PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A

Phil. Trans. R. Soc. A (2007) **365**, 1957–1970
doi:10.1098/rsta.2007.2068
Published online 14 June 2007

Ensembles and probabilities: a new era in the prediction of climate change

By Mat Collins*

Met Office, Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK

Predictions of future climate are of central importance in determining actions to adapt to the impacts of climate change and in formulating targets to reduce emissions of greenhouse gases. In the absence of analogues of the future, physically based numerical climate models must be used to make predictions. New approaches are under development to deal with a number of sources of uncertainty that arise in the prediction process. This paper introduces some of the concepts and issues in these new approaches, which are discussed in more detail in the papers contained in this issue.

Keywords: climate change; prediction; uncertainty; probability

1. Introduction and motivation

The existence of human-induced climate change is now established. The evidence base is made up of theoretical understanding of greenhouse gases and their radiative effects in the atmosphere, high-quality observations of the climate system over time and physically based modelling (the subject of this issue), which has been used to interpret those observations. What next then for climate change science? There is an emerging political will to take action. What is the role for the science of climate prediction[1] in enabling action?

We may divide action on climate change into two classes: adaptation actions and mitigation actions. In order for society to adapt to climate change which is happening and to which we are already committed, detailed predictions of future climate variability and change (whether they be human-induced or natural in origin) are needed. In many cases, information is required for local areas—how high should we build new sea defences, what should be the capacity of a new reservoir? Adaptation planning horizons are generally (although not exclusively) of the order of years to decades and may involve dependencies on other non-climate factors. In contrast, mitigation actions are often considered to involve long time-scale questions. At what level must CO_2 concentrations be stabilized to avoid dangerous climate change, possibly involving

*matthew.collins@metoffice.gov.uk

[1] Some prefer the word projection to prediction to indicate that there are always some assumptions that are made in assessing future climate change. Some choose to label the probabilistic assessments as predictions and the single-model experiments as projections. It is this nomenclature that is adopted here.

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

the crossing of thresholds in the system, which may trigger catastrophes such as the collapse of the Atlantic Ocean overturning circulation, or the death of the Amazon rainforest?

Such questions are complex to answer. They sit at the intersection of a number of disciplines: atmospheric physics, oceanography, ecosystems science, etc. Notwithstanding the issues of nonlinearity and chaos in fluid flows such as the atmosphere, the coupling between physical and biological processes means that it is very difficult (if not impossible) to develop an elegant theory for answering the key questions posed by society. We are therefore left with the rather inelegant, but at least practical, solution of developing numerical models of the climate system that represent the major processes and their interactions and using those models to make predictions.

The very fact that a team of people can take the fundamental laws of fluid flow, electromagnetic radiation, etc., approximate the equations, discretize them on a grid, code them up on a computer and produce a simulation that bears a passing resemblance to the world we live in is, in retrospect, a significant feat. Yet, perfectionism dominates and we are quick to point out deficiencies in our simulations of climate. These deficiencies, often called model errors (although perhaps better called as model imperfections—they are not simple coding slip-ups that can be easily corrected but rather fundamental limitations that arise owing to finite computer-processing capacity) reduce our confidence in the projections we make of future climate with those models. Model imperfections, coupled with fundamental limitations on the initial-value prediction of chaotic weather and the unknown path that society may take in terms of future emissions of greenhouse gases, imply that it is not possible to be certain about future climate.

The purpose of this issue is to showcase some of the new approaches in dealing with these uncertainties in climate prediction. The general methodology is to use ensembles of projections that attempt to sample uncertain inputs and processes, and are then used to produce predictions in the form of the probability of different outcomes. The use of ensembles and probabilities arose in weather forecasting some years ago (Palmer & Hagedorn 2006). Ensembles, in which different initial conditions are sampled (recognizing the sensitive dependence on initial conditions in atmospheric flows), are now routinely produced and used in risk-based assessment of weather impacts. More recently, model uncertainties are also sampled. Hence, the extension to the climate problem is a natural one. Nevertheless, the key distinction is that for the weather forecasting problem, the ensemble prediction system may be verified over many cycles and may be corrected and improved. For the climate change prediction, no such verification is possible—a problem that has caused much debate about the best way to make predictions.

