

Hierarchical number of the contribution: Theme

The Philosophy of Climate Science

Richard Bradley and Roman Frigg and Katie Steele

*Department of Philosophy, Logic and Scientific Method,
London School of Economics and Political Science, UK*

Erica Thompson

*Centre for the Analysis of Time Series,
London School of Economics and Political Science, UK*

Charlotte Werndl

*Department of Philosophy, University of Salzburg, AUT
and*

*Department of Philosophy, Logic and Scientific Method,
London School of Economics and Political Science, UK*

Authors are listed alphabetically. This work is fully collaborative.

Keywords

Definition of climate, climate change, climate models, attribution, detection, natural variability, fingerprinting, IPCC, confirmation, prediction, projection, model ensemble, observation, tuning, calibration, values in science, uncertainty, climate change economics, discounting the future, expert disagreement, expected utility theory, imprecise probability, decision making under uncertainty/disagreement, robust decisions, precautionary principle.

Contents

1. Introduction
 2. Defining Climate and Climate Change
 3. Climate Models
 4. Detection and Attribution of Climate Change
 5. Confirmation and Predictive Power
 6. Understanding and Quantifying Uncertainty
 7. Conceptualising Decisions Under Uncertainty
 8. Managing Uncertainty
 9. Conclusion
- References

Summary

The aim of this article is to provide an up-to-date summary of the main problems and questions in the foundations of climate science. We first discuss the problem of defining climate. After having introduced climate models, we review problems in connection with detecting and attributing climate change. We examine the confirmation of climate models and the limits of predictability, and then review classifications of uncertainty and the use of model ensembles. After these science-oriented topics we turn to decision theory. We discuss the framing of climate decision problems and offer an examination of alternative decision rules.

1. Introduction

Climate science is an umbrella term referring to scientific disciplines studying aspects of the earth's climate. It includes, among others, parts of atmospheric science, oceanography, and glaciology. In the wake of public discussions about an appropriate reaction to climate change, parts of decision theory and economics have also been brought to bear on issues of climate, and thus contributions from these disciplines can be considered part of climate science broadly construed. At the heart of the philosophy of climate science lies a reflection on the methodology used to reach various conclusions about how the climate may evolve and what we should do about it. The philosophy of climate science is a new sub-discipline of the philosophy of science that began to crystalize at the turn of the century, when philosophers of science started having a closer look at methods used in climate modelling. It now comprises a reflection on almost all aspects of climate science, including observation and data, methods of detection and attribution, model ensembles and decision-making under uncertainty. Since the devil is always in the detail, the philosophy of climate science operates in close contact with science itself and pays careful attention to the scientific details. For this reason, there is no clear separation between climate science and the philosophy thereof, and conferences in the field are often attended by both scientists and philosophers.

The aim of this article is to provide an up-to-date summary of the main problems and questions in the foundations of climate science. In Section 2 we talk about the problem of defining climate. In Section 3 we introduce climate models. In Section 4 we discuss the problem of detecting and attributing climate change. In Section 5 we examine the confirmation of climate models and the limits of predictability. In Section 6 we review classifications of uncertainty and the use of model ensembles. In Section 7 we turn to decision theory and discuss the framing of climate decision problems. In Section 8 we introduce alternative decision rules. In Section 9 we offer a few conclusions.

Two qualifications are order. First, we review issues and questions that arise in connection with climate science from a philosophy of science perspective, and with special focus on epistemological and decision theoretic problems. Needless to say, this is not the only perspective. Much can be said about climate science from other points of view, most notably science studies, sociology of science, political theory, and ethics. For want of space we cannot review contributions from these fields.

Second, to guard against possible misunderstandings, it ought to be pointed out that engaging in a critical philosophical reflection on the aims and methods of climate science – the objective of this chapter – is in no way tantamount to adopting a position known as *climate scepticism*. Climate sceptics are a heterogeneous group of people who do not accept the results of 'mainstream' climate science, encompassing a broad spectrum from those who flat out deny the basic physics and of the greenhouse effect (and the influence of human activities on the world's climate) to a small minority who actively engage in scientific research and debate and reach conclusions at the lowest end of climate impacts. Critical philosophy of science is not the handmaiden of climate scepticism; nor are philosophers *ipso facto* climate sceptics. So it should be stressed here that we don't endorse climate scepticism. We aim to understand how climate science works, reflect on its methods, and understand the questions that it raises.

2. Defining Climate and Climate Change

Climate talk is ubiquitous, in the popular media as well as in academic discourse, and climate change has become a homely topic. This veils the fact that *climate* is a complex concept and that the correct definitions of climate and climate change are a matter of controversy. To gain an understanding of the notion of climate, it is important to distinguish it from weather. Intuitively speaking, the weather at a particular place and a particular time is the state of the atmosphere at that place and at the given time. For instance, the weather in central London at 2pm on 1 January 2015 can be characterised by saying that the temperature is 12 centigrade, the humidity 65%, etc. By

contrast, climate is an aggregate of weather conditions: it is a *distribution* of certain variables (called the climate variables) arising for a certain configuration of the climate system.

The question is how to make this basic idea precise, and this is where different approaches diverge. Current approaches to defining climate can be divided into two groups: those that define climate as a *distribution over time*, and those that define climate as an *ensemble distribution*. The climate variables in both approaches include those that describe the state of the atmosphere and the ocean, and sometimes also variables such as those describing the state of glaciers and ice sheets [IPCC 2013].

Distribution over time. The state of the earth depends on *external conditions* of the system such as the amount of energy received from the sun and volcanic activity. Assume that there is period of time over which the external conditions are relatively stable in that they exhibit small fluctuations around a constant mean value c . One can then define the climate over this time period as the distribution of the climate variables over that period under constant external conditions c [e.g. Lorenz 1995]. Climate change then amounts to successive time periods being characterised by different distributions. However, in reality the external conditions are not constant and even when there are just slight fluctuations around c , the resulting distributions may be very different. Hence this definition is unsatisfactory [Werndl 2014].

This problem can be avoided by defining climate as the empirically observed distribution over a certain period of time, where external conditions are allowed to vary. Again, climate change amounts to different distributions for successive time periods. This definition is popular because it is easy to estimate from the observations, e.g., from the statistics taken over thirty years that are published by the World Meteorological Organisation [Hulme et al. 2009]. A major problem of this definition can be illustrated by the example where in the middle of a period of time the Earth is hit by a meteorite and thus becomes a much colder place. Clearly, the climate before and after the hit of the meteor differ. Yet this definition has no means to require that such a regime change cannot occur because all it says is that climate is a distribution arising over a certain time period.

To circumvent this problem, Werndl [2015] introduces the idea of regimes of varying external conditions and suggests defining climate as the distribution over time of the climate variables arising under a certain regime of varying external conditions. The challenge for this account is to spell out what exactly is meant by a regime of varying external conditions.

Ensemble Distribution. An ensemble of climate systems (not to be confused with a model ensemble to which we turn below) is an infinite collection of virtual copies of the climate system. Consider the sub-ensemble of these that satisfy the condition that the present values of the climate variables lie in a certain interval around the *values measured* in the actual climate system (i.e. the values compatible with the measurement accuracy). Now assume again that there is period of time over which the external conditions are relatively stable in that they exhibit small fluctuations around a constant mean value c . Then climate at future time t is often defined as the distribution of values of the climate variables that arises when all systems in the ensemble evolve from now to t under constant external conditions c [e.g., Lorenz 1995]. In other words, the climate in the future is the distribution of the climate variables over all possible climates that are consistent with current observations under the assumption of constant external conditions c .

However, in reality external conditions are not constant and even when there are just small fluctuations around a mean value, this can lead to different distributions [Werndl 2014]. This worry can be addressed by tracing the development of the initial condition ensemble under *actual* external conditions. The climate at future time t then is the distribution of the climate variables that arises

when the initial conditions ensemble is evolved forward for *the actual path taken by the external conditions* [e.g. Daron and Stainforth 2013].

This definition faces a number of conceptual challenges. First, it makes the world's climate dependent on our knowledge (via measurement accuracy), but this is counterintuitive because we think of climate as something objective that is independent of our knowledge. Second, the above definition is a definition of *future* climate, and it is difficult to see how the present and past climate should be defined. Yet without a notion of the present and past climate one cannot define climate change. A third serious problem is that ensemble distributions do not relate to the past time series of observations and this would imply that the climate cannot be estimated from them [cf. Werndl 2015].

These considerations show that defining climate is nontrivial and there is no generally accepted or uncontroversial definition of climate.

3. Climate Models

A climate model is a representation of certain aspects of the climate system. One of the simplest climate models is an energy-balance model, which treats the earth as a flat surface with one layer of atmosphere above it. It is based on the simple principle that in equilibrium the incoming and outgoing radiation must be equal (see Dessler [2011], Chapters 3-6, for a discussion of such models). This model can be refined by dividing the earth into zones, allowing energy transfer between zones, or describing a vertical profile of the atmospheric characteristics. Despite their simplicity, these models provide a good qualitative understanding of the greenhouse effect.

Modern climate science aims to construct models that integrate as much as possible of the known science (for an introduction to climate modelling see [McGuffie and Henderson-Sellers 2005]). Typically this is done by dividing the earth (both the atmosphere and ocean) into grid cells. At the time of writing, state-of-the-art global climate models had a horizontal grid scale of around 150km. Climatic processes can then be conceptualised as flows of physical quantities such as heat or vapour from one cell to another. These flows are mathematically described by equations. These equations form the 'dynamical core' of a global circulation model (GCM). The equations typically are intractable with analytical methods and powerful supercomputers are used to solve them. For this reason they are often referred to as 'simulation models'. To solve equations numerically, time is discretised. Current state-of-the-art simulations use time steps of approximately 30 minutes, taking weeks or months in real time on supercomputers to simulate a century of climate evolution.

In order to compute a single hypothetical evolution of the climate system (a 'model run'), we also require *initial conditions* and *boundary conditions*. The former are a mathematical description of the state of the climate system (projected into the model's own domain) at the beginning of the period being simulated. The latter are values for any variables which affect the system but which are not directly output by the calculations. These include, for instance, the concentration of greenhouse gases in the atmosphere at a given time, the amount of aerosols, and the amount of solar radiation received by the earth. Since these are drivers of climatic change, they are often referred to as *external forcings* or *external conditions*.

Where processes occur on a smaller scale than the grid, they may be included via *parameterisation*, whereby the net effect of the process is separately calculated as a function of the grid variables. For instance, cloud formation is a physical process which cannot be directly simulated because typical clouds are much smaller than the grid. So the net effect of clouds is usually parameterised (as a function of temperature, humidity, etc.) in each grid cell and fed back into the calculation. Sub-grid processes are one of the main sources of uncertainty in climate models.

There are currently around twenty climate models, which are under continuous development by national modelling centres like NASA, the UK Met Office, and the Beijing Climate Center. In order to be able to compare outputs of these different models, the Coupled Model Intercomparison Project (CMIP) defines a suite of standard experiments to be run for each climate model. One standard experiment is to run each model using the historical forcings experienced during the twentieth century, so that the output can be directly compared against real climate system data.

Climate models are used in many places in climate science, and their use gives rise to important questions. We discuss these questions in next three sections.

