step (iii) below involves running a simple climate model, which places computational restrictions on the sample size. In selecting the 25,000 members, we use the interim weights from (i) to ensure that different parts of parameter space are sampled with a likelihood approximately consistent with their likely final contribution to the final probabilistic projections.

iii. Obtain realisations of time-dependent climate changes for the 21st century (such as those shown in Figure 3.2) by applying our timescaling technique to each of the 25,000 members from (ii). This is done by forcing our simple climate model from 1860 to 2100 with time series of historical and future forcing agents, using emulated values of regional equilibrium responses and land and ocean climate sensitivities (see Section 3.2.4), and sampling values of timescaling error, ocean heat uptake, carbon cycle feedback and sulphate aerosol forcing from the distributions described in Sections 3.2.4 and 3.2.6. Calculate the final weight to be assigned to each point in parameter space, given by the emulated values of present-day climate observables from step (i), plus the Braganza *et al*. (2003) indices measuring changes in surface temperature patterns for the period 1970–1999 relative to 1910–1939 (see Section 3.2.9).

iv. Sub-sample the 25,000 points according to the ratio of the final weights from (iii) to the interim weights from (i). This produces a final sample of 10,000 points which can be treated as a set of individual estimates of equal likelihood, based on the final weights. This further restriction of the sample size is done in order to provide a dataset which can be processed by users without placing an excessive burden on their data processing facilities.

v. Ideally, step (iv) would provide, for relevant GCM grid boxes, 10,000 samples of the joint variations between all the future variables of interest, at all times of the year (see Table 1.1), for all future periods of interest (Figure 1.3). However, such a large joint calculation is not computationally feasible, so the data are split into smaller batches. Each of the five GCM land boxes and nine marine boxes is treated separately, in two distinct batches containing different subsets of the required variables, making 28 batches in all. For a given grid box, the first batch includes all variables relating to temperature and precipitation in Table 1.1, and the additional variables required as input to the UKCP09 weather generator (with the exception of the correlation between successive daily precipitation amounts), for all times of the year and all future periods. The second batch covers the remaining variables. Within a given batch, the sampled values for different variables, months/seasons and future periods include a fully consistent treatment of covariances between both the best estimate values of the variables (driven by variations in the various climate and simple model parameters controlling the relevant physical and biogeochemical processes), and between their sampled errors. Many of these errors are actually assumed independent of one another (e.g. we assume no relationship between emulation errors, timescaling errors, observational errors or discrepancy values), however we do account for covariances between emulation errors for different variables, months (or seasons) and locations in parameter space, and between timescaling errors for different variables for a given month/season and future period. Data in different batches (e.g. projections of a given variable for a given month and future period, but at different GCM boxes), will account for physically-driven covariances between the variables, but not for the statistical error covariances identified above. The implications of handling variables from separate batches are discussed further in the UKCP09 User Guidance.

vi. Sampled climate changes for a given batch are then converted into 10,000 equiprobable Monte Carlo estimates for UKCP09 target locations (i.e. 25 km squares or aggregated regions, see Figure 1.2) using our downscaling relationships, sampling values for the regression coefficients and residuals assuming Gaussian distributions with means and variances determined from the fitting procedure described in Section 3.2.11. Joint probabilities can be estimated from these downscaled samples for changes in two or more variables in the same batch.

vii. Marginal posterior probabilities for individual climate variables for each UKCP09 target location and period are generated by a slightly different procedure. In this case, we start from probabilistic projections of the relevant variable from the appropriate GCM grid box, adjusting values associated with different probability levels of the cumulative distribution function (CDF) according to the slope and uncertainty in the appropriate downscaling relationship, and hence generating an updated CDF appropriate to the required 25 km grid box or administrative region. This procedure provides a robust numerical approximation to a full (but unfeasible) integration over the entire model parameter space.

viii. The sampled data were not considered robust either below the 1% probability level or above the 99% probability level, so we prevented the sampled data from going outside that range. That is, for a given combination of variable, location, time of year, future period and emission scenario, the values of sampled data below the 1% probability level are set to the value of the 1% probability level from the corresponding CDF, and values above the 99% probability level are set to the value of the 99% probability level. Three variables used by the weather generator (variance and skewness of daily precipitation and variance of daily temperature) are higher order statistics than the other variables, and were considered less robust; for these three variables we set the limits at the 5 and 95% probability levels.