While this issue contains papers from many of the authors who have contributed to the development of ensemble and probabilistic techniques, the list is by no means exhaustive. References to other studies may be found in the individual papers and there are certainly a number of papers in press and to be published in future. The issue begins with four longer papers that are a mixture of reviews of previous work and 'position' papers that describe certain approaches to the problem. As suggested above, there has been considerable debate regarding the most appropriate methodological approach, and it is hoped that this collection of papers will aid the informed reader in interpreting the strengths and weaknesses of those various methodologies. The shorter papers contain more original research (although some authors do not shy away from methodological issues) and are indicative of the fast pace of development in this area. This introductory article is designed to set the scene and draw out some of the major and common points contained in the papers.

2. The hierarchy of climate models

Climate models, at their heart, solve equations derived from physical laws, for example, the nonlinear Navier–Stokes equations of fluid flow. This is in contrast with some predictive models that may be purely based on the fitting of (perhaps complex) functions to data. Empirical approaches are not valid for climate change prediction as such models could not be reliably used to make extrapolations outside the historical training period (some authors even level this criticism at physically based models). Owing to finite computing capacity, approximations must be made to the equations and because they must be solved using computational techniques, certain processes must be simplified or *parameterized* for reasons of practicality (and here an element of empiricism can creep in). We may identify the following three classes of models for the purpose of this issue.

(i) *Energy balance models* (EBMs) are very low-order models that may only solve equations for global or hemispheric mean quantities. They are computationally very cheap and are therefore can be used to produce large ensembles in which key uncertain parameters are comprehensively sampled. They are useful for testing hypotheses about techniques and can be used to approximate the output of more complex models. They cannot provide information at regional scales or provide information on quantities such as winds that are not part of their simplified equation set.

(ii) *Earth system models of intermediate complexity* (EMICs) are more complex and hence more computationally expensive, but nevertheless are usually cheap enough for large ensembles to be produced (Stott & Forest 2007). They may, for example, be formulated by dimensionality reduction of the atmosphere and/or ocean (from three spatial to two spatial dimensions) or by applying a coarse numerical grid. They provide a useful step between the EBM and highly complex global climate model (GCM), although many also include approximations of biological processes and feedbacks which have only been included in a limited number of complex GCMs—hence the 'Earth system'.

(iii) *Global climate models* or global/general circulation models (GCMs) are the most complex of climate models (Huebener *et al.* 2007). Resolving the three-dimensional flow of the ocean and atmosphere in the finest grid practicable and including representations of unresolved sub-grid-scale processes, they are the most computationally expensive. Standard supercomputer resources can be employed to produce ensembles with tens to hundreds of members (Murphy *et al.* 2007; Tebaldi & Knutti 2007): much larger ensembles have only been possible using highly novel computing techniques (Frame *et al.* 2007; Stainforth *et al.* 2007*b*). They are the only way of providing predictions at regional scales and for 'exotic' variables such as extreme precipitation, for example.

It may sometimes be hard to pigeonhole a particular model. For the purpose of ensemble climate prediction, it is important to consider what measures of climate change are the emergent properties of the model and what must be specified as, or trivially related to, an input parameter. In EBMs, for example, it is usual to specify the climate sensitivity (the global mean temperature change for a doubling of atmospheric CO_2—a beloved measure of climate researchers), whereas for a GCM, the climate sensitivity is a function of the interaction between resolved and parameterized physical processes and cannot be specified *a priori*. In the former case, it is possible to assume something about the distribution of the climate sensitivity, but in the latter case, it is not (or at least it is not easy). Another important consideration is whether the model is computationally cheap enough to produce a complete mapping of input parameters to output parameters, or does some novel sampling technique have to be employed?

3. Drivers of uncertainty

Model predictions of climate change are uncertain for the following three principal reasons.

(*a*) *Natural variations in climate*

The climate system has internal variability (an oft-quoted example being the El Niño phenomenon) that arises via the accumulation of variations in day-to-day weather and interactions between components of the system (the ocean and atmosphere in the case of El Niño). This class of uncertainty is often labelled as initial condition uncertainty, in keeping with the idea of the sensitive dependence on initial conditions encountered in chaotic nonlinear systems. It is certainly not possible, even given a perfect model and perfect initial conditions, to predict the timing of a particular weather system or El Niño event 30 years into the future. Rather, we seek to quantify the uncertainty in the prediction that arises owing to these random unpredictable events.