4. Detection and Attribution of Climate Change

Every empirical study of climate has to begin by observing the climate. Meteorological observatories measure a number of variables such as air temperature near the surface of the Earth using thermometers. But more or less systematic observations are available only since about 1750, and hence to reconstruct the climate before then scientists have to rely on *proxy data*: data for climate variables that are derived from observing other natural phenomena such as tree rings, ice cores and ocean sediments.

The use of proxy data raises a number of methodological problems centred around the statistical processing of such data, which are often sparse, highly uncertain, and several inferential steps away from the climate variable of interest. One particular study, a proxy-based reconstruction of the Northern Hemisphere temperature record [Mann, Bradley and Hughes, 1998], gave rise to a heated debate known as the *Hockey Stick controversy*. The sceptics pursued two lines of argument. They cast doubt on the reliability of the available data, and they argued that the methods used to process the data are such that they would produce a hockey-stick-shaped curve from almost any data [e.g., McIntyre and McKittrick 2003]. The papers published by the sceptics raised important issues and stimulated further research, but were found to contain serious flaws undermining their conclusions. There are now more than two dozen reconstructions of this temperature record using various statistical methods and proxy data sources. Although there is indeed a wide range of plausible past temperatures, due to the constraints of the data and methods, these studies do robustly support the consensus that temperatures during the late 20th century are likely to have been the warmest in the past 1400 years [Frank et al. 2010].

Do rising temperatures indicate that there is climate change and if so, can the change be attributed to human action? These two problems are known as the problems of *detection* and *attribution*. The Intergovernmental Panel on Climate Change (IPCC) defines these as follows:

“Detection of change is defined as the process of demonstrating that climate or a system affected by climate has changed in some defined statistical sense without providing a reason for that change. An identified change is detected in observations if its likelihood of occurrence by chance due to internal variability alone is determined to be small” [...]. Attribution is defined as “the process of evaluating the relative contributions of multiple causal factors to a change or event with an assignment of statistical confidence”.’ [IPCC 2013]

These definitions raise a host of issues. The root cause of the difficulties is the clause that climate change has been detected only if an observed change in the climate is unlikely to be due to internal variability. *Internal variability* is the phenomenon that climate variables such as temperature and precipitation would change over time due to the internal dynamics of the climate system even in the absence of climate change: there have been (and would be) hotter and colder years even if there were no humans at all.

Taken at face value, this definition of detection has the consequence that there cannot be internal climate change. The ice ages, for instance, would not count as climate change if they occurred because of internal variability. This is not only at odds with basic intuitions about climate and with the most common definitions of climate as a finite distribution over a relatively short time period (where internal climate change is possible); it also leads to difficulties with attribution: if detected climate change is *ipso facto* change *not* due to internal variability, then from the very beginning it is excluded that certain factors (namely, internal climate dynamics) can lead to a change in the climate, which seems to be an unfortunate conclusion.

For the case of the ice ages many would stress that internal variability is different from natural variability. Since, say, orbital changes explain the ice ages, and orbital changes are natural but external, this is a case of external climate change. While this move solves some of the problems, there remains the problem that there is no generally accepted way to separate internal and external factors, and the same factor is sometimes classified as internal and sometimes as external. For instance, glaciation processes are sometimes treated as internal factors and sometimes as prescribed external factors. Likewise, sometimes the biosphere is treated as an external factor, but sometimes it is also internally modelled and treated as an internal factor. One could even go so far to ask whether human activity is an external forcing on the climate system or an internally-generated earth system process. Research studies usually treat human activity as an external forcing, but it could consistently be argued that human activities are an internal dynamical process. The appropriate definition simply depends on the research question of interest.

The effects of internal variability are present on all timescales, from the sub-daily fluctuations experienced as weather to the long-term changes due to cycles of glaciation. Since internal variability is the outcome of a highly complex nonlinear system, it is also unlikely that the statistical properties of internal variability are constant over time, which further compounds methodological difficulties. State-of-the-art climate models run with constant forcing show significant disagreements both on the magnitude of internal variability and the timescale of variations. (On <http://www.climate-lab-book.ac.uk/2013/variable-variability/#more-1321> the reader finds a plot showing the internal variability of all CMIP5 models. The plot indicates that models exhibit significantly different internal variability, leaving considerable uncertainty.) The model must be deemed to simulate pre-industrial climate (including variability) sufficiently well before it can be used for such detection and attribution studies, but we do not have thousands of years of detailed observations upon which to base that judgement. Estimates of internal variability in the climate system are produced from climate models themselves [Hegerl et al. 2010], leading to potential circularity. This underscores the difficulties in making attribution statements based on the above definition, which recognises an observed change as climate change only if it is unlikely to be due to internal variability.

Since the IPCC's definitions are widely used by climate scientists, the discussion about detection and attribution in the remainder of this section is based on these definitions. Detection relies on statistical tests, and detection studies are often phrased in terms of the likelihood of a certain event or sequence of events happening in the absence of climate change. In practice, the challenge is to define an appropriate null hypothesis (the expected behaviour of the system in the absence of changing external influences), against which the observed outcomes can be tested. Because the climate system is a dynamical system with processes and feedbacks operating on all scales, this is a non-trivial exercise. An indication of the importance of the null hypothesis is given by the results of Cohn and Lins [2005], who compare the same data against alternate null hypotheses, with results differing by 25 orders of magnitude of significance! This does not in itself show that either null is more appropriate, but it demonstrates the sensitivity of the result to the null hypothesis chosen. This, in turn, underscores the importance of the choice of null hypothesis and the difficulty of making any such choice if the underlying processes are poorly understood.

In practice the best available null hypothesis is often the best available model of the behaviour of the climate system, including internal variability, which for most climate variables usually means a state of the art GCM. This model is then used to perform long *control runs* with constant forcings in order to quantify the internal variability of the model (see discussion above). Climate change is then said to have been detected when the observed values fall outside a predefined range of the internal variability of the model. The difficulty with this method is that there is no single “best” model to choose: many such models exist, they are similarly well developed, but, as noted above, they have appreciably different patterns of internal variability.

The differences between different models are relatively unimportant for the clearest detection results such as recent increases in global mean temperature. Here, as stressed by Parker [2010] detection is robust across different models (for a discussion of robustness see Section 6), and, moreover, there is a variety of different pieces of evidence all pointing to the conclusion that the global mean temperature has increased. However, the issues of which null hypothesis to use and how to quantify internal variability, can be important for the detection of subtler local climate change.

If climate change has been detected, then the question of attribution arises. This might be an attribution of any particular change (either a direct climatic change such as increased global mean temperature, or an impact such as the area burnt by forest fires) to any identified cause (such as increased CO₂ in the atmosphere, volcanic eruptions, or human population density). Where an indirect impact is considered, a two-step or multi-step approach may be appropriate, first attributing a climate change to some forcing agent, and then an indirect impact to the climate. An example of this, taken from the IPCC Good Practice Guidance paper [Hegerl et al. 2010], is the attribution of coral reef calcification impacts to rising CO₂ levels, in which an intermediate stage is used by first attributing changes in the carbonate ion concentration to rising CO₂ levels, then attributing calcification processes to changes in the carbonate ion concentration. This also illustrates the need for a clear understanding of the physical mechanisms involved, in order to perform a reliable multi-step attribution in the presence of many potential confounding factors.

In the interpretation of attribution results, in particular those framed as a question of whether human activity has influenced a particular climatic change or event, there is a tendency to focus on whether the confidence interval of the estimated anthropogenic effect crosses zero. The absence of such a crossing indicates that change is likely to be due to human factors. This results in conservative attribution statements, but it reflects the focus of the present debate where, in the eyes of the public and media, “attribution” is often understood as confidence in ruling out non-human factors, rather than as giving a best estimate or relative contributions of different factors.

Statistical analysis quantifies the strength of the relationship, given the simplifying assumptions of the attribution framework, but the level of confidence in the simplifying assumptions must be assessed outside that framework. This level of confidence is standardised by the IPCC into discrete (though subjective) categories (“very high”, “high”, etc.), which aim to take account of the process knowledge, data limitations, adequacy of models used, and the presence of potential confounding factors. The conclusion that is reached will then have a form similar to the IPCC’s headline attribution statement:

‘It is *extremely likely* [95% probability] that more than half of the observed increase in global average surface temperature from 1951 to 2010 was caused by the anthropogenic increase in greenhouse gas concentrations and other anthropogenic forcings together.’ [IPCC 2013; Summary for Policymakers, section D.3].

One method to reach such results is *optimal fingerprinting*. The method seeks to define a spatio-temporal pattern of change (fingerprint) associated with each potential driver (such as the effect of greenhouse gases or of changes in solar radiation), normalised relative to the internal variability, and then perform a statistical regression of observed data with respect to linear combinations of these patterns. The residual variability after observations have been attributed to each factor should then be consistent with the internal variability; if not, this suggests that an important source of variability remains unaccounted for. Parker [2010] argues that climate change fingerprint studies are similar to Steel's [2008] streamlined comparative process tracing. That is, in fingerprint studies computer simulation models are used to quantify and characterise the expected effects of mechanisms (such as from the greenhouse gas effect or the cooling of aerosols when they scatter radiation). These effects serve then as fingerprints that are used to test claims about the causes of climate change.

As emphasised by Parker [2010], fingerprint studies rely on several assumptions. The first one is linearity, i.e. that the response of the climate system when several forcing factors are present is equal to a linear combination of the forcings. Because the climate system is nonlinear, this is clearly a source of methodological difficulty, although for global-scale responses (in contrast to regional-scale responses) additivity has been shown to be a good approximation. Another assumption is that climate models simulate the causal processes that are at work in Earth's climate system accurately enough. As Parker argues, the very success of fingerprint studies (which is nontrivial) can help putting aside worries about this assumption. A further problem is, once more, the need to define internal variability characteristics (see also discussions in IPCC [2013, §10.2.3]). As argued above, the practice of estimating internal variability properties (in the form of a covariance matrix) from long unforced climate model runs introduces an element of circularity.

Levels of confidence in these attribution statements are primarily dependent on physical understanding of the processes involved. Where there is a clear, simple, well-understood mechanism, there should be greater confidence in the statistical result; where the mechanisms are loose, multi-factored or multi-step, or where a complex model is used as an intermediary, confidence is correspondingly lower. The Guidance Paper cautions that,

‘Where models are used in attribution, a model’s ability to properly represent the relevant causal link should be assessed. This should include an assessment of model biases and the model’s ability to capture the relevant processes and scales of interest.’ [Hegerl 2010, 5]

As Parker [2010] argues, there is also higher confidence in attribution results when the results are robust and there is a variety of evidence. For instance, the finding that late twentieth century temperature increase was mainly caused by greenhouse gas forcing is found to be robust given a wide range of different models, different analysis techniques and different forcings and there is a variety of evidence all supporting this claim. Thus our confidence that greenhouse gases explain global warming is high. (For further useful extended discussion of detection and attribution methods in climate science, see pages 872-878 of IPCC [2013], and in the Good Practice Guidance paper by Hegerl et al. [2010].)

