Forthcoming in Routledge Encyclopedia of Philosophy

Philosophy of Climate science

Richard Bradley, Roman Frigg, Katie Steele, Erica Thompson and Charlotte Werndl

Keywords: climate change, climate models, detection, attribution, uncertainty, projection, climate decisions, ensembles

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. The philosophy of climate science is a new sub-discipline of the philosophy of science that began to crystalize at the turn of the 21st century when philosophers started having a closer look at climate science. It comprises a reflection on almost all aspects of climate science, including observation and data, methods of detection and attribution, kinds of uncertainties, model ensembles, and decision-making under uncertainty. The philosophy of climate science operates in close contact with science itself and pays careful attention to the scientific details.

1. Climate and Climate Models

Global climate models (GCMs) are representations of the Earth's climate system. Climate science aims to develop GCSs that integrate as much scientific knowledge as possible. Nowadays there are dozens of global climate models under continuous development by national modelling centres like NASA, the UK Met Office, and the Beijing Climate, and by smaller institutions. For these models typically the Earth is divided into grid cells (in 2021 the horizontal grid scale is around 150km). The dynamics of the climate is then conceptualised as flows of physical quantities such as heat or vapour between the cells. Mathematically, these flows are described by equations. One also often tries to replace processes that are too small-scale to be physically represented in the model by adding a simplified process (one then says that *parameterisations* have been included in the model).

Since the resulting equations usually cannot be solved analytically, supercomputers are used. It may take weeks or months in real time to simulate a century of climate evolution. In order to compute a hypothetical evolution of the climate system, one also needs as input an initial condition (a mathematical description of the state at the beginning of the simulation period) and external conditions (factors such as the concentration of greenhouse gases, the amount of solar radiation received by the Earth etc. that affect the system but are not directly calculated by the simulations).

What is the climate that these models represent? Climate is a complex concept and how to define climate and climate change is controversial. To start with, it is important to distinguish climate from weather. Intuitively, the weather is about the state of the atmosphere. By contrast, climate is a *distribution* of particular variables (called the climate variables) arising for a particular configuration of the climate system. Climate variables concern domains like the atmosphere, the ocean, glaciers, and ice shields (IPCC 2013).

This basic idea can be made precise in different ways. Definitions of climate fall into two groups: those that define climate as a distribution over time and those that define climate as an ensemble distribution. The former posits that climate is the empirically observed distribution of the climate variables over a specific period of time. Climate change then is the difference between the distributions of two time periods. The advantages of this definition are its simplicity, its intuitive appeal and the fact that it is easy to estimate from the observations (Lorenz 1995).

A major drawback of this definition is that it is insensitive to abrupt changes. Suppose that in the middle of a period of time the Earth is hit by a meteorite and, as a consequence, becomes a much colder place. It is obvious that the climate before and after the hit of the meteorite is different. Yet, this definition does not tell us that this is so because it simply says that averages values should be taken over a preset period. To alleviate these problems, Werndl (2016) introduces the idea of regimes of varying external conditions. She then defines climate as the distribution of the climate variables arising under specific regimes of varying external conditions. The challenge here then is to say precisely what a regime of varying external conditions amounts to.

The ensemble approach defines climate with respect to an infinite collection of virtual copies of the climate system. This collection is called an *ensemble* of climate systems. Now consider all virtual copies where the present values of the climate variables are compatible with the values measured in the actual climate system. The climate at a time *t* is the distribution of the values of the climate variables that arises when all copies in the ensemble are evolved from now to that time *t* for the predicted path taken by the external conditions (Daron and Stainforth 2013).

While useful for predictive purposes, this definition faces several challenges. First, we think of climate as something out there in the world independent on our knowledge, but according to this definition climate is dependent on our knowledge (because of the measurement accuracy). Second, the definition is only about the future climate and it is unclear how the present and past climate should be defined. But without such definitions it is unclear how to think of climate change. Third, climate thus defined does not have any relationship to the past time series of observations, which seems counterintuitive (Werndl 2016).

2. Detection, Attribution, Projection

The three key research uses of climate models are detection, attribution and projection. *Detection* is the process of demonstrating that the climate has changed in some defined statistical sense without providing a reason for that change. *Attribution* of climatic change (once detected) is the identification of the causes of that change, such as anthropogenic emissions of greenhouse gases and natural internal variability. A *projection* of future climatic changes is a forecast given expected changes in greenhouse gas concentrations.

In the 2020s, detection of climate change is unequivocal in many aspects of the climate system and particularly in temperature changes at large scales. The challenge for detection is to define an appropriate null hypothesis (the expected behaviour in the absence of changing external influences), against which the observed outcomes can be tested. In practice, the best available null hypothesis is often provided by a state of the art GCM. As stressed by Parker (2010), detection is robust across different models, and there is a variety of evidence pointing to the conclusion that the global mean temperature has increased beyond that which can be attributed to internal variability. 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.