It may, however, be possible to predict the phase of some longer time-scale variations in some aspects of the system, perhaps the ocean meridional overturning circulation or heat content. This has led to the idea of initializing climate models with observations (Troccoli & Palmer 2007) which, while still being at an early stage, is a candidate for quantifying uncertainty in near-term predictions, in comparison with the usual approach of assuming an envelope of natural variability that surrounds the signal of climate change.

(*b*) *Uncertain pathways for forcing agents such as greenhouse gas emissions*

This involves factors outside the realm of climate science: diverse areas of research such as economics, population dynamics and 'green technologies' to name but a few. Formal quantification of future anthropogenic forcing agents has been said to be impossible, leading many to dismiss the idea that we will ever be able to produce probabilistic predictions with all uncertainties quantified. However, the problem is not as grave as it might first appear. For adaptation questions, the inertia of the climate system means that predictions are relatively insensitive (at least at global scales) to the precise details of future emissions over the next few decades (Stott & Forest 2007). For mitigation questions, the emissions scenario may considered in some sense to be the independent variable. That is to say, given an emissions pathway, the problem is one of calculating the risk of exceeding a threshold.

The problem of uncertain natural forcing agents, principally large volcanic eruptions and variations in solar luminosity, is a real one. While the timing of eruptions and the variations in the Sun on time-scales longer than the quasi-repeatable 11-year cycle may not be predictable in a deterministic sense, they can at least be treated by statistical methods. Nevertheless, most studies have yet to include their potential influence.

(*c*) *Uncertainty in feedback processes in models, or put more simply, uncertainty in models*

It is this uncertainty that has driven the development of ensemble and probabilistic techniques in climate prediction. Different climate models produce significantly different projections for the same emission scenario, both in terms of the magnitude of global mean changes and in terms of the detailed spatial and temporal patterns of change. That is not to say, however, that there is a complete divergence of projections and 'anything goes'. There is some commonality of responses, particularly in terms of the broad-scale patterns of warming and even some regional patterns of precipitation change.

In the past, the range of different model projections has been interpreted as the bound of uncertainty in the prediction, but this is far from ideal when one is, for example, considering how big to build a reservoir. Is the collection of the world's climate models an adequate sample of the space of all possible models (and, indeed, is it even possible to define such a space)? Should all models be considered equally probable when assessing uncertainty? These and other questions have been central in framing the development of the techniques outlined in this issue.

4. Ensembles of climate models

It is common to sub-divide model uncertainty further and consider *parameter uncertainty*, i.e. uncertainty in the parameters that control the parameterized physical processes in climate models and *structural uncertainties*, i.e. uncertainties in choices made when coding the resolved processes (although we note that the two are inexorably linked).

Given resolved (or large-scale) variables such as temperature and humidity at any time, sub-grid-scale parameterization schemes calculate the impact of unresolved processes such as clouds and give back a partial time derivative (or tendency) for the large-scale fields. For some processes, such as the transfer of electromagnetic radiation through the atmosphere, it is possible to calculate the tendencies quite accurately using line-by-line radiation codes. However, owing to computational constraints, the line-by-line codes have to be approximated, so that they do not take an excessive amount of the total processing time for the whole model. Other processes like turbulent energy transport are much less well understood and are highly situation-dependent, meaning that their parameterization is a field in itself within numerical weather and climate modelling, relying on a mixture of theoretical understanding and empirical fitting to observational studies or high-resolution processes model simulations. In terms of the parameters in climate models, some parameters have direct counterparts in the physical world and may be measured and are thus subject to measurement errors, while others may be simplifications of processes and are thus subject to uncertainties associated with their representivity.

Given uncertain parameters, an obvious approach is to perturb those parameters and measure the impact on the projection of climate change. The so-called *perturbed physics ensembles* have formed a cornerstone of a number of approaches to quantifying uncertainty and there are a number of papers in this issue which use them. The full hierarchy of models has been employed from the computationally efficient EBMs and EMICs for which it is possible to fully map model parameter space, to complex GCMs for which it is not. The problem is considerable, as complex GCMs contain of the order of tens to hundreds of uncertain parameters. Moreover, those parameters are not independent and their interdependence is non-trivial; there are complex interactions between physical processes such as clouds and radiation, for example. The main advantage of the perturbed physics approach is that it is at least theoretically possible to estimate the uncertainty in climate change prediction across all plausible values of uncertain parameters (the idea is explored further in Rougier & Sexton 2007). In addition, it is possible to use efficient techniques for sampling the parameter space of a model (Annan & Hargreaves 2007). More practically, the research group has tight control over the ensemble generation in terms of the specification of both inputs and outputs, enabling the testing of specific hypotheses.