There is an interesting question concerning the status of attribution methodologies like fingerprinting. The greenhouse effect is well-documented and indeed easily observable in laboratory experiments. This, some argue, provides a good qualitative understanding of the climate system, which is enough to say with confidence that global warming is real and that anthropogenic CO₂ has been identified as a cause of the increase in global mean temperature [e.g. Betz 2013]. This would render statistical debates about attribution methodologies second-order in terms of the key finding that anthropogenic CO₂ emissions cause global warming. However, this line of argument is not universally accepted. Winsberg and Goodwin [2015] argue that that fingerprinting is crucial in attributing an increase in global mean temperature to anthropogenic CO₂ emissions.

5. Confirmation and Predictive Power

Two questions arise in connection with models: how are models confirmed and what is their predictive power? *Confirmation* concerns the question whether, and to what degree, a certain model is supported by the data. Lloyd [2009] argues that many climate models are confirmed by past data. Parker [2009] objects to this claim. She argues that the idea that climate models *per se* are confirmed cannot be seriously entertained because all climate models are known to be wrong and empirically inadequate. Parker urges a shift in thinking from confirmation to *adequacy for purpose*: only hypotheses of climate models for particular purposes can be confirmed. For example, one might claim that a certain climate model adequately predicts the global temperature increase by 2100 (when run from certain initial conditions and relative to a certain emission scenario). Yet, at the same time, one might hold that the predictions of global mean precipitation by 2100 by the same model cannot be trusted.

Katzav [2014] cautions that adequacy for purpose assessments are of limited use. He claims that these assessments are typically unachievable because it is far from clear which of the model's observable implications can possibly be used to show that the model is adequate for the purpose. Instead he argues that climate models can at best be confirmed as providing a range of possible futures. Katzav is right to stress that adequacy for purpose assessments are more difficult than appears at first sight. But the methodology of adequacy for purpose cannot be dismissed wholesale; in fact it is used successfully across the sciences (e.g. when ideal gas models are confirmed to be useful for certain purposes). Whether or not adequacy for purpose assessment is possible depends on the case at hand.

If one finds that one model predicts certain variables well and another model doesn't, then one would like to know the reasons why the first model is successful and the second not. Lenhard and Winsberg [2010] argue that this is hopeless: For complex climate models a *strong version of confirmation holism* makes it impossible to tell where the failures and successes of climate models lie. In particular, they claim that it is impossible to assess the merits and problems of sub-models and the parts of models. They support their argument with case studies (e.g., about the coupled model intercomparison project) and argue that this strong form of confirmation holism is here to stay. There is a question, though, whether this confirmation holism affects all models and whether it is here to stay. Complex models have different modules for the atmosphere, the ocean and ice. These modules can be run individually and also together. The aim of the many new Model Intercomparison Projects (MIPs) is, by comparing individual and combined runs, to obtain an understanding of the performance and physical merits of separate modules, which it is hoped will identify areas for improvement and eventually result in better performance of the entire model.

Another problem concerns the use of data in the construction of models. The values of model parameters are often estimated using observations, a process known as *calibration*. For example, the magnitude of the aerosol forcing is often estimated from data. When data have been used for calibration, the question arises *whether the same data can be used again to confirm the model*. If data are used for confirmation that have not already been used for calibration, they are *use-novel*. If data are used for both calibration and confirmation, this is referred to as *double-counting*.

Scientists and philosophers alike have argued that double-counting is illegitimate and that data have to be use-novel to be confirmatory [Lloyd 2010; Shackley et al. 1998; Worrall 2010]. Steele and Werndl [2013] oppose this conclusion and argue that on *Bayesian* and relative-likelihood accounts of confirmation double-counting is legitimate. Furthermore, Steele and Werndl [2015] argue that *model selection theory* presents a more nuanced picture of the use of data than the commonly endorsed positions. Frisch [forthcoming] criticises this claim on the basis that evidence used for

confirmation cannot be ‘old’, as in climate science. But note that we need not interpret Bayesian probabilities in such a purist manner, as representing the actual beliefs of scientists at some particular point in time. The important point, as stressed by Steele and Werndl [2013], is that the same data cannot inform a prior probability for a hypothesis and also further (dis)confirm the hypothesis. There are two cases. First, there are methods such as cross-validation where the data are required to be use-novel. For cross-validation the data are split up into two groups: the first group is used for calibration and the second for confirmation. Second, there are the methods such as the Akaike Information Criterion for which the data need not be use-novel, although information criteria methods are hard to apply in practice to climate models because the number of degrees of freedom is poorly defined.

This brings us to the second issue: *prediction*. In the climate context this is typically framed as the issue of *projection*. ‘Projection’ is a technical term in the climate modelling literature and refers to a prediction that is conditional on a certain forcing scenario. The forcing scenario is specified either by the amount of greenhouse gas emissions and aerosols added to the atmosphere or directly by their atmospheric concentrations, and these in turn depend on future socioeconomic and technological developments.

Models can be put to different uses. Simulation results can be used either to develop physical insight about the system by means of experiment and comparison with reality, or to make projections of future climate (see Held [2005] for a discussion of this contrast). However, much research these days is undertaken with the aim of generating projections about the actual future evolution of the Earth system upon which policies are made and real-life decisions are taken. In these cases it is necessary to quantify and understand how good those projections are likely to be.

It is doubtful that this question can be answered along traditional lines. One such line would be to refer to the confirmation of a model against historical data (Chapter 9 of IPCC [2013] discusses model evaluation in detail) and argue that the ability of a model to successfully reproduce historical data should give us confidence that it will perform well in the future too. It is unclear at best whether this is a viable answer. The problem is that climate projections for high forcing scenarios take the system well outside any previously experienced state, and at least *prima facie* there is no reason to assume that success in low forcing contexts is a guide to success in high-forcing contexts; for example, a model calibrated on data from a world with the Arctic Sea covered in ice might no longer perform well when the sea ice is completely melted and the relevant dynamical processes are quite different. For this reason calibration to past data has at most limited relevance for a model’s predictive success [Oreskes et al. 1994; Stainforth et al. 2007a, 2007b, Steele and Werndl 2013].

This brings into focus the fact that there is no general answer to the question of the trustworthiness of model outputs. There is widespread consensus that predictions are better for longer time averages, larger spatial averages, low specificity and better physical understanding; and, all other things being equal, shorter lead times (nearer prediction horizons) are easier to predict than longer ones. Global mean temperature trends are considered trustworthy, and it is generally accepted that the observed upward trend will continue [Oreskes 2007], although the basis of this confidence is usually a physical understanding of the greenhouse effect with which the models are consistent, rather than a direct reliance on the output of models themselves. The latest IPCC report [IPCC 2013, Summary for Policymakers, section D.1] professes that modelled surface temperature patterns and trends are trustworthy on the global and continental scale.

There still are interesting questions about the epistemic grounds on which such assertions are made (and we return to them in the next section). A harder problem, however, concerns the use of models as providers of detailed information about the future *local* climate. The *United Kingdom Climate Impacts Program’s* UKCP09 project, for instance [Sexton et al. 2012, Sexton and Murphy 2012],

aims to make high-resolution probabilistic projections of the climate up to 2100 based on HadCM3, a global climate model developed at the UK Met Office Hadley Centre. Probabilities are given for events on a 25km grid for finely defined specific events such as changes in the temperature of the warmest day in summer, the precipitation of the wettest day in winter, or the change in summer-mean cloud amount, with projections blocked into overlapping thirty year segments which extend to 2100. It is projected, for instance, that under a medium emission scenario the probability for a 20-30% reduction in summer mean precipitation in central London in 2080 is 0.5.

The trustworthiness – and policy-relevance – of such projections has been disputed. A model has structural model error if the model’s dynamics differs from the dynamics in the target system. Frigg et al. [2014a] point out that any structural model error in nonlinear models may compromise the ability to generate decision-relevant predictions. Furthermore, there is little reason to expect that post-processing of model outputs can correct for the consequences of such errors [Frigg et al. 2014b, 2015]. This casts doubt on our ability, today, to make trustworthy, high-resolution probabilistic projections out to the end of this century. The research question is to determine the timescales on which such projections are likely to be reliable, and beyond those timescales to estimate the effect of model inadequacy. Where is the boundary between trustworthy and non-trustworthy projections? That is, where in between global temperature trends and precise projections on a 25km grid does trustworthiness come to an end? This is a crucial – and eminently policy-relevant – question in the epistemology of climate science, and one that is hitherto unsolved. This conclusion is challenged by Winsberg and Goodwin [2015] who submit that Frigg et al. overstate the limitations imposed by model error.

6. Understanding and Quantifying Uncertainty

Uncertainty features prominently in discussions about climate models, and yet is a concept that is poorly understood and that raises many difficult questions. In most general terms uncertainty is lack of knowledge. The first challenge is to circumscribe more precisely what is meant by ‘uncertainty’ and what the sources of uncertainty are. A number of proposals have been made, but the discussion is still in a ‘pre-paradigmatic’ phase. Smith and Stern [2011, 4821-4824] identify four relevant varieties of uncertainty: imprecision, ambiguity, intractability and indeterminacy. Spiegelhalter and Riesch [2011] consider a five-level structure with three within-model levels – event, parameter and model uncertainty – and two extra-model levels concerning acknowledged and unknown inadequacies in the modelling process. Wilby and Dessai [2010] discuss the issue with reference to what they call the cascade of uncertainty, studying how uncertainties magnify as one goes from assumptions about future global emissions of greenhouse gases to the implications of these for local adaptation. Petersen [2012, Chapters 3 and 6] introduces a so-called uncertainty matrix listing the sources of uncertainty in the vertical and the sorts of uncertainty in the horizontal direction. Lahsen [2005] looks at the issue from a science studies point of view and discusses the distribution of uncertainty as a function of the distance from the site of knowledge production. And these are but a few of the many proposals currently available.

The next problem is the one of measuring and quantifying uncertainty in climate predictions. Among the approaches that have been devised in response to this challenge, ensemble methods occupy centre stage. Current estimates of climate sensitivity and increase in global mean temperature under various emission scenarios, for instance, include information derived from ensembles containing multiple climate models. The reason to use ensembles is the acknowledged uncertainties in individual models, which concerns both the model structure and the values of parameters in the model. It is a common assumption that ensembles help mitigate the effects of these uncertainties either by producing and identifying “robust” predictions, or by providing estimates of this uncertainty about future climate change. (Parker [2013] provides an excellent discussion of ensemble methods and the problems that attach to them.)

Before discussing the epistemic function of ensembles, a distinction needs to be made between two types of ensembles. As we have seen above, a climate model has a number of parameters in it. Some represent physical magnitudes such as the viscosity of water, while others are ‘effective summaries’ of sub-grid processes that are not explicitly resolved (such as cloud coverage). A *perturbed parameter ensemble* (PPE, sometimes alternatively a “perturbed physics ensemble”) explores how sensitively the outputs of one model depend on the parameters by running the model a number of times, each time with different parameter values. In this way the ensemble explores the impact of parametric uncertainty on predictions (i.e. a sensitivity analysis with respect to the chosen parameters). By contrast, a *multi model ensemble* (MME) consists of several different models – i.e. models that differ in mathematical structure and physical content rather than only in parameter values. Such an ensemble is used to investigate how predictions of relevant climate variables are impacted by uncertainty about the model structure.