3.2.13 Probabilistic projections for the SRES B1 and A1FI emissions scenarios

The ensemble simulations of Sections 3.2.4 and 3.2.5 are all driven by future emissions and/or concentrations of anthropogenic forcing agents consistent with the SRES A1B emissions scenario. In order to provide probabilistic projections for the B1 and A1FI scenarios, the 17 member PPE_A1B ensemble was re-run using appropriate time-dependent concentrations of greenhouse gases, and emissions of sulphate aerosol precursors. These ensembles were used to re-calibrate key timescaling statistics (specifically the correction and error terms) for the B1 and A1FI scenarios by comparing the HadCM3 simulations against timescaled estimates derived from corresponding HadSM3 simulations in conjunction with our simple climate model, as described in Section 3.2.4.

Probabilistic projections were then obtained by following the procedure of Section 3.2.12, specifying time series of forcing agents for B1 or A1FI in the simple climate model in step (iii). Apart from the timescaling aspects referred to above, all sources of uncertainty were all assumed to be the same as those specified for the A1B scenario. Some of these sources would clearly be independent of future emissions, such as emulation errors derived from our HadSM3 simulations, or the discrepancy attached to simulations of historical observables. The discrepancy for future projection variables is assumed independent of future emissions as a basic constraint of our experimental design. Further uncertainties could be specified separately for different emissions scenarios in principle, but were not in practice. These include global mean sulphate aerosol forcing, ocean heat uptake efficiency and carbon cycle feedback strengths, and regional downscaling relationships, for which resources to run additional ensemble simulations for B1 and A1FI were not available.

These assumptions are generally likely to be reasonable if global feedback strengths, and regional patterns of change per unit global warming (e.g. Mitchell, 2003), can be assumed independent of the chosen emissions scenario. Results from the latest IPCC assessment suggest that this is a reasonable assumption to leading order (e.g. Figure 10.9 of Meehl *et al*. 2007); however, our assumptions render the results for SRES B1 and A1FI somewhat less robust than those for A1B, particularly for projections in the latter decades of the 21st century, when the applied forcing and simulated response for different SRES scenarios diverges significantly (Figure 2.14).

3.3 Interpretation of UKCP09 probabilistic climate projections

UKCP09 provides a state-of-the-art basis for assessing the risk of different outcomes consistent with current climate modelling capability and understanding. However it is not yet possible to provide probabilistic projections for all variables of interest. As knowledge improves in future, the projections are liable to change.

In this chapter we have described our methodology for probabilistic projection in UKCP09, based on perturbed physics ensembles of climate model simulations specifically designed to sample uncertainties in key physical and biogeochemical processes. This is done by perturbing poorly constrained parameters in a number of configurations of one particular climate model (HadCM3), to which is added a strategy for the sampling of structural modelling uncertainties (discrepancy, explained in Section 3.2.8) by using results from one of our perturbed physics ensembles to predict the results of an alternative ensemble of climate change simulations from models developed at different climate research institutes.

Our ensemble projections are converted into probabilistic projections using a Bayesian statistical framework developed to support inference of future information about real systems from complex but imperfect models (Goldstein and Rougier, 2004; Rougier, 2007). This process allows our projections to be constrained by a set of observations of past climate (Section 3.2.9), and also involves the use of expert judgements, for example in specifying prior distributions for uncertain model parameters. The probabilities which emerge from this approach represent the relative credibility of a family of different possible outcomes, taking into account our understanding of physics, chemistry, biology, observational evidence, and expert judgement. Climate change probabilities cannot be verified in the same way as (say) probabilistic weather forecasts, because we do not have the opportunity to test our projections over many historical forecast cycles. Rather, they should be interpreted as an attempt to quantify the relative risk of different future outcomes, consistent with climate modelling technology, physical understanding and observational evidence currently available.