In some perturbed physics ensembles, different parts of complex model code have been switched on and off. However, in general, it is not possible to sample all the so-called structural choices made in designing a GCM. Choices are numerous and examples include the specification of the numerical solver for dynamical fluid flow, the parameterization of convective processes and the representation of sea ice. The production of an ensemble that adequately samples all the possible (or even most commonly used) combinations of structural choices in GCMs would require a significant world-wide investment in research: an investment many orders of magnitude more than that currently spent on climate research. Nevertheless, effort has been expended to coordinate experimental design and the output from the 20 or so complex climate models in existence. This *multi-model ensemble* has the advantage that it does sample structural uncertainties, although the sampling is not systematic and, moreover, different models share common components and published algorithms for solving aspects of the resolved and unresolved processes. Notwithstanding this, the ensemble is an invaluable guide to the magnitude of uncertainties and has been used, in itself, to produce probabilistic predictions (Min *et al.* 2007; Tebaldi & Knutti 2007).

A further possibility yet to be applied in the context of predicting long-term climate change is the so-called *stochastic parameterization* or stochastic physics. Recognizing that there may not be a smooth invertible relationship between the grid-scale resolved variables and the sub-grid-scale unresolved processes, modellers have introduced random perturbations to sub-grid-scale tendencies. These have been shown to improve both the mean simulation characteristics of models and short-term weather forecast skill and more sophisticated representations are currently under development. How such schemes would influence climate change predictions is still an open question.

The above discussion generally applies to uncertainties in the 'known–knowns' of climate modelling, principally those physical processes associated with ocean, land, sea ice and atmospheric feedbacks. There are other processes or known–unknowns like biological feedbacks associated with the earth's carbon cycle which have been included in a number of computationally efficient models and a handful of complex GCM models but are not generally implemented as a matter of course. Other known–unknowns like potential feedbacks associated with the release of methane from sub-ocean reservoirs of methane hydrates have yet to be studied in any significant way by modelling centres. To complete the list, there are the unknown–unknowns, i.e. processes and feedbacks that have yet to be even discovered (or rediscovered). For those we have no information (by definition), but it is fair to say that it is very unlikely that any major feedbacks have been missed so far.

5. Production of probabilistic predictions

As ensemble and probabilistic prediction matured in weather forecasting, it has been possible to compare different types of ensemble and different perturbation techniques by examining verification scores over large numbers of forecasting cases. Such verification can then feed directly back to improve the forecasting system. As stated above, this is not practical in climate prediction.

The basic approach to generating probability density functions (PDFs) can be characterized using Bayes' theorem

$$p(s|\text{data}) \propto p(s)p(\text{data}|s), \tag{5.1}$$

where data is some collection of observed climate variables; s is the prediction variable of interest; and p denotes the probability. The formula says that the posterior probability of s is proportional to the prior probability of s (without reference to observed data) multiplied by the likelihood, the probability of obtaining the data for a given value of s. In practice, the application of the formula can get very complicated and more details can be found in the papers in this issue and the references therein. However, equation (5.1) is useful in discussing some of the generic issues of generating PDFs.

The role of the *prior* has been much discussed. In the presence of a strong concentration of probability density via a tight likelihood function or observational constraint (see below), the prior distribution of s should have little impact on the posterior predictive distribution. However, this is not generally the case (so far) in climate change prediction, so the prior distribution can have a leading-order influence on the prediction. One obvious way of calculating the prior would be to draw from the 'space of all possible models'. Owing to the complexity of climate models, the definition of a metric of distance between models cannot (or at least has yet to) be defined.