With detection becoming clearer , there is more focus on attribution to causal factors. An important method for attribution is optimal fingerprinting. It 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) notes that fingerprint studies rely on several assumptions, chief among them linearity.

Attribution studies have also tackled the question of individual extreme weather events: it is well-known that events such as hurricanes or heatwaves in general cannot be directly said to be “caused by climate change” but it is possible to statistically analyse the degree to which a similar event is more or less likely in a changed climate (van Oldenborgh et al. 2021). In the interpretation of attribution results, there is a tendency to focus on whether or not the confidence interval of the estimated anthropogenic effect crosses zero. This results in conservative attribution statements, reflecting public discourse where “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 (Lloyd and Oreskes 2019; Lusk 2017). As Parker (2010) argues, there is higher confidence in attribution results when the results are robust and there is a variety of evidence, as there is for the finding that late twentieth-century temperature increase was mainly caused by greenhouse gases.

Currently the most urgent task for climate models is projecting future climatic change to inform decisions about mitigation and adaptation, and other climate-related decisions. “Projection” is a technical term referring to a forecast that is conditional on a particular forcing scenario (and other conditions such as land use), specified either by the amount of greenhouse gas emissions and aerosols added to the atmosphere or directly by their atmospheric concentrations (Werndl 2018).

If using modelled projections to inform high-impact decisions, it is necessary to understand how accurate they are likely to be. There is no general answer to the question of the trustworthiness of model outputs. Variously one might consider the *ability to successfully reproduce past data* (Oreskes et al. 1994; Stainforth et al. 2007a; Katie Steele and Werndl 2013), *expert judgement* about the potential limitations of models (Thompson et al. 2016), *model-based understanding* of fidelity of the internal process representations and *physical understanding* of the theoretical behaviour of the climate system with which we expect models to be consistent. None of these constitute sufficient criteria. Parker (2009) urges a shift in thinking from confirmation to *adequacy for purpose*: models can only be found to be adequate for specific purposes, such as for global mean temperature projection in 2100, but they cannot be confirmed wholesale for all variables. Katzav (2014) cautions that adequacy for purpose assessments are of limited use, and that climate models can at best be confirmed as providing a range of possible futures.

Given these difficulties at global scale, there are further questions about the use of models as providers of detailed information about the future *local* climate. The question of whether high-resolution projections such as those generated by the *United Kingdom Climate Impacts Programme* (Sexton et al. 2012; Sexton and Murphy 2012) are trustworthy has sparked ongoing debate. Frigg et al. have urged caution (2014; 2015), where Winsberg (2018) and Winsberg and Goodwin (2016) criticise these arguments as overstating the limitations imposed by mathematical considerations. Meanwhile some agencies continue to produce detailed regional projections, and others are exploring different information provision including storyline approaches, which are based on information from climate models but without assigning detailed probabilities to outcomes (Shepherd 2019).

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*. When data have been used for calibration, the question arises whether the same data can be used again to confirm the model. Scientists and philosophers alike have argued that such *double-counting* is illegitimate (Lloyd 2010; 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 (2018) argue that model selection theory presents a more nuanced picture of the use of data than the commonly endorsed positions. Frisch (2015) cautions that Bayesian as well as other inductive logics can be applied in better and worse ways to real problems such as climate prediction.

3. Understanding Uncertainty

Uncertainty features prominently in discussions about climate models, and yet is a concept that is poorly understood with different authors disagreeing on what is meant by "uncertainty" and on how to classify different kinds of uncertainty (for different proposals see, for instance, Stern and Smith (2011) and Spiegelhalter and Riesch (2011)). Setting matters of classification aside, one of the main challenges is to measure and quantify uncertainty in climate projections. Model ensembles play a crucial role in this. Multi-model ensembles are sets of several different models which differ in mathematical structure and physical content. Such an ensemble is used to investigate how predictions of relevant climate variables vary (or do not vary) according to model structure and assumptions.

A model-result is *robust* if all or most models in the ensemble show the same result. If, for instance, all models in an ensemble show more than four degree increase in global mean temperature by the end of the century when run under a specific emission scenario, this result is robust across the specified ensemble. Does robustness justify increased confidence? Lloyd (2015) argues that robustness arguments are powerful in connection with climate models and lend credibility to core claims. Parker (2011), by contrast, argues that agreement does not warrant the conclusion that relevant claims are likely to be true. 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)).

When ensembles do not yield robust predictions, then the spread of results within the ensemble is sometimes used to estimate quantitatively the uncertainty of the outcome. In Fifth Assessment Report of the IPCC, for instance, result of CMIP5 ensemble are used produce probabilities for the climate sensitivity and increase in global mean temperature under various emission scenarios to lie in certain ranges. A problem with this approach is that current ensembles are "ensembles of opportunity", grouping together existing models. Even the best ensembles are not designed to systematically explore all possibilities. The IPCC acknowledges this