The credibility of the UKCP09 projections should be judged, therefore, on whether the underlying experimental design captures the leading known drivers of uncertainty, and on the extent to which the projections are robust to reasonable variations in the experimental choices and assumptions. These have been highlighted throughout the chapter, and Annex 2 contains a number of tests of key assumptions, including our expert prior distributions for model parameters, our method of estimating discrepancy, and our method of selecting the appropriate level of detail in the observational information used to constrain our projections (specifically the number of eigenvectors retained in our analysis, as explained in Section 3.2.9). This Annex also tests our results by comparing them against an approach based on a different philosophy, in which probabilities of future change are sought using a method designed to maximize the role of the constraining observations, and to be as independent as possible from the set of climate models used (e.g. Allen *et al*. 2000; Stott *et al*. 2006a).

Some of our experimental choices are not yet testable, and arise from unavoidable limitations imposed by limited scientific understanding or modelling capability. For example, while we believe that our experimental design caters for the leading known drivers of uncertainty in 21st century climate change (in particular physical atmospheric feedback processes, and carbon cycle feedbacks), there are other possible forcing agents (e.g. non-sulphate aerosol species), or feedbacks (e.g. through methane cycle processes) which are not included in UKCP09. We have no positive evidence that such factors would, if included, provide sources of uncertainty comparable with those included in UKCP09 (at least for projection time scales of a century or less), but this remains an issue for future research.

Further assumptions are imposed by limitations in computational resource. In particular, we sample uncertainties in surface and atmospheric physical processes more comprehensively than uncertainties in other earth system modules (ocean, sulphur cycle, carbon cycle), because it was not feasible to run the large ensembles of time-dependent climate change simulations which would be required. Thus we characterise uncertainties in these modules using simpler methods, applying the greater sophistication of our Bayesian calculations only to the treatment of surface and atmospheric uncertainties. In the case of the carbon cycle, however, we do make a simple attempt to account for variations in historical simulation skill between different ensemble members, and to account for structural modelling uncertainties by including results from a multi-model ensemble of projections (Friedlingstein *et al*. 2006), in addition to those from our perturbed physics ensemble.

We also assume that non-linear interactions between uncertainties in different components of the Earth System are important at the global scale, but not at the regional scale, because our finite computing resources were not able to support ensembles of climate projections with a comprehensive Earth System Model (ESM) in which uncertain processes in different components were simultaneously co-varied. Such an experiment is now in progress with HadCM3C, but UKCP09 relies on the assumption that regional interactions between earth system components are likely to be small compared with uncertainties arising when each component is sampled in isolation.

It is important that such caveats are clearly recognized. However, we believe that the UKCP09 methodology represents the most systematic and comprehensive attempt yet to provide climate projections which combine the effects of key sources of uncertainty, are constrained by a set of observational metrics representative of widely-accepted tests of climate model performance, and provide a state-of-the-art basis for the assessment of risk, within limits of feasibility imposed by current modelling capability and computing facilities.

Another key point is that we cannot make a universal assumption that probabilistic predictions can be provided for all variables that users might be interested in. As discussed in Section 3.2.10, our method is based on the assumption that robust probabilities cannot be inferred from small multi-model ensembles in isolation (see Section 3.1), and that larger perturbed physics ensembles can be used as an alternative means of sampling key process uncertainties to first order. If this is the case, then we would expect that: (a) the spread of changes simulated by the 12 member multi-model ensemble used in UKCP09 should lie more or less within that simulated by our corresponding perturbed physics ensemble; (b) even if (a) is satisfied, the discrepancy term calculated from the multi-model ensemble results should supply a modest (albeit non-trivial) component to the total uncertainty reflected in our probability distributions. With the exception of the latent heat flux variable (see Section 3.2.10), we find that criteria (a) and (b) are satisfied for the UKCP09 projection variables.

However, there were two further variables for which probabilities could not be provided, for different reasons. In the case of soil moisture content, the issue was that different models define this variable in slightly different ways, so it was not possible to calculate a discrepancy term by comparing the perturbed physics results against simulations of a consistently defined quantity in the multi-model ensemble. Secondly, it was not possible to provide probabilistic projections of fractional changes in snowfall. This is because the logarithmic transformation applied prior to our statistical calculations (in order to avoid the possibility of projecting reductions below the absolute bound of –100% — see Section 3.2.3) sometimes resulted in distributions with a highly skewed upper tail. This suggested a non-negligible probability for substantial increases in snowfall, not supported by the climate model results. This arose because the logarithm of snowfall varies rapidly at small snowfall values, and small values are often simulated in the climate model runs. This in turn means that statistical uncertainties (variances resulting from emulation error, downscaling error and timescaling error) calculated in the transformed variable tend to have large values. However our method does not account for changes in this variance as a function of the value of the projection variable, so these large variances are then assumed to apply to all projected values, leading to an unrealistic inflation of the upper tail of the attempted probabilistic projection. Changes in snowfall derived from our 11 member regional climate model ensemble projections are discussed in Chapter 4, noting that these simulations sample only a subset of the uncertainties considered in the fuller probabilistic analysis applied to other variables.