Some authors have defined the prior distribution in the perturbed physics ensemble case in terms of the output of the ensemble computed for some assumed (usually expert specified) distributions for the range of the model parameters

(Murphy *et al.* 2007). The calculation of such distributions is computationally prohibitive and the concept of an *emulator*, which estimates the relationship between the model inputs and the model outputs, has been developed (Rougier & Sexton 2007). It is argued, however, that this is open to the risk of 'double counting' of the data as the experts must have had some idea of how certain parameter settings would have performed in a model–data comparison. This has led to the idea of uniform, non-informative or flat priors (Frame *et al.* 2007). Other studies (often those which use reduced-complexity models in which the prediction variable is also a specified input parameter) use prior distributions specified by experts or taken from previous studies (Stott & Forest 2007). As the choice is important, some authors choose to present PDFs for different prior assumptions and leave the reader to decide which is most appropriate. Other authors (Stainforth *et al.* 2007*a*) suggest that it is a barrier to the production of any type of PDF.

This introductory article does not attempt to pass judgement on which approach is best. It is simply pointed out that this is a strongly argued point and there are some firmly held beliefs. It may take some time to resolve such issues (they may never be resolved), and this is, and will continue to be, a stumbling block to the acceptance of probabilistic climate change prediction.

The role of the *likelihood* is in reweighting the prior distribution probabilities to produce the final prediction probabilities. This is desirable as it can have the feature of downweighting the less-believable model predictions and upweighting the more believable model predictions to reduce the total uncertainty. The obvious difficulty is that now we do not know which model predictions of, e.g. 2050, climate are any better than others. Hence, the challenge is to develop weighting techniques that are based on the ability of ensemble members to reproduce some feature of the observed mean climate, natural climate variability or forced climate change.

Some studies have used rather simple metrics based on a small number of climate variables, the advantage being that they are at least simple to understand, but the problem being that they can suffer from the problem of 'getting the right answer for the wrong reason', in the sense that a model which has a good present day simulation of, for example, surface air temperature trends may do so owing to some cancellation of errors in some other closely related fields (Min *et al.* 2007). Studies that attempt to use a large number of fields in the definition of the metric should suffer less from this potential error but may suffer the double counting problem if care is not taken to use truly independent constraints. Under the assumption that the observed sensitivity of the warming attributable to increasing greenhouse gases will be the same in future, it is possible to use studies that estimate this quantity in the likelihood weighting (Stott & Forest 2007). However, if this assumption is violated because of some nonlinearity in the system or because the rate of change of greenhouse gas concentrations changes, that approach is not valid. Other elements of weighting methodologies that have been proposed include, for example, criteria based on how similar projections from different models are (called convergence; Tebaldi & Knutti 2007).

The concept of an *observational constraint* is an alternative way of presenting the notion of likelihood weighting. Following Piani *et al.* (2005), we can illustrate the concept using a schematic (figure 1). Assume that we are in possession of an

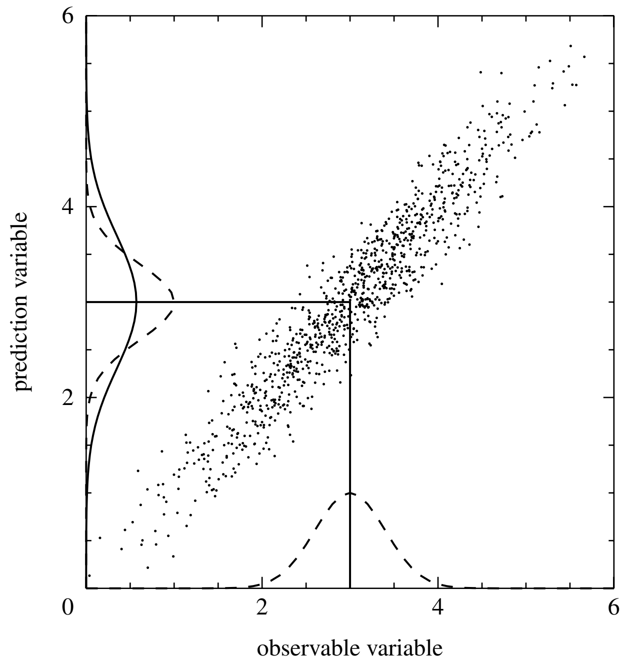

Figure 1. A schematic example of a climate prediction constrained by observations. An ensemble of climate model simulations of the past and the future is searched to find a relationship between the variable to be predicted (the y-axis) and a variable that might be observed (x-axis). The relationship is then used to translate the observational estimate into a prediction, taking into account uncertainty in the observation (dashed distributions on both axes) and the relationship (the solid distribution on the y-axis).