A result is *robust* if all or most models in the ensemble show the same result; for general discussion of robustness analysis see Weisberg [2006]. If, for instance, all models in an ensemble show more than 4° increase in global mean temperature by the end of the century when run under a certain emission scenario, this result is robust across the specified ensemble. Does robustness justify increased confidence? Lloyd [2010, 2015] argues that robustness arguments are powerful in connection with climate models and lend credibility at least to core claims like the reality of global warming in the of the 20th Century. Parker [2011], by contrast, reaches a more sober conclusion: ‘When today’s climate models agree that an interesting hypothesis about future climate change is true, it cannot be inferred [...] that the hypothesis is likely to be true or that scientists’ confidence in the hypothesis should be significantly increased or that a claim to have evidence for the hypothesis is now more secure’ [*ibid.* 579]. One of the main problems is that if today’s models share the same technological constraints posed by today’s computer architecture and understanding of the climate system, then they inevitably share some common errors. Indeed such common errors have been widely acknowledged (see, for instance, Knutti et al. [2010]) and studies have demonstrated and discussed the lack of model independence [Bishop and Abramowitz 2013; Jun et al. 2008a; 2008b]. But if models are not independent, then agreement between them carries little, if any, epistemic weight.

When ensembles do not yield robust predictions, then the spread of results within the ensemble is often used to estimate quantitatively the uncertainty of the outcome. There are two main approaches to this. The first approach aims to translate the histogram of model results directly into a probability distribution. These methods assign higher probabilities to outcomes on which more ensemble members agree: in effect the guiding principle is that the probability of an outcome is proportional to the fraction of models in the ensemble which produce that result. The thinking behind this method seems to be to first treat models as exchangeable sources of information (in the sense that there is no reason to trust one ensemble member any more than any other), and then invoke some sort of frequentist approach to probabilities. As we have seen above, the assumption that models are independent is problematic for a number of reasons. But there is a further problem: current MMEs are ‘ensembles of opportunity’, grouping together existing models. Even the models of the best currently available ensembles such as CMIP5 are not designed to systematically explore all possibilities, and it is therefore imaginable that there are vast classes of models that produce entirely different results. Unless a reason can be found to rule this out, it is not clear why the frequency of ensemble results should double as a guide to probability. The IPCC acknowledges this limitation (see discussion in Chapter 12 of IPCC [2013]) and thus downgrade the assessed likelihood of ensemble-derived confidence intervals.

A more modest approach regards ensemble outputs as a guide to possibility rather than probability. On this view the spread of an ensemble presents the range of outcomes that cannot be ruled out. The

bounds of this set of results – often referred to as a ‘non-discountable envelope’ – provide a lower bound of the uncertainty [Stainforth et al. 2007b]. In this spirit Katzav [2014] argues that a focus on prediction is misguided and that models ought to be used to show that certain scenarios are real possibilities.

While undoubtedly less committal than the probability approach, also non-discountable envelopes raise questions. The first is the relation between non-discountability and possibility. Non-discountable results are ones that cannot be ruled out. How is this judgment reached? Do results which, given current knowledge, cannot be ruled out indicate possibilities? If not, what is their relevance for estimating lower bounds? Furthermore, it is important to keep in mind that the envelope just represents some possibilities. Hence it does not indicate the *complete* range of possibilities, making certain types of formalised decision-making procedures impossible. For a further discussion of these issues see Betz [2009, 2010].

Finally, a number of authors emphasise the limitations of model-based methods (such as ensemble methods) and submit that any realistic assessment of uncertainties will also have to rely on other factors, most notably expert judgement. Petersen [2012, Chapter 4] outlines the approach of the Netherlands Environmental Assessment Agency (PBL), which sees expert judgment and problem framings as essential components of uncertainty assessment. Aspinall [2010] suggests using methods of structured expert elicitation.

In light of the issues raised above, how should uncertainty in climate science be communicated to decision-makers? The most prominent framework for communicating uncertainty is the IPCC’s. The latest version of the framework, which is used throughout the Fifth Assessment Report (AR5), is explicated in the ‘Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties’ and further explicated in [Mastrandrea et al. 2011]. The framework appeals to two measures for communicating uncertainty. The first, a qualitative ‘confidence’ scale, depends on both the type of evidence and the degree of agreement amongst experts. The second measure is a quantitative scale for representing statistical likelihoods (or more accurately, fuzzy likelihood intervals) for relevant climate/economic variables. The following statement exemplifies the use of these two measures for communicating uncertainty in AR5: ‘The global mean surface temperature change for the period 2016–2035 relative to 1986–2005 is similar for the four RCPs and will *likely* be in the range 0.3°C to 0.7°C (*medium confidence*). [IPCC 2013] A discussion of this framework can be found in Budescu et al. [2014] and Adler and Hadorn [2014].

At this point it should also be noted that the role of ethical and social values in relation to uncertainties in climate science is controversially debated. Winsberg [2012] appeals to complex simulation modelling to argue that it is infeasible for climate scientists to produce results that are untinted by their ethical and social values. More specifically, he argues that assignments of probabilities to hypotheses about future climate change are influenced by ethical and social values because of the way these values come into play in the building and evaluating of climate models. Parker objects that often pragmatic factors rather than social or ethical values play a role when evaluating and building climate models. She further objects that the focus on precise probabilistic uncertainty estimates is misguided. Uncertainty about future climate change is more appropriately depicted with coarser estimates and these coarser estimates are less influenced by values. She argues that these objections show that while Winsberg’s argument is not mistaken, it exaggerates the influence of ethical and social values. Parker then goes on to argue that a more traditional challenge to the value-free ideal of science fits to the climate case. Namely, one could argue that estimates of uncertainty are themselves always somewhat uncertain, and that the decision to offer a particular estimate of uncertainty thus involves value judgements [cf. Douglas 2009].

7. Conceptualising Decisions Under Uncertainty

The decisions that we make in response to climate change have consequences affecting both individuals and groups at different places and times. Moreover, the circumstances of many of these decisions involve uncertainty and disagreement that is sometimes both severe and wide-ranging, concerning not only the state of the climate (as discussed above) and the broader social consequences of any action or inaction on our part, but also the range of actions available to us and what significance we should attach to their possible consequences. These considerations make climate decision-making both important and hard. The stakes are high, and so too are the difficulties for standard decision theory—plenty of reason for philosophical engagement with this particular application of decision theory.

Let us begin by looking at the actors in the climate domain and the kinds of decision problems that concern them. When introducing decision theory, it is common to distinguish three main domains: individual decision theory (which concerns the decision problem of a single agent who may be uncertain of her environment), game theory (which focuses on cases of strategic interaction amongst rational agents), and social choice theory (which concerns procedures by which a number of agents may ‘think’ and act collectively). All three realms are relevant to the climate-change predicament, whether the concern is adapting to climate change or mitigating climate change or both.

Determining the appropriate agential perspective and type of engagement between agents is important, because otherwise decision-modelling efforts may be in vain. For instance, it may be futile to focus on the plight of individual citizens when the power to affect change really lies with states. It may likewise be misguided to analyse the prospects for a collective action on climate policy, if the supposed members of the group do not see themselves as contributing to a shared decision that is good for the group as a whole. It would also be misleading to exclude from an individual agent’s decision model the impact of others who perceive that they are acting in a strategic environment. This is not, however, to recommend a narrow view of the role of decision models—that they must always realistically represent the key decisions at hand; the point is rather that we should not employ decision models with particular agential framings in a naïve way.

Getting the agential perspective right is just the first step in framing a decision problem so that it presents convincing reasons for action. There remains the task of representing the details of the decision problem from the appropriate epistemic and evaluative perspective. Our focus is individual decision theory, for reasons of space, and because most decision settings ultimately involve the decision of an individual, whether this be a single person or a group.

The standard model of (individual) decision-making under uncertainty used by decision theorists derives from the classic work of von Neumann and Morgenstern [1944] and Leonard Savage [1954]. It treats actions as functions from possible states of the world to consequences, these being the complete outcomes of performing the action in question in that state of the world. All uncertainty is taken to be uncertainty about the state of the world and is quantified by a single probability function over the possible states, where the probabilities in question measure either objective risk or the decision maker’s degrees of belief (or a combination of the two). The relative value of consequences is represented by an interval-scaled utility function over these consequences. Decision-makers are advised to choose the action with maximum *expected utility* (EU); where the EU for an action is the sum of the probability-weighted utility of the possible consequences of the action.

It is our contention that this model is inadequate for many climate-oriented decisions, because it fails to properly represent the multidimensional nature and severity of the uncertainty that decision-makers face. To begin with, not all the uncertainty that climate decision-makers face is empirical

uncertainty about the actual state of the world (state uncertainty). There may be further empirical uncertainty about what options are available to them and what are the consequences of exercising each option for each respective state (option uncertainty). In what follows we use term ‘empirical uncertainty’ to cover both state uncertainty and option uncertainty. Furthermore, decision-makers face a non-empirical kind of uncertainty – ethical uncertainty – about what values to assign to possible consequences.

Let us now turn to empirical uncertainty. As noted above, standard decision theory holds that all empirical uncertainty can be represented by a probability function over the possible states of the world. There are two issues here. The first is that confining all empirical uncertainty to the state space is rather unnatural for complex decision problems such as those associated with climate change. In fact, decision models are less convoluted if we allow the uncertainty about states to depend on the actions that might be taken (cf. Richard Jeffrey’s [1965] expected utility theory), and if we also permit further uncertainty about what consequence will arise under each state, given the action taken (an aspect of option uncertainty). For instance, consider a crude version of the mitigation decision problem faced by the global planner: it may be useful to depict the decision problem with a state-space partition in terms of possible increases in average global temperature over a given time period. In this case our beliefs about the states (how likely they each are) would be conditional on the mitigation option taken. Moreover, for each respective mitigation option, the consequence arising in each of the states depends on further uncertain features of the world, for instance the extent to which, on average, regional conditions would be favourable to food production and whether social institutions would facilitate resilience in food production.

The second issue is that using a precise probability function to represent uncertainty about states (and consequences) can misrepresent the severity of this uncertainty. For instance, even if one assumes that the position of the scientific community may be reasonably well represented by a precise probability distribution over the state space, conditional on the mitigation option, precise probabilities over the possible food productions and other economic consequences, given this option and average global temperature rise, are less plausible. Note that the global social planner’s mitigation decision problem is typically analysed in terms of a so-called Integrated Assessment Model (IAM), which does indeed involve dependencies between mitigation strategies and both climate and economic variables. As regards empirical uncertainty, Nordhaus’s [2008] reliance on ‘best estimates’ for parameters like climate sensitivity can be compared with Stern’s [2007] use of ‘confidence intervals’. Frisch [2013] nonetheless argues that *all extant* IAMs inadequately represent the uncertainty surrounding model parameters [cf. Weitzman 2009], and recommends the use of sets of plausible probabilities and utilities, in line with our comments in this section. Popular among philosophers is the use of sets of probability functions to represent severe uncertainty, whether the uncertainty is due to evidential limitations or due to evidential/expert disagreement. This is a minimal generalisation of the standard decision model, in the sense that probability measures still feature: roughly, the more severe the uncertainty, the more probability measures over the space of possibilities needed to conjointly represent the epistemic situation (see, for instance, Walley [1991]). For maximal uncertainty all possibilities are on a par—they are effectively assigned probability $[0, 1]$. Indeed it is a strength of the imprecise probability representation that it generalises the two extreme cases, i.e. the *precise probabilistic* as well as the *possibilistic* frameworks. In some contexts, it may be suitable to weight the possible probability distributions in terms of plausibility (as required for some of the decision rules discussed below). (See Halpern [2003] for a thorough treatment of frameworks, both qualitative and quantitative, for representing uncertainty.) This latter approach may in fact better match the IPCC’s representation of the uncertainty surrounding decision-relevant climate and economic variables. Indeed, an important question is whether and how the IPCC’s representation of uncertainty can be translated into the kind of *imprecise probabilistic* framework discussed here and in the next section. An alternative to the aforementioned proposal is that the IPCC’s confidence and likelihood measures for relevant

variables should be combined to form an unweighted imprecise set of probability distributions, or even a precise probability distribution, suitable for input into an appropriate decision model.