For users, an important question concerns how climate projections will change in future. Should planners make decisions now, based on estimates showing a wide range of possible changes, or should they delay in the hope that more precise information will be available in (say) 10 yr time? On the one hand, modellers have striven successfully to improve their models over the past decade or so (e.g. Reichler and Kim, 2008), yet the range of future global projections in the IPCC AR4 (Meehl *et al*. 2007) was not significantly narrower than in the previous IPCC assessment, and the range of projected changes over the UK has certainly not narrowed. On the other hand, some of the errors in climate models tend to be systematic across different models, partly due to shared features such as limited resolution. Examples, including a tendency to underestimate the frequency of blocking anticyclones over Europe in winter, are given in Annex 3. The presence of common errors gives rise to the possibility that ensemble climate projection exercises of the future might give different results to those deriving from the current generation of models, at least for some aspects of climate.

In practice, therefore, the prospects for better projections will depend on which variables or which future periods users are most interested in. For example, uncertainties in the UKCP09 projections are substantial even for a couple of decades ahead (Sections 4.4.2 and 4.5), due to the significant influence of internal variability at regional scales, and then grow larger through the 21st century due to the additional influence of uncertain climate change feedbacks (Box 2.1). Prospects for reducing uncertainties in near-term changes are likely to rest mainly on constraining projections of internal variability by initializing climate models with ocean observations (Smith *et al*. 2007; Keenlyside *et al*. 2008), and through improvements in the ability of models to simulate regional modes of variability. For example, increased horizontal or vertical resolution might lead to better simulation of features such as the North Atlantic storm track, or the coupling between sea surface temperature anomalies and atmospheric circulation anomalies. At longer lead times progress would also depend on improvements in our ability to represent thermodynamic climate feedbacks and carbon cycle processes, and their complex interactions. An active dialogue between users and climate research scientists will therefore be crucial, in order to ensure that adaptation decisions are taken on the basis of up-to-date information concerning the potential for emerging research to update projections currently available, such as UKCP09.

As mentioned above, improvements in climate models are one potential route to improved projections in future. By *improved*, we mean both more comprehensive sampling of climate feedbacks (through the use of comprehensive ESMs), and smaller uncertainties through the development of models with higher resolution and better representations of sub-grid scale processes. Initialisation of climate models with observations (also mentioned above) has potential to improve projections of near-term climate over the next decade or so, and possibly longer. Uncertainties could also be reduced by developments in experimental design, subject to available computing resources. For example, future exercises of this type could potentially be based entirely on simulations in which the atmosphere model is coupled to a full dynamical ocean component, rather than a simple mixed layer ocean (see Section 3.2.3). This would remove the need for scaling approaches to infer time-dependent climate changes from equilibrium changes, and hence narrow the probability distributions significantly, as our timescaling procedure is responsible for a significant component of the total uncertainty captured in our probabilities (see Annex 2). It would also allow a wider range of observational metrics to be used in constraining the projections.

In summary, the UKCP09 projections should be seen as a comprehensive summary of possible climate futures consistent with understanding, models and resources available at present, but users should be aware that the projections could change in future, as the basis for climate prediction evolves over time.

3.4 References

Allen, M. R., Stott, P. A., Mitchell, J. F. B., Schnur, R. & Delworth, T. L. (2000). Quantifying the uncertainty in forecasts of anthropogenic climate change. *Nature*, **407**, 617–620.

Allen, M. R. & Stainforth, D. A. (2002). Towards objective probabilistic climate forecasting. *Nature*, **419**, 228.

Anderson, T. L. *et al*. (2003). Climate forcing by aerosols — a hazy picture. *Science*, **300**, 110–111.