ensemble of climate model simulations (a perturbed physics, multi-model or even a collection of both) for both past and future conditions. We may search the database of model output to find observable variables that are related (in the schematic in a linear way) to the variable we are interested in predicting. We have then hopefully found a relationship between the past (which we can observe) and the future (which we wish to predict). Given an observation of the real world, we can simply read off the value of the future prediction variable using the relationship (as indicated by the vertical and horizontal lines in the figure). Of course, the observational part will be uncertain, possibly owing not only to random observation error but also complexities in comparing models with observed data. Hence, we use the relationship to translate this uncertainty (as shown by the dashed distribution on the x-axis) into uncertainty in the prediction (as shown by the dashed distribution on the y-axis). Furthermore, because the relationship between the observable and the prediction variable is not perfect, we must inflate the uncertainty in the prediction to account for this 'statistical uncertainty' (as shown by the solid distribution on the y-axis).

It is immediately obvious then how to improve (i.e. how to further constrain) predictions. One can take better observations that are less uncertain, or one can reduce the 'statistical' uncertainty by finding better relationships between observables and things we wish to predict. The latter is essentially a signal-to-noise maximization problem for which there are standard statistical approaches (canonical correlation analysis, for example). The problem is that because the dimension of the problem needs to be reduced to make it tractable, the observable may end up being some rather complex function of many different variables, perhaps projected onto some variance-maximizing set of basis vectors, and hence difficult to unravel and understand. Further problems may arise if the observations are biased, and of course, the approach is predicated on the ensemble producing a realistic representation of the relationship between the past and the future. Nevertheless, the approach does facilitate the quantitative comparison of different observational constraints beyond the usual expert judgement of the 'goodness' of different models.

Of course, models are certainly not (and never will be) perfect representations of the real world. The problem of the 'perfect model assumption' is often ignored, although not because it is assumed that models are perfect, rather it is just realized to be a very difficult issue to deal with. However, there has been some progress recently in attacking the problem using perturbed physics ensembles by defining the concept of the 'best model', i.e. by imagining that there is a set of model parameters to give us the best historical (observed and unobserved) simulation of climate and the most probable prediction. The distance between this best model and real world is termed the *discrepancy*. In practice, it is impossible to find this set of parameters (recall that we cannot observe the future); however, it may be possible to approximately estimate the degree of uncertainty it might add to predictions and to test the sensitivity of the predictions to different choices.

6. Predictions for users

If any of the ensemble and probabilistic techniques discussed in this issue are of any worth, they must produce information that can be directly used by society in making decisions. Government agencies, businesses and private individuals are not often concerned with equilibrium global climate sensitivity but rather with much more local climate variability and change and with systems for which climate is one of many controlling influences.

The downscaling of climate information to the local scale may be achieved through the use of embedded high-resolution dynamical models derived from complex GCMs (Murphy *et al.* 2007) or by the training of statistical techniques based on past observed data. The former is technically complex and there are a number of issues to do with the way the physical parameterization schemes operate at higher spatial resolution for example, that may introduce further uncertainties. The latter is simpler to implement, but empirical relationships derived using past data may be hard to find owing to the lack of observations or instability of relationships between large- and small-scale variables. Such present-day empirical relationships may not even be appropriate to the future.

There is a significant area of research in the impacts or applications of climate change which takes the raw climate information and uses further modelling techniques to produce societal-relevant information (New *et al.* 2007; Stainforth *et al.* 2007*b*). Impacts models are diverse and examples are available in the prediction of river flow and other hydrological variables, crop yield and energy demand. The key issue to note is that there is a 'cascade of uncertainty' from the large GCM scales, through the local scales and down to the impact variable of interest and hence it is important to consider fully the uncertainties at all stages of the process, including those uncertainties in the impacts models themselves. Some impact variables may be simply related to climate variables, so a 'response surface' approach whereby climate uncertainty is mapped directly onto impact variable uncertainty may be possible. In the case of complex relationships between impact and climate variables, or in cases of other controlling factors, ensembles of impacts models must be driven using direct output from GCMs or from downscaling models or 'weather generators'. The quantification of uncertainty in impact variables is still in its infancy, but we can expect to see more studies in the future.

Another challenge is in educating the users of climate change predictions so that they can make full use of ensemble or probabilistic information. The use of scenarios has pervaded climate change prediction and it may take a while for many to get out of the habit. However, the assessment of risk based on probabilities is a task that most of us undertake regularly in everyday life, and we are seeing a number of users who are actively requesting quantitative information about uncertainty in predictions.