Decision makers face uncertainty not only about what will or could happen, but also about what value to attach to these possibilities – in other words, they face ethical uncertainty. Such value or ethical uncertainty can have a number of different sources. The most important ones arise in connection with judgments about how to distribute the costs and benefits of mitigation and adaptation amongst different regions and countries, about how to take account of persons whose existence depends on what actions are chosen now, and about the degree to which future wellbeing should be discounted. (For discussion and debate about the ethical significance of various climate outcomes, particularly at the level of global rather than regional or national justice, see the articles in Gardiner et al.'s [2010] edited collection, *Climate Ethics*.) Of these, the latter has been the subject of the most debate, because of the extent to which (the global planner's) decisions about how drastically to cut carbon emissions are sensitive to the discount rate used in evaluating the possible outcomes of doing so (as highlighted in Broome [2008]). Discounting thus provides a good illustration of the importance of ethical uncertainty.

In many economic models, a discount rate is applied to a measure of total wellbeing at different points in time (the 'pure rate of time preference'), with a positive rate implying that future wellbeing carries less weight in the evaluations of options than present wellbeing. Note that the overall 'social discount rate' in economic models involves other terms besides the pure rate of time preference; these other discounting terms apply to goods or consumption rather than wellbeing per se. See Broome [1992] and Parfit [1984] for helpful discussions of the reasons for discounting *goods* that do not imply discounting *wellbeing*. Many philosophers regard any pure discounting of future wellbeing as completely unjustified from an objective point of view. This is not to deny that temporal location may nonetheless correlate with features of the distribution of wellbeing that are in fact ethically significant. If people will be better off in the future, for instance, it is reasonable to be less concerned about their interests than those of the present generation, much as one might prioritise the less well-off within a single generation. But the mere fact of a benefit occurring at a particular time cannot be relevant to its value, at least from an impartial perspective.

Economists do nonetheless often discount wellbeing in their policy-oriented models, although they disagree considerably about what pure rate of time preference should be used. One view, exemplified by the Stern Review and representing the impartial perspective described above, is that only a very small rate (in the order of 0.5%) is justified, and this on the grounds of the small probability of the extinction of the human population. Other economists, however, regard a partial rather than an impartial point of view more appropriate in their models. A view along these lines, exemplified by Nordhaus [2007] and Arrow [1995a], is that the pure rate of time preference should be determined by the preferences of current people. But typical derivations of average pure time discounting from observed market behaviour are much higher than those used by Stern (around 3% by Nordhaus's estimate). Although the use of this data has been criticised for providing an inadequate measure of people's reasoned preferences (see, e.g., Sen [1982], Drèze and Stern [1990], Broome [1992]), the point remains that any plausible method for determining the current generation's attitude to the wellbeing of future generations is likely to yield a rate higher than that advocated by the Stern Review. To the extent that this debate about the ethical basis for discounting remains unresolved, there will be ethical uncertainty about the discount rate in climate policy decisions. This ethical uncertainty may be represented analogously to empirical uncertainty—by replacing the standard precise utility function with a set of possible utility functions.

8. Managing Uncertainty

How should a decision maker choose amongst the courses of action available to her when she must make the choice under conditions of severe uncertainty? The problem that climate decision makers face is that, in these situations, the precise utility and probability values required by standard EU theory may not be readily available.

There are, broadly speaking, three possible responses to this problem.

Firstly, the decision maker can simply bite the bullet and try to settle on precise probability and utility judgements for the relevant contingencies. Orthodox decision theorists argue that rationality requires that decisions be made as if they maximise the decision maker's subjective expectation of benefit relative to her precise degrees of belief and values. Broome [2012, 129] gives an unflinching defence of this approach: "The lack of firm probabilities is not a reason to give up expected value theory [...] Stick with expected value theory, since it is very well founded, and do your best with probabilities and values." (When it comes to environmental decision making, in particular, it may already be a big step in terms of representing uncertainty to introduce well-founded precise probability and utility functions (cf. Steele [2006] on the Precautionary Principle). Weitzman [2009], for instance, argues that proper acknowledgment of the non-negligible probability of catastrophic climate consequences (in terms of a 'fat tailed' probability distribution) radically changes the assessment of mitigation options.) But in many circumstances there remains the question of how to follow Broome's advice: How should the decision maker settle, in a non-arbitrary way, on a precise opinion on decision-relevant issues in the face of an effectively 'divided mind'? There are two interrelated strategies: she can deliberate further and/or aggregate conflicting views. The former aims for convergence in opinion, while the latter aims for an acceptable compromise in the face of persisting conflict. (For a discussion of deliberation see Fishkin and Luskin [2005]; for more on aggregation see, for instance, Genest and Zidek [1986], Mongin [1995], Sen [1970], List and Puppe [2009]. There is a comparatively small formal literature on deliberation, a seminal contribution being Lehrer and Wagner's [1981] model for updating probabilistic beliefs.)

Secondly, the decision maker can try to delay making a decision, or at least postpone parts of it, in the hope that her uncertainty will become manageable as more information becomes available, or as disagreements resolve themselves through a change in attitudes. The basic motive for delaying a decision is to maintain *flexibility* at zero cost (see Koopmans [1962], Kreps and Porteus [1978], Arrow [1995b]). Suppose that we must decide between building a cheap but low sea wall or a high, but expensive, one, and that the relative desirability of these two courses of action depends on unknown factors, such as the extent to which sea levels will rise. In this case it would be sensible to consider building a low wall first but leave open the possibility of raising it in the future. If this can be done at no additional cost, then it is clearly the best option. In many adaptation scenarios, the analogue of the 'low sea wall' may in fact be social-institutional measures that enable a delayed response to climate change, whatever the details of this change turn out to be. In many cases, however, the prospect of cost-free postponement of a decision (or part thereof) is simply a mirage, since delay often decreases rather than increases opportunities due to changes in the background environment. This is often true for climate-change adaptation decisions, not to mention mitigation decisions.

Finally the decision maker can employ a different decision rule to that prescribed by EU theory; one that is much less demanding in terms of the information it requires. A great many different proposals for such rules exist in the literature, involving more or less radical departures from the orthodox theory and varying in the informational demands they make. It should be noted from the outset that there is one widely-agreed rationality constraint on these non-standard decision rules: '(EU)-dominated options' are not admissible choices, i.e., if an option has lower expected utility than another option according to all permissible pairs of probability and utility functions, then the former dominated option is not an admissible choice. This is a relatively minimal constraint, but it may well yield a unique choice of action in some decision scenarios. In such cases, the severe

uncertainty is not in fact decision relevant. For example, it may be the case that, from the global planner's perspective, a given mitigation option is better than continuing with business as usual, whatever the uncertain details of the climate system. This is even more plausible to the extent that the mitigation option counts as a 'win-win' strategy [Maslin and Austin 2012], i.e., to the extent that it has other positive impacts, say, on air quality or energy security, regardless of mitigation results. In many more fine-grained or otherwise difficult decision contexts, however, the non-EU-dominance constraint may exclude only a few of the available options as choice-worthy.

A consideration that is often appealed to in order to further discriminate between options is *caution*. Indeed, this is an important facet of the popular but ill-defined Precautionary Principle. (The Precautionary Principle is referred to in the latest IPCC [2014b] ARC-5 WGII report. See, for instance, Gardiner [2006] and Steele [2006] for discussion of what the Precautionary Principle does/could stand for.) Cautious decision rules give more weight to the 'down-side' risks; the possible negative implications of a choice of action. The Maxmin-EU rule, for instance, recommends picking the action with greatest minimum expected utility (see Gilboa and Schmeidler [1989], Walley [1991]). The rule is simple to use, but arguably much too cautious, paying no attention at all to the full spread of possible expected utilities. The α -Maxmin rule, in contrast, recommends taking the action with the greatest α -weighted sum of the minimum and maximum expected utilities associated with it. The relative weights for the minimum and maximum expected utilities can be thought of as reflecting either the decision maker's pessimism in the face of uncertainty or else their degree of caution (see Binmore [2009]). (For a comprehensive survey of non-standard decision theories for handling severe uncertainty in the economics literature, see Gilboa and Marinacci [2012].)

A more informationally-demanding set of rules are those that draw on considerations of *confidence* and/or *reliability*. The thought here is that an agent is more or less confident about the various probability and utility functions that characterise her uncertainty. For instance, when the estimates derive from different models or experts, the decision maker may regard some models as better corroborated by available evidence than others or else some experts as more reliable than others in their judgments. In these cases it is reasonable, *ceteris paribus*, to favour actions of which you are more confident that they will have beneficial consequences. One (rather sophisticated) way of doing this is to weight each of the expected utilities associated with an action in accordance with how confident you are about the judgements supporting them and then choose the action with the maximum confidence-weighted expected utility (see Klibanoff et al. [2005]). This rule is not very different from maximising expected utility and indeed one could regard confidence weighting as an aggregation technique rather than an alternative decision rule. But considerations of confidence may be appealed to even when precise confidence weights cannot be provided. Gärdenfors and Sahlin [1982/ 1988], for instance, suggest simply excluding from consideration any estimates that fall below a reliability threshold and then picking cautiously from the remainder. Similarly, Hill [2013] uses an ordinal measure of confidence that allows for stake-sensitive thresholds of reliability that can then be combined with varying levels of caution.