Andreae, M. O., Jones, C. D. & Cox, P. M. (2005). Strong present-day aerosol cooling implies a hot future. *Nature*, **435**, 1187–1190.

Archer, D. & Buffett, B. (2005). Time-dependent response of the global ocean clathrate reservoir to climatic and anthropogenic forcing. *Geochemistry, Geophysics, Geosystems*, **6**, Qo3002. (doi:10.1029/2004GC000854).

Betts, R. A., Cox, P. M., Collins, M., Harris, P., Huntingford, C. & Jones C. D. (2004). The role of ecosystem–atmosphere interactions in simulated Amazonian precipitation decrease and forest dieback under global climate warming. *Theoretical and Applied Climatology*, **78**, 157–175.

Boer, G. J. & Yu, B. (2003). Dynamical aspects of climate sensitivity. *Geophysical Research Letters*, **30**, 1135.

Bony, S. *et al*. (2006). How well do we understand and evaluate climate change feedback processes? *Journal of Climate*, **19**, 3445–3482.

Braganza, K., Karoly, D. J., Hirst, A. C., Mann, M. E., Stott, P., Stouffer, R. J. & Tett, S. F. B. (2003). Simple indices of global climate variability and change: Part I — variability and correlation structure. *Climate Dynamics*, **20**, 491–502.

Brohan, P., Kennedy, J. J., Harris, I., Tett, S. F. B. & Jones, P. D. (2006). Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850. *Journal of Geophysical Research*, **111**, D12106.

Christensen, T. R. *et al*. (2004). Thawing sub-arctic permafrost: Effects on vegetation and methane emissions. *Geophysical Research Letters*, **31**, doi:10.1029/2003GL018680.

Collins, M., Booth, B. B. B., Harris, G., Murphy, J. M., Sexton, D. M. H. & Webb, M. J. (2006). Towards quantifying uncertainty in transient climate change. *Climate Dynamics*, **27**, 127–147.

Collins, M., Brierley, C. M., MacVean, M., Booth, B. B. B. & Harris, G. R. (2007). The sensitivity of the rate of transient climate change to ocean physics perturbations. *Journal of Climatology*, **20**, 2315–2320.

Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A. & Totterdell, I. (2000). Acceleration of global warming due to carbon cycle feedbacks in a coupled climate model. *Nature*, **408**, 184–187.

Cox, P. M. (2001). Description of the TRIFFID dynamic global vegetation model. Technical Note 24, Hadley Centre, Met Office, UK.

Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A. & Totterdell, I. J. (2001). Modelling vegetation and the carbon cycle as interactive elements of the climate system. In: Meteorology at the millennium. Pearce, R. (Ed). Academic Press, 259–279.

Essery, R. L. H., Best, M. J., Betts, R. A., Cox, P. M. & Taylor, C. M. (2003). Explicit representation of subgrid heterogeneity in a GCM land-surface scheme. *Journal of Hydrometeorology*, **4**, 530–543.

Forster, P. *et al*. (2007). Changes in atmospheric constituents and in radiative forcing. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M. & Miller, H. L. (Eds.) Cambridge University Press, Cambridge, UK and New York, NY, USA.

Friedlingstein, P. *et al*. (2006). Climate–carbon cycle feedback analysis: Results from the C4MIP model intercomparison. *Journal of Climatology*, **19**, 3337–3353.

Furrer, R., Sain, S. R., Nychka, D. & Meehl, G. A. (2007). Multivariate Bayesian analysis of atmosphere–ocean general circulation models. *Environmental Ecological Statistics*, **14**, 249–266.

Goldstein, M. & Rougier, J. C. (2004). Probabilistic formulations for transferring inferences from mathematical models to physical systems. *SIAM Journal of Scientific Computing*, **26**, 467–487.

Gordon, C., Cooper, C., Senior, C. A., Banks, H. T., Gregory, J. M., Johns, T. C., Mitchell, J. F. B. & Wood, R. A. (2000). The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics*, **16**, 147–168.

Greene, A. M., Goddard, L. & Lall, U. (2006). Probabilistic multimodel regional temperature change projections. *Journal of Climate*, **19**, 4326–4346.

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

Harris, G., Sexton, D. M. H., Booth, B. B. B., Collins, M., Murphy, J. M. & Webb, M. J. (2006). Frequency distributions of transient regional climate change from perturbed physics ensembles of general circulation model simulations. *Climate Dynamics*, **27**, 357–375.