7. The future

The science of ensemble and probabilistic climate change prediction is still rather new and there are many methodological issues to sort out. We are yet to routinely produce PDFs of user-specified variables that are actually used to make decisions on a day-to-day basis. However, we do believe that this is an important area to make progress in, so that we may move forward from the simple approach of presenting indicative scenarios of what *might* happen in the future. Climate models will always be imperfect and owing to their inherent three-dimensional nature, modest increases in model spatial resolution have only been possible in the lifetime of climate prediction. High-resolution climate models that can be used for prediction in which sub-grid-scale processes like convection are explicitly resolved rather than parameterized are still decades away: decades during which climate change will intensify and people will have to adapt.

In the future, we expect to see more studies that produce estimates of probabilities at sub-global and even regional local scales. We also expect to see quantitative comparisons of different methods and approaches to better address which are the most appropriate. Specific areas are likely to be

- comparisons of perturbed physics ensembles generated using structurally different models;
- comparison and combination of different observational constraints;
- more use of probabilistic methods in impacts assessments; and
- 'seamless' prediction of weather and climate from days to seasons to decades to centuries.

The application of the approaches described in this short article and in the papers in this issue also provides a framework by which climate models and predictions may be improved. The derivation of observational constraints for

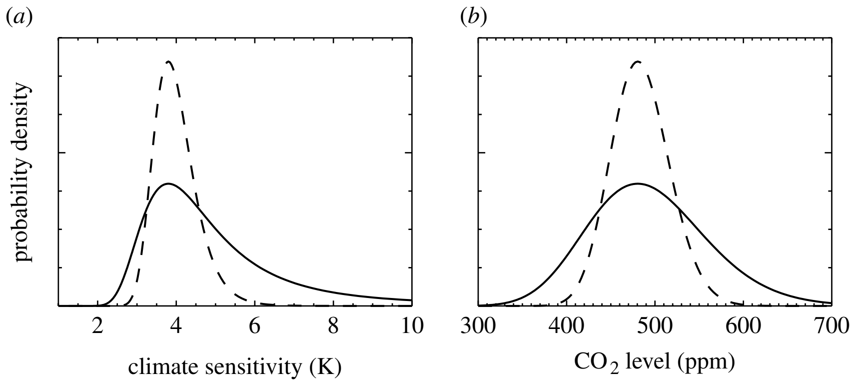

Figure 2. A simple example of the importance of quantifying and reducing uncertainty in climate change predictions. (*a*) Distributions of climate sensitivity produced from assumed normally distributed estimates of the climate feedback parameter (equation (7.1)). The dashed line has half the uncertainty in feedback parameter when compared with the solid line. (*b*) Distributions of the CO_2 stabilization level that would have to be achieved in order to limit global warming 3 K, given the distributions of climate sensitivity in (*a*). For the solid line, the lower 5 percentile is at 386 ppm; for the reduced uncertainty dashed line, it is at 431 ppm.

example (figure 1) allows us to understand which observable variables are of most use in constraining predictions. Thus, we might target new or more accurate observations of climate and target model development for those specific variables. For many, the approaches will seem excessively statistical, making it difficult to conceptually understand how a prediction is constrained by a set of observations. It is therefore crucial that we explore sensitivities and unravel some of the methods in order to advance understanding in climate change.

The ultimate goal, of course, is to reduce uncertainty in climate predictions. This will be done by improving climate models, taking better observations and, now, by implementing and improving probabilistic prediction techniques. This is more than just an academic exercise however. Every uncertain climate prediction is a possible risk to society that we might have to plan to adapt to or work to prevent. If that risk can be eliminated through better, more certain predictions, then there is a saving in real terms.