One might finally distinguish decision rules that are cautious in a slightly different way—that compare options in terms of 'robustness' to uncertainty, relative to a problem-specific satisfactory level of expected utility. Better options are those that are more assured of having an expected utility that is good enough or regret-free, in the face of uncertainty. The 'information-gap theory' developed by Ben-Haim [2001] provides one formalisation of this basic idea that has proved popular in environmental management theory. Another prominent approach to robust decision-making is that developed by Lempert, Popper and Bankes [2003]. These two frameworks are compared in Hall et al. [2012]. Recall that the uncertainty in question may be multi-faceted, concerning probabilities of states/outcomes, or values of final outcomes. Most decision rules that appeal to robustness assume that a best estimate for the relevant variables is available (perhaps

achieved by one of the methods described in 3.3 A), and then consider deviations away from this estimate. A robust option is one that has a satisfactory expected utility relative to a class of estimates that deviate from the best one to some degree; the wider the class in question, the more robust the option. Much depends on what expected utility level is deemed satisfactory. For mitigation decision making, one salient satisfactory level of expected utility is that associated with a 50% chance of average global temperature rise of 2 degrees Celsius or less. Note that one may otherwise interpret any such mitigation temperature target in a different way, namely as a constraint on what counts as a *feasible* option. In other words, mitigation options that do not meet the target are simply prohibited options, not suitable for consideration. For adaptation decisions, the satisfactory level would depend on local context, but roughly speaking, robust options are those that yield reasonable outcomes for all the inopportune climate scenarios that have non-negligible probability given some range of uncertainty. These are plausibly adaptation options that focus on resilience to any and all of the aforesaid climate scenarios, perhaps via the development of social institutions that can coordinate responses to variability and change. (Robust decision-making is endorsed, for instance, by Dessai et al. [2009] and Wilby and Dessai [2010], who indeed associate this kind of decision rule with resilience strategies. See also Linkov et al. [2014] for discussion of resilience strategies vis-à-vis risk management.)

9. Conclusion

In this article we reviewed, from a philosophy of science perspective, issues and questions that arise in connection with climate science. Most of these issues are the subject matter of ongoing research, and they indeed deserve further attention. Rather than repeating these points, we would like to mention a topic that has not received the attention that it deserves: the epistemic significance of consensus in the acceptance of results. Not much attention has been paid to this issue in the context of climate science. Yet, as the recent controversy over the Cook et al. [2013] paper shows, many people do seem to think that the level of expert consensus is an important reason to believe in climate change given that they themselves are not expert; and conversely, attacking the consensus and sowing doubt is a classic tactic of the other side. The role of consensus in the context of climate change deserves more attention than it has received hitherto.

Acknowledgments

We would like to thank the editors and two referees for this volume for their constructive feedback on earlier drafts. Many thanks also to Reto Knutti, Wendy Parker, Lenny Smith, and Dave Stainforth valuable discussions and suggestions. Funding support for the research was provided by the Arts and Humanities Research Council via Grant AH/J006033/1, and by the ESRC Centre for Climate Change Economics and Policy, funded by the Economic and Social Research Council. Frigg further acknowledges financial support from the Project and Grant FFI2012-37354 of the Spanish Ministry of Science and Innovation (MICINN). Finally we would like to thank Markus Hierl for invaluable editorial assistance with the preparation of the manuscript for publication.

Glossary

Attribution (of climate change): The process of evaluating the relative contributions of multiple causal factors to a change or event with an assignment of statistical confidence.

Boundary conditions: Values for any variable which affect the system but which are not directly output by the calculations.

Calibration: The process of estimating the values of model parameters by using observations.

Climate model: A representation of certain aspects of the climate system.

Detection (of climate change): The process of demonstrating that climate or a system affected by climate has changed in some defined statistical sense without providing a reason for that change.

Double counting: The use of data for both calibration and confirmation.

Expected utility (for an action): The sum of the probability-weighted utility of the possible consequences of the action.

External conditions (of the climate system): Conditions that influence the state of the earth such as the amount of energy received from the sun.

Initial conditions: A mathematical descriptions of the state of the climate system at the beginning of the period being simulated.

Internal variability: The phenomenon that climate variables such as temperature and precipitation would change over time due to the internal dynamics of the climate system even in the absence of climate change.

Null hypothesis: The expected behaviour of the climate system in the absence of changing external influences.

Projection: The prediction of a climate model that is conditional on a certain forcing scenario.

Proxy data: The data for climate variables that derived from observing natural phenomena such as tree rings, ice cores and ocean sediments.

Robustness (of a result): A result is robust if all (or most) models show the same result.

Use novel data: Data that are used for confirmation and have not been used for calibration.

Bibliography

- Adler C. E. and G. H. Hadorn. (2014). The IPCC and treatment of uncertainties: topics and sources of dissensus. *Wiley Interdisciplinary Reviews: Climate Change* **5.5**, 663 – 676. [Reviews debate about procedure, scope and purpose of IPCC approach to uncertainty communication].
- Arrow K. J. (1995b). A Note on Freedom and Flexibility. *Choice, Welfare and Development*. (eds. K. Basu, P. Pattanaik, and K. Suzumura), 7 – 15. Oxford: Oxford University Press. [Outlines the rationale for a preference for flexibility in future choice sets, based on uncertainty about one’s future utility function].
- Arrow K. J. (1995a). ‘Discounting Climate Change: Planning for an Uncertain Future. *Lecture given at Institut d’Économie Industrielle, Université des Sciences Sociales, Toulouse*. Available online at <http://idei.fr/doc/conf/annual/paper_1995.pdf> (accessed 22 August 2014). [Discusses proposals for what is the appropriate rate for discounting future costs/benefits in public investment decisions, and argues for a non-negligible pure-time discounting rate].
- Aspinall W. (2010). A route to more tractable expert advice. *Nature* **463**, 294 – 295. [Suggests that expert advice should be used to quantify uncertainty and improve decision-making].
- Ben-Haim Y. (2001). *Information-Gap Theory: Decisions Under Severe Uncertainty*, 330 pp. London: Academic Press. [Proposes a non-probabilistic decision theory for handling situations of severe uncertainty that focuses on robustness to uncertainty in reaching satisfactory decision outcomes].
- Betz G. (2013). Chaos, plurality and model metrics in climate science. *Models, Simulation, and the Reduction of Complexity*, (eds. Ulrich v. Gähde et al.), 255 – 263. Berlin: de Gruyter. [Discusses the role of chaos and model metrics in climate science].
- Betz G. (2009). What range of future scenarios should climate policy be based on? Modal falsificationism and its limitations. *Philosophia naturalis* **46**, 133 – 158. [Criticises the IPCC’s climate scenarios and proposes an alternative method].
- Betz G. (2010). What’s the worst case?. *Analyse und Kritik* **32**, 87 – 106. [Discusses worst case scenarios in Knightian framework].
- Binmore K. (2009). *Rational Decisions*, 216 pp. Princeton, NJ: Princeton University Press. [Exposition and critical discussion of Bayesian decision theory, with some new ideas about how to handle ambiguity].
- Bishop C. H. and G. Abramowitz. (2013). Climate model dependence and the replicate Earth paradigm. *Climate Dynamics* **41**, 885 – 900. [Considers the causes and effects of model dependence, and suggests an empirical ensemble correction method].
- Broome J. (2012). *Climate Matters: Ethics in a Warming World*, 192 pp. New York: Norton. [A discussion of the causes and consequences of climate change, and associated individual and government responsibilities].
- Broome J. (1992). *Counting the Cost of Global Warming*, 147 pp. Cambridge: The White Horse Press. [A philosophical analysis of welfare economics as it applies to long-range environmental problems such as global warming, with particular attention to our obligations to future generations].
- Broome J. (2008). The Ethics of Climate Change. *Scientific American* **298**, 96 – 102. [Discusses the framing of climate-change decisions, including key ethical inputs and the treatment of risk].
- Budescu, D. V., H. Por, S. B. Broome and M. Smithson. (2014). The interpretation of IPCC probabilistic statements around the world. *Nature Climate Change* **4**, 508 – 512. [Discusses how to interpret the probabilistic statements in the IPCC].
- Cohn T. A. and H. F. Lins. (2005). Nature’s style: naturally trendy. *Geophysical Research Letters* **32**, L23402. [Demonstrates the importance of null hypothesis for statistical significance testing of trends in time series, especially with respect to long-term persistence].
- Cook J. et al. (2013). Quantifying the consensus on the anthropogenic global warming in the scientific literature. *Environmental Research Letters* **8**, 1 – 7. [Discusses the various sources of the consensus about global warming]

- Daron J. D. and D. Stainforth. (2013). On predicting climate under climate change. *Environmental Research Letters* **8**, 1 – 8. [Discusses how the climate can be predicted given that there is climate change].
- Dessai S. et al. (2009). Do We Need Better Predictions to Adapt to a Changing Climate?. *Eos* **90.13**, 111 – 112. [Argues that society can and must make effective adaptation decisions even in the absence of accurate and precise climate predictions].
- Dessler A. (2011). *Introduction to Modern Climate Change*, 238 pp. Cambridge: Cambridge University Press. [Introductory textbook covering aspects of science, economics and policy of climate change].
- Drèze J., and Stern, N. (1990). Policy reform, shadow prices, and market prices. *Journal of Public Economics* **42.1**, 1 – 45. [Argues for the evaluation of policy reforms in terms of appropriate shadow prices, where market prices are a misleading indicator of social value].
- Douglas H. (2009). *Science, Policy, and the Value-Free Ideal*, 256 pp. Pittsburgh: Pittsburgh University Press. [This book provides a detailed and comprehensive discussion of values in science and policy making].
- Fishkin J. S., and R. C. Luskin. (2005). Experimenting with a Democratic Ideal: Deliberative Polling and Public Opinion. *Acta Politica* **40**, 284 – 298. [Discusses views on the process and effects of group deliberation, and presents empirical evidence against ‘defeatist’ views that deliberation has no positive impact].
- Frank D., J. Esper, E. Zorita and R. Wilson. (2010). A noodle, hockey stick, and spaghetti plate: A perspective on high-resolution paleoclimatology. *Wiley Interdisciplinary Reviews: Climate Change* **1.4**, 507 – 516. [Discusses the Hockey Stick Controversy and shows that there is now a robust consensus that temperatures during the late 20th century are likely to have been the warmest in the past 1400 years].
- Frigg R. P., D. A. Stainforth and L. A. Smith. (2015). An Assessment of the Foundational Assumptions in High-Resolution Climate Projections: The Case of UKCP09 2015, draft under review. [Offers sustained criticism of the UKCP09 project].
- Frigg R. P., S. Bradley, H. Du and L. A. Smith. (2014a). Laplace’s Demon and the Adventures of His Apprentices. *Philosophy of Science* **81.1**, 31 – 59. [Argues that if a nonlinear model has only the slightest structural model error, then its ability to generate decision-relevant predictions is compromised].
- Frigg R. P., D. A. Stainforth and L. A. Smith. (2014b). The Myopia of Imperfect Climate Models: The Case of UKCP09. *Philosophy of Science* **80.5**, 886 – 897. [Criticises the methodology used to produce the UKCP09 climate projections].
- Frisch M. (2013). Modeling Climate Policies: A Critical Look at Integrated Assessment Models. *Philosophy and Technology* **26**, 117 – 137. [Offers a critical perspective on integrated assessment models].
- Frisch M. (forthcoming). Predictivism and old evidence: a critical look at climate model tuning. Forthcoming in: *European Journal for Philosophy of Science*. [Argues for a version of weak predictivism and is a response to Steele and Werndl (2013)].
- Gärdenfors P. and N.-E. Sahlin. [1982] (1988). Unreliable probabilities, risk taking, and decision making. *Decision, Probability and Utility*, (eds. P. Gärdenfors and N.-E. Sahlin), 313 – 334. Cambridge: Cambridge University Press. [Early discussion of how the reliability of probability judgements figures in decision making under ambiguity].
- Gardiner S. (2006). A Core Precautionary Principle. *The Journal of Political Philosophy* **14.1**, 33 – 60. [Addresses criticisms that the Precautionary Principle is confused by identifying its core commitments: cautionary decision-making in situations of high stakes and severe uncertainty].
- Genest C. and J. V. Zidek. (1986). Combining Probability Distributions: A Critique and Annotated Bibliography. *Statistical Science* **1.1**, 113 – 135. [Very useful survey of the literature on combining probability functions].
- Gilboa I. and M. Marinacci. (2012). Ambiguity and the Bayesian Paradigm. *Advances in Economics and Econometrics: Theory and Applications, Tenth World Congress of the Econometric Society* (eds. D. Acemoglu, M. Arellano and E. Dekel), 179 – 242 Cambridge: Cambridge University Press. [Useful survey of the (primarily economics) literature on decision making under uncertainty and especially ambiguity].
- Gilboa I. and D. Schmeidler. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics* **18**, 141 – 153. [Classic axiomatisation of the maximin expected utility model].
- Hall J. W., Lempert, R. J., Keller, K., Hackbarth, A., Mijere, C., McInerney, D. J. (2012). Robust Climate Policies Under Uncertainty: A Comparison of Robust Decision-Making and Info-Gap Methods. *Risk Analysis* **32.10**, 1657 – 1672. [Comparison of two methods for decision making under severe uncertainty that both turn on robustness to uncertainty in achieving a satisfactory decision outcome].
- Halpern J. Y. (2003). *Reasoning About Uncertainty*, 483 pp. Cambridge, MA: MIT Press. [A comprehensive analysis of different representations of uncertainty and associated logics of belief, learning and decision making].
- Hegerl G. C., O. Hoegh-Guldberg, G. Casassa, M. P. Hoerling, R. S. Kovats, C. Parmesan, D. W. Pierce, P. A. Stott. (2010). Good Practice Guidance Paper on Detection and Attribution Related to Anthropogenic Climate Change. *Meeting Report of the Intergovernmental Panel on Climate Change Expert Meeting on Detection and Attribution of Anthropogenic Climate Change* (eds. T. F. Stocker, C. B. Field, D. Qin, V. Barros, G.-K. Plattner, M. Tignor, P. M. Midgley and K. L. Ebi. Bern). Switzerland: IPCC Working Group I Technical Support Unit, University of Bern. [Clarifies methods, definitions and terminology for use by the IPCC Working Groups and outlines recommendations for good practice in detection and attribution studies].