Haywood, J. & Shulz, M. (2007). Causes of the reduction in uncertainty in the anthropogenic radiative forcing of climate between IPCC (2001) and IPCC (2007). *Geophysical Research Letters*, **34** (doi:10.1029/2007/GL030749).

Jacob, D. *et al*. (2007). An intercomparison of regional climate models for Europe: design of the experiments and model performance. *Climatic Change*. (doi:10.1007/s10584-006-9213-4).

Johns, T. C. *et al*. (2003). Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emissions scenarios. *Climate Dynamics*, **20**, 583–612.

Johnson, C. E., Stevenson, D. S., Collins, W. J. & Derwent, R. G. (2001). Role of climate feedback on methane and ozone studied with a coupled Ocean-Atmosphere-Chemistry model. *Geophysical Research Letters*, **28**, 1723–1726.

Jones, A., Roberts, D. L., Woodage, M. J. & Johnson, C.E. (2001). Indirect sulphate aerosol forcing in a climate model with an interactive sulphur cycle. *Journal of Geophysical Research*, **106**, 20293–20310.

Jones, C. D., Cox, P. M., Essery, R. L. H., Roberts, D. L. & Woodage, M. A. (2003). Strong carbon cycle feedbacks in a climate model with interactive CO_2 and sulphate aerosols. *Geophysical Research Letters*, **30**, 1479.

Jones, G. S., Jones, A., Roberts, D. L., Stott, P. A. & Williams, K. D. (2005). Sensitivity of global-scale climate change attribution results to inclusion of fossil fuel black carbon aerosol. *Geophysical Research Letters*, **32**, L14701.

Jones, R. G., Murphy, J. M. & Noguer, M. (1995). Simulation of climate change over Europe using a nested regional climate model. In: Assessment of control climate, including sensitivity to location of lateral boundaries. *Quarterly Journal of the Royal Meteorological Society*, **121**, 1413–1449.

Keenlyside, N. S., Latif, M., Jungclaus, J., Kornblueh, L. & Roeckner, E. (2008). Advancing decadal-scale climate prediction in the North Atlantic sector. *Nature*, **453**, 84–88.

Kettleborough, J. A., Booth, B. B. B., Stott, P. A. & Allen, M. R. (2007). Estimates of uncertainty in predictions of global mean surface temperature. *Journal of Climate*, **20**, 843–855.

Knutti, R., Meehl, G. A., Allen, M. R. & Stainforth, D. A. (2006). Constraining climate sensitivity from the seasonal cycle in surface temperature. *Journal of Climate*, **19**, 4224–4233.

Lambert, S. J. & Boer, G. J. (2001). CMIP1 evaluation and intercomparison of coupled climate models. *Climate Dynamics*, **17**, 83–106.

McKay, M. D., Beckman, R. J. & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, **21**, 239–245.

Meehl, G. A., Covey, C., McAvaney, B., Latif, M. & Stouffer, R. J. (2005). Overview of the coupled model intercomparison project. *Bulletin of the American Meteorological Society*, **86**, 89–93.

Meehl, G. A. *et al*. (2007). Global Climate Projections. In: *Climate Change 2007: The Physical Science Basis.Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M. & Miller, H. L. (Eds.) Cambridge University Press, Cambridge, UK and New York, NY, USA.

Mitchell, T. D. (2003). Pattern scaling: an examination of the accuracy of the technique for describing future climates. *Climatic Change*, **60**, 217–242.

Moberg, A. & Jones, P. D. (2004). Regional climate model simulations of daily maximum and minimum near-surface temperatures across Europe compared with observed station data 1961–1990. *Climate Dynamics*, **23**, 695–715.

Murphy, J. M., Sexton, D. M. H., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M. & Stainforth, D. A. (2004). Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, **429**, 768–772.

Murphy, J. M., Booth, B. B. B., Collins, M., Harris, G. R., Sexton, D. M. H. & Webb, M. J. (2007). A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles. *Philosophical Transactions of the Royal Society A*, **365**, 1993–2028.

Palmer, J. R. & Totterdell, I. J. (2001). Production and export in a global ocean ecosystem model. *Deep Sea Research Part I: Oceanographic Research Papers*, **48**, 5, 1169–1198.