A simple illustration follows. Assume that, after all our efforts, we have produced an estimate of the uncertainty in the climate feedback parameter

$$\lambda = \frac{Q}{\Delta T}, \tag{7.1}$$

defined for an equilibrium climate with a radiative forcing of Q divided by the temperature change ΔT, which is normally distributed with a mean of $1.0\ \mathrm{W\,m^{-2}\,K^{-1}}$ and a 5–95% range of 0.6–$1.4\ \mathrm{W\,m^{-2}\,K^{-1}}$. This is translated simply to a right-skewed distribution for climate sensitivity (owing to the inverse relationship in equation (7.1)) that has a 5–95% range of 2.7–6.3 K (figure 2*a*). Now, say, that we wish to stabilize atmospheric CO_2 levels so as to avoid a global mean temperature rise, ΔT, of 3 K above pre-industrial values as that is deemed to be some 'dangerous' level of climate change. The radiative forcing, Q, for doubled CO_2 is approximately $3.8\ \mathrm{W\,m^{-2}}$ and is known to vary logarithmically with CO_2 concentration. It is straightforward to write

$$C_{\text{stab}} = C_{\text{pre}} \exp\left(\frac{\Delta T \lambda \ln 2}{Q}\right), \tag{7.2}$$

where C_{stab} is future and the stabilized level of CO_2 and C_{pre} is the pre-industrial level. Thus, in order to be 95% sure of not passing 3 K of global warming (or alternatively, accepting a 5% risk that we do), the CO_2 must be stabilized at 1.4 pre-industrial levels or around 386 ppm by volume (figure 2*b*). This is approximately the CO_2 concentration today.

However, assume that we might improve our estimate of the feedback parameter, producing a distribution with the same mean value but with half the 5–95% range (0.8–1.2 $\text{W m}^{-2}\,\text{K}^{-1}$). Then, to be 95% sure that we do not exceed 3 K warming, we only have to stabilize at 431 ppm (figure 2*b*). We have bought society some time to develop technologies to limit emissions and allowed economic development for poorer societies.

This is, of course, a simple example. However, the numbers are indicative of many studies in the literature and the 3 K threshold exceeds the value of 2 K global warming that has been considered to be dangerous. What it is meant to illustrate is the importance of being able to measure the uncertainty in climate change predictions, so that we know that we have improved them in the future.

The work was supported by the UK Department of the Environment, Food and Rural Affairs under contract PECD/7/12/37 and by the European Community ENSEMBLES (GOCE-CT-2003-505539) project under the Sixth Framework Programme. Many of the leading authors of the papers in the issue provided insightful comments on a draft.

References

Annan, J. D. & Hargreaves, J. C. 2007 Efficient estimation and ensemble generation in climate modelling. *Phil. Trans. R. Soc. A* **365**, 2077–2088. (doi:10.1098/rsta.2007.2067)

Frame, D. J., Faull, N. E., Joshi, M. M. & Allen, M. R. 2007 Probabilistic climate forecasts and inductive problems. *Phil. Trans. R. Soc. A* **365**, 1971–1992. (doi:10.1098/rsta.2007.2069)

Huebener, H., Cubasch, U., Langematz, U., Spangehl, T., Nierhorster, F., Fast, I. & Kunze, M. 2007 Ensemble climate simulations using a fully coupled ocean–troposphere–stratosphere general circulation model. *Phil. Trans. R. Soc. A* **365**, 2089–2101. (doi:10.1098/rsta.2007.2078)

Min, S.-K., Simonis, D. & Hense, A. 2007 Probabilistic climate change predictions applying Bayesian model averaging. *Phil. Trans. R. Soc. A* **365**, 2103–2116. (doi:10.1098/rsta.2007.2070)

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

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

Palmer, T. N. & Hagedorn, R. (eds) 2006. *Predictability of weather and climate*. Cambridge, UK: Cambridge University Press.

Piani, C., Frame, D. J., Stainforth, D. A. & Allen, M. R. 2005 Constraints on climate change from a multi-thousand member ensemble of simulations. *Geophys. Res. Letts.* **32**, L23 825. (doi:10.1029/2005GL024452)

Rougier, J. & Sexton, D. S. 2007 Inference in ensemble experiments. *Phil. Trans. R. Soc. A* **365**, 2133–2143. (doi:10.1098/rsta.2007.2071)

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

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

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

Tebaldi, C. & Knutti, R. 2007 The use of the multi-model ensemble in probabilistic climate projection. *Phil. Trans. R. Soc. A* **365**, 2053–2075. (doi:10.1098/rsta.2007.2076)

Troccoli, A. & Palmer, T. N. 2007 Ensemble decadal prediction from analysed initial conditions. *Phil. Trans. R. Soc. A* **365**, 2179–2191. (doi:10.1098/rsta.2007.2079)