- Held I. M. (2005). The Gap between Simulation and Understanding in Climate Modeling. *Bulletin of the American Meteorological Society* **80**, 1609 – 1614. [Expert perspective on the present and future role of climate models in developing both process resolution and physical understanding, advocating a “hierarchy of lasting value”].
- Hill B. (2013). Confidence and Decision. *Games and Economic Behavior* **82**, 675 – 692. [Develops an ordinal notion of confidence and a model for decision making under ambiguity that is sensitive to considerations of confidence].
- Hulme M., S. Dessai, I. Lorenzoni and D. Nelson. (2009). Unstable Climates: exploring the statistical and social constructions of climate. *Geoforum* **40**, 197 – 206. [Discusses statistical considerations as well as social factors that were involved in the formulation and acceptance of definitions of climate].
- IPCC. (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge and New York: Cambridge University Press. [The most recent assessment report about the physical science basis of climate change].
- IPCC. (2014). *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge and New York: Cambridge University Press. [The most recent assessment report about impact, adaptation, and vulnerability of climate change].
- Jeffrey R. (1965). *The Logic of Decision*, 231 pp. Chicago: University of Chicago Press [Provides an extensive introduction and overview of decision theory].
- Jun M., R. Knutti, D. W Nychka. (2008). Local eigenvalue analysis of CMIP3 climate model errors. *Tellus A* **60.5**, 992 – 1000. [Assessment of correlations in error patterns from climate models].
- Katzav J. (2014). The epistemology of climate models and some of its implications for climate science and the philosophy of science’ *Studies in History and Philosophy of Modern Physics* **46**, 228 – 238. [Discusses problems that arise when confirming models in climate science].
- Klibanoff P., M. Marinacci and S. Mukerji. (2005). A smooth model of decision making under ambiguity. *Econometrica* **73**, 1849 – 1892. [Presents an influential model of decision making under ambiguity that uses second-order probabilities to weight first-order risks].
- Knutti R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl. (2010). Challenges in Combining Projections from Multiple Climate Models. *Journal of Climate* **23.10**, 2739 – 2758. [Outlines the challenges in use and interpretation of climate model ensembles].
- Koopmans T. C. (1962). On flexibility of future preference. *Cowles Foundation for Research in Economics, Yale University, Cowles Foundation Discussion Papers* **150**. [Discusses and formalises the value of maintaining flexibility or opportunity in future choice situations that depend on one’s present choices].
- Kreps D. M. and E. L. Porteus. (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica* **46.1**, 185 – 200. [Discusses and formalises a preference for when uncertainty associated with an option is resolved over time].
- Lahsen M. (2005). Seductive Simulations? Uncertainty Distribution Around Climate Models. *Social Studies of Science* **35.6**, 895 – 922. [Drawing on participant observation and interviews with climate modelers, this study offers a science studies perspective on understanding uncertainty].
- Lehrer K. and Wagner, C. (1981). *Rational Consensus in Science and Society*, 165 pp. Dordrecht: Reidel. [Discusses the role of consensus in many domains and presents a model for the achievement of consensus through iterative linear averaging].
- Lempert R. J., Popper, S. W., Bankes, S. C. (2003). *Shaping the Next One Hundred Years: New Methods for Quantitative Long-Term Policy Analysis*, 208 pp. Santa Monica, CA: RAND Corporation, MR-1626-RPC. [Proposes a decision making approach that focuses on robustness to uncertainty in achieving a satisfactory decision outcome].
- Lenhard J. and E. Winsberg. (2010). Holism, entrenchment, and the future of climate model pluralism. *Studies in History and Philosophy of Modern Physics* **41**, 253 – 262. [Discusses confirmation holism in climate science].
- Linkov I. et al. (2014). Changing the resilience program. *Nature Climate Change* **4**, 407 – 409. [Discusses the need to manage resilience as well as risk, where the former concerns how quickly a system recovers from shocks rather than the incidence of shocks].
- List C. and C. Puppe. (2009). Judgment aggregation: a survey. *Oxford Handbook of Rational and Social Choice* (eds. P. Anand, C. Puppe and P. Pattanaik). Oxford: Oxford University Press. [Useful survey of the literature on judgement aggregation in both economics and philosophy].
- Lorenz E. (1995). Climate is what you expect. Prepared for publication by NCAR. Unpublished, 1 – 33. [Provides a detailed and comprehensive review of definitions of climate].
- Lloyd E. A. (2010). Confirmation and robustness of climate models. *Philosophy of Science* **77**, 971 – 984. [Discusses confirmation and robustness in climate science].
- Lloyd E. A. (2015). Model robustness as a confirmatory virtue: The case of climate science. *Studies in History and Philosophy of Science* **49**, 58 – 68. [Argues that robustness is a confirmatory virtue in climate science].
- Lloyd E. A. (2009). Varieties of Support and Confirmation of Climate Models. *Proceedings of the Aristotelian Society Supplementary Volume LXXXIII*, 217 – 236. [Discusses confirmation in climate science].
- Mann M. E., R. S. Bradley and M.K. Hughes (1998). Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* **392**, 779 – 787. [A study that gives a proxy-based reconstruction of the Northern Hemisphere temperature record, which gave rise to the Hockey Stick controversy].

- Maslin M. and P. Austin. (2012). Climate models at their limit?. *Nature* **486**, 183 – 184. [Argues that estimates of climate-change impacts will get less, rather than more, certain].
- Mastrandrea M. D., K. J. Mach, G.-K. Plattner, O. Edenhofer, T. F. Stocker, C. B. Field, K. L. Ebi, and P. R. Matschoss. (2011). The IPCC AR5 guidance note on consistent treatment of uncertainties: a common approach across the working groups. *Climatic Change* **108**, 675 – 691. [Perspectives on the application of the IPCC's recommendations].
- McGuffie K. and A. Henderson-Sellers. (2005). *A Climate Modelling Primer*, 217 pp. New Jersey: Wiley. [Provides an introduction into climate modelling].
- McIntyre S. and R. McKittrick. (2003). Corrections to the Mann et. al. (1998) proxy data base and northern hemispheric average temperature series. *Energy & Environment* **14.6**, 751 – 771. [A paper that argued that methods used to process the data are such that they would produce a hockey-stick-shaped curve from almost any data. The paper was later found to contain serious flaws].
- Mongin P. (1995). Consistent Bayesian Aggregation. *Journal of Economic Theory* **66.2**, 313 – 51. [Presentation of important possibility and impossibility results in the aggregation of probabilities and utilities].
- Nordhaus W. D. (2007). A Review of the Stern Review on the Economics of Climate Change. *Journal of Economic Literature* **45.3**, 686 – 702. [Identifies and critiques the key parameters of Stern's framework for assessing climate-change mitigation options, in particular the pure time discounting rate].
- Nordhaus W. C. (2008). *A Question of Balance*, 366 pp. New Haven, CT: Yale University Press. [Presents a model (an alternative to Stern's) for assessing climate-change mitigation options from the point of view of the global planner, called the Dynamic Integrated model of Climate and the Economy (DICE)].
- Oreskes N. and E. M. Conway. (2012). *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*, 355 pp. New York: Bloomsbury Press. [Tells the story of how a group of high-level scientists and advisers ran effective campaigns to mislead the public and deny well-established scientific knowledge, among others about climate change].
- Oreskes N. (2007) The Scientific Consensus on Climate Change: How Do We Know We're Not Wrong? *Climate Change: What It Means for Us, Our Children, and Our Grandchildren* (eds. J. F. C. DiMento and P. Doughman), 65 – 99. Boston: MIT Press. [Documents a widespread consensus on the reality of climate change in the scientific community].
- Oreskes N., K. Shrader-Frechette and K. Belitz. (1994). Verification, validation, and confirmation of numerical models in the Earth Science. *Science New Series* **263.5147**, 641 – 646. [Discusses the role of verification, validation and confirmation in climate science].
- Parfit D. (1984). *Reasons and Persons*, 560 pp. Oxford: Clarendon Press. [A classic work on temporal issues associated with rationality, personal identity and ethics; discusses potential rationales for future discounting in an appendix].
- Parker W. S. (2010). Comparative Process Tracing and Climate Change Fingerprints. *Philosophy of Science* **77**, 1083 – 1095. [Provides an in depth discussion of detection attribution studies from a philosophical point of view].
- Parker W. S. (2009). Confirmation and Adequacy for Purpose in Climate Modelling. *Aristotelian Society Supplementary Volume* **83.1** 233 – 249. . [Discusses confirmation in climate science and argues that climate models can only be confirmed relative to a certain purpose].
- Parker W. S. (2013). Ensemble modeling, uncertainty and robust predictions. *Wiley Interdisciplinary Reviews: Climate Change* **4.3**, 213 – 223. [Discusses the role of ensembles, uncertainty and robustness in climate science].
- Parker W. S. (2014). Values and Uncertainties in Climate Prediction, Revisited. *Studies in History and Philosophy of Science Part A* **46**, 24 – 30. [Discusses the role of ethical and social values in climate science].
- Parker W. S. (2011). When Climate Models Agree: The Significance of Robust Model Predictions. *Philosophy of Science* **78.4**, 579 – 600. [Discusses the significance of robust predictions in climate science].
- Petersen A. C. (2012). *Simulating Nature: A Philosophical Study of Computer-Simulation Uncertainties and Their Role in Climate Science and Policy Advice*, 210 pp. Boca Raton, Florida: CRC Press. [Offers a methodological reflection on climate modelling and the use of its results in policy contexts].
- Resnik M. (1987). *Choices: an introduction to decision theory*, 221 pp. Minneapolis: University of Minnesota Press. [An introductory text on decision theory, social choice theory and game theory].
- Savage L. J. (1954). *The Foundations of Statistics*, 310 pp. New York: John Wiley & Sons. [Presents a classic version of expected utility theory that can account for subjective uncertainty, measured by a probability function].
- Sen A. (1982). Approaches to the choice of discount rate for social benefit–cost analysis. *Discounting for Time and Risk in Energy Policy* (ed. R. C. Lind), 325 – 353. Washington, DC: Resources for the Future. [Discusses the shortcomings of using the market rate as a measure of the social rate for discounting the future].
- Sen A. (1970). *Collective Choice and Social Welfare*, 225 pp. San Francisco: Holden-Day Inc. [A classic early work on social choice and welfare theory, including relevant mathematical results and discussion of their significance].
- Sexton D. M. H., J. M. Murphy, M. Collins and M. J. Webb. (2012). Multivariate Probabilistic Projections Using Imperfect Climate Models. Part I: Outline of Methodology. *Climate Dynamics* **38**, 2513 – 2542. [Summarises the method used to generate the UKCP09 climate projections].
- Sexton D. M. H., and J. M. Murphy. (2012). Multivariate Probabilistic Projections Using Imperfect Climate Models. Part II: Robustness of Methodological Choices and Consequences for Climate Sensitivity. *Climate Dynamics* **38**, 2543 – 2558. [Assesses the Robustness of the UKCP09 climate projections].