Piani, C., Frame, D. J., Stainforth, D. A. & Allen, M. R. (2005). Constraints on climate change from a multi-thousand member ensemble of simulations. *Geophysical Research Letters*, **32**, L23825.

Pittock, A. B., Jones, R. N. & Mitchell, C. D. (2001). Probabilities will help us plan for climate change. *Nature*, **413**, 249.

Randall, D. A. *et al*. (2007). Climate Models and Their Evaluation. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M. & Miller, H. L. (Eds.) Cambridge University Press, Cambridge, UK and New York, NY, USA.

Raper, S. C. B., Gregory, J. M. & Stouffer, R. J. (2002). The role of climate sensitivity and ocean heat uptake on AOGCM transient temperature response. *Journal of Climate*, **15**, 124–130.

Reichler, T. & Kim, J. (2008). How well do coupled models simulate today's climate? *Bulletin of the American Meteorological Society*, **89**, 303–311.

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

Rougier, J. C. & Sexton, D. M. H. (2007). Inference in ensemble experiments. *Philosophical Transactions of the Royal Society A*, **365**, 2133–2144.

Rougier, J. C., Sexton, D. M. H., Murphy, J. M. & Stainforth, D. A. (2009). Analysing the climate sensitivity of the HadSM3 climate model using ensembles from different but related experiments. *Journal of Climate* (doi:10.1175/2008JCLI2533.1, in press).

Rowell, D. P. & Jones, R. G. (2006). Causes and uncertainty of future summer drying over Europe. *Climate Dynamics* **27**, 281–299.

Rowntree, P. R. (1998). Global average climate forcing and response since 1750. *International Journal of Climatology*, **18**, 355–377.

Sabine, C. L. *et al*. (2004). The oceanic sink for anthropogenic CO_2. *Science*, **305**, 367–371.

Schlesinger, M. E., Andronova, N. G., Entwistle, B., Ghanem, A., Ramankutty, N., Wang, W. & Yang, F. (1997). Modelling and simulation of climate and climate change. In: *Past And Present Variability Of The Solar-Terrestrial System*. Catagnoli, G. A. and Provenzale, A. (Eds.) International School of Physics Enrico Fermi, Vol 133. IOS Press, Amsterdam, 389–429.

Smith, D. M., Cusack, S., Colman, A. W., Folland, C. K., Harris, G. R. & Murphy, J. M. (2007). Improved surface temperature prediction for the coming decade from a global climate model. *Science*, **317**, 796–799.

Soden, B. J. & Held, I. M. (2006). An assessment of climate feedbacks in coupled ocean-atmosphere models. *Journal of Climate*, **19**, 3354–3360.

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

Stott, P. A., Mitchell, J. F. B., Allen, M. R., Delworth, T. L., Gregory, J. M., Meehl, G. A. & Santer, B. D. (2006a). Observational constraints on past attributable warming and predictions of future global warming. *Journal of Climate*, **19**, 3055–3069.

Stott, P. A., Kettleborough, J. A. & Allen, M. R. (2006b). Uncertainty in continental-scale temperature predictions. *Geophysical Research Letters*, **33**, L02708.

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, **106**, 7183–7192.

Tebaldi, C. & Knutti, R. (2007). The use of the multi-model ensemble in probabilistic climate projections. *Philosphical Transactions of the Royal Society A*, **365**, 2053–2076.

Tebaldi, C., Smith, R. W., Nychka, D. & Mearns, L. O. (2005). Quantifying uncertainty in projections of regional climate change: A Bayesian approach to the analysis of multi-model ensembles. *Journal of Climate*, **18**, 1524–1540.

Watterson, I. G. (2008). Calculation of probability density functions of temperature and precipitation change under global warming. *Journal of Geophysical Research,* **113**. (doi:10.1029/2v007JD009254).

Webb, M. J. *et al*. (2006). On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles. *Climate Dynamics*, **27**, 17–38.

Wilby, R. L. *et al*. (2004). Guidelines for the use of climate scenarios developed from statistical downscaling methods, IPCC Task Group on Data and Scenario Support for Impact and Climate Analysis (TGICA). http://ipcc-ddc.cru.uea.ac.uk/guidelines/StatDown_Guide.pdf.