- Shackley S., P. Young, S. Parkinson and B. Wynne. (1998). Uncertainty, Complexity and Concepts of Good Science in Climate Change Modelling: Are GCMs the Best Tools? *Climatic Change* **38**, 159 – 205. [Discusses the complexity, confirmation and the various sources of uncertainty in climate science].
- Smith L. A. and N. Stern. (2011). Uncertainty in science and its role in climate policy. *Phil. Trans. R. Soc. A* **369**, 4818 – 4841. [Discusses the impact of scientific uncertainty on policy making].
- Spiegelhalter D. J. and H. Riesch. (2011). Don't know, can't know: embracing deeper uncertainties when analysing risks. *Phil. Trans. R. Soc. A* **369**, 4730 – 4750. [Proposes a five-level structure for assessing and communicating uncertainties].
- Stainforth D. A., M. R. Allen, E. R. Tredger and L. A. Smith. (2007a). Confidence, Uncertainty and Decision-support Relevance in Climate Predictions. *Philosophical Transactions of the Royal Society A* **365**, 2145 – 2161. [Discusses the policy-relevance of modelling results and the various sources of uncertainty in climate science].
- Stainforth D. A., T. E. Downing, R. Washington, A. Lopez and M. New. (2007b). Issues in the Interpretation of Climate Model Ensembles to Inform Decisions. *Philosophical Transactions of the Royal Society A* **365**, 2163 – 2177. [Discusses how to interpret climate model ensembles].
- Steele K. (2006). The precautionary principle: a new approach to public decision-making?. *Law Probability and Risk* **5**, 19 – 31. [Discusses the ambiguity of the Precautionary Principle, and argues that it is best interpreted as a guide to framing public decision problems rather than an alternative choice rule].
- Steel D. P. (2008). *Across the Boundaries: Extrapolation in Biology and Social Science*, 256 pp. New York: Oxford University Press. [Provides an in-depth discussion of how causal knowledge can be carried over from one context into another one].
- Steele K. and C. Werndl. (2013). Climate Models, Confirmation and Calibration. *The British Journal for the Philosophy of Science* **64**, 609 – 635. [Discusses confirmation and estimation of parameters in climate science and argues that in the Bayesian framework it is fine to use the same data both for confirmation and estimation of parameters].
- Steele K. and C. Werndl. forthcoming (2015). The Need for a More Nuanced Picture on Use-Novelty and Double-Counting. *Philosophy of Science*. [Discusses confirmation and estimation of parameters in climate science and argues that model selection theory presents us with a more nuanced account about the use of data].
- Stern N. (2007). *The Economics of Climate Change: The Stern Review*, 692 pp. Cambridge: Cambridge University Press. [A highly influential analysis of the global costs of abating greenhouse gas emissions relative to continuing with 'business as usual'].
- von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behaviour*, 739 pp. Princeton: Princeton University Press. [A classic and seminal work on decision and game theory, including the theory of utility measures of preferences and solution concepts in game theory].
- Walley P. (1991). *Statistical Reasoning with Imprecise Probabilities*, 706 pp. New York: Chapman and Hall. [A comprehensive account of the foundations of imprecise probabilities as a representation of belief and basis for inference and decision].
- Weitzman M. L. (2009). On Modeling and Interpreting the Economics of Catastrophic Climate Change. *The Review of Economics and Statistics* **91.1**, 1 – 19. [Argues that global cost-benefit analyses of climate change are significantly altered when the probability for catastrophic consequences is non-negligible ('fat tailed' distribution) rather than negligible].
- Werndl C. (2015). On defining climate and climate change. *The British Journal for the Philosophy of Science*, doi:10.1093/bjps/axu48. [Provides a comprehensive and detailed assessment of various definitions of climate].
- Wilby R. L. and S. Dessai. (2010). Robust adaptation to climate change. *Weather* **65.7**, 180 – 185. [Introduce the widely discussed envelope of uncertainty and consider robust adaption to climate change].
- Weisberg Michael. (2006). Robustness Analysis. *Philosophy of Science* **73**, 730 – 742. [Offers a general philosophical discussion of robustness analysis].
- Winsberg E. (2012). Values and Uncertainties in the Predictions of Global Climate Models. *Kennedy Institute of Ethics Journal* **22**, 111 – 127. [Argues that climate science is not value free and that social and ethical values influence modelling in climate science].
- Winsberg E. and W. Goodwin (2015). Missing the Forest for the Fish. Draft in preparation. [Provide a critical discussion of Frigg et al (2014)].
- Worrall J. (2010). Error, Tests, and Theory Confirmation. *Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science* (eds. D. G. Mayo and A. Spanos), 125 – 154. Cambridge: Cambridge University Press. [Discusses the significance of 'use-novel' evidence and argues that evidence used to construct or fix the free parameters of a general theory cannot also provide support for the general theory].

Biographical Sketches

Richard Bradley hails from Johannesburg in South Africa. After completing a BA (Hons) in Political Science at the University of Witwatersrand, he did his MSc in Social Philosophy at the London School of Economics and Political

Science, graduating in 1988, and, after a period of employment in consulting, a PhD in Philosophy at the University of Chicago, graduating in 1997.

He is currently Professor of Philosophy in the Department of Philosophy, Logic and Scientific Method at the London School of Economics and Political Science, having joined the department in 1997. He is one of the editors of *Economics and Philosophy* and serves on a number of editorial and advisory boards. He publishes in decision theory, including both individual choice under uncertainty and social choice, in formal epistemology and on the semantics of conditionals.

Prof. Bradley is a member of the British Society for the Philosophy of Science, the American Philosophical Association and the Philosophy of Science Association.

Roman Frigg was born in Basel, Switzerland. He holds a PhD in Philosophy from the University of London and masters degrees both in theoretical physics and philosophy from the University of Basel.

He is Professor of Philosophy in the Department of Philosophy, Logic and Scientific Method, Director of the Centre for Philosophy of Natural and Social Science (CPNSS), and Co-Director of the Centre for the Analysis of Time Series (CATS) at the London School of Economics and Political Science. He is a permanent visiting professor in the Munich Centre for Mathematical Philosophy of the Ludwig-Maximilians-University Munich. He held visiting appointments in the Rotman Institute of Philosophy of the University of Western Ontario, the Descartes Centre for the History and Philosophy of the Sciences and the Humanities of the University of Utrecht, the Sydney Centre for the Foundations of Science of the University of Sydney, and the Department of Logic, History and Philosophy of Science of the University of Barcelona. He is associate editor of the *British Journal for the Philosophy of Science*, member of the steering committee of the European Philosophy of Science Association, and he serves on a number of editorial and advisory boards.

Professor Frigg is a member of the European Society for the Philosophy of Science, the Philosophy of Science Association, and the British Society for the Philosophy of Science.

Katie Steele was born in Ipswich, Australia. She completed a PhD in philosophy at the University of Sydney in 2007 and before that, a Masters degree in philosophy and an undergraduate degree in mathematics at the University of Queensland.

She is Associate Professor in the Department of Philosophy, Logic and Scientific Method and an affiliate of the Grantham Research Institute on Climate Change and the Environment at the London School of Economics. Previously she was a Postdoctoral Research Fellow at the University of Sydney. She has published papers on rational choice and scientific inference, and the interface between science and policy decision-making.

Erica Thompson was born in Scotland. She has a PhD in physics from Imperial College, London (2013), and an MMath from the University of Cambridge (2007).

She is presently a Research Officer in the Centre for the Analysis of Time Series within the Department of Statistics at the London School of Economics and Political Science. She has previously worked for the Grantham Institute for Climate Change at Imperial College and the UK Energy Research Centre. Her current research interests include quantification of scientific uncertainty, the use and utility of climate information for real-world decision-making, and the interpretation and communication of climate model output.

Dr Thompson is or has previously been a member of the Royal Meteorological Society, the American Geophysical Union, Scientists for Global Responsibility, and the British Beekeepers' Association.

Charlotte Werndl was born in Salzburg. She completed a PhD in Philosophy at the University of Cambridge in 2010 and master's degrees both in mathematics and philosophy at the University of Salzburg in 2006.

She is Professor of Logic and Philosophy of Science at the Department of philosophy at the University of Salzburg, Austria, a Visiting Professor at the Centre for Philosophy of Natural and Social Science (CPNSS) at the London School of Economics, and an affiliate of the Grantham Research Institute on Climate Change and the Environment at the London School of Economics. Previously she was an Associate Professor at the Department of Philosophy, Logic and Scientific Method at the London School of Economics and before that a research fellow at the University of Oxford. She is an editor of the *Review of Symbolic Logic* and an associate editor of the *European Journal for the Philosophy of Science* and serves on a number of editorial and advisory boards. She has published papers on climate change, statistical mechanics, mathematical knowledge, chaos theory, predictability, confirmation, evidence, determinism, indeterminism, observational equivalence and underdetermination. Her current work focuses on the philosophy of climate science, evidence and the philosophy of statistics and the foundation of statistical mechanics. Professor Werndl is a member of the Aristotelian Society, the British Society for the Philosophy of Science, the European Society for the Philosophy of Science, and the Philosophy of Science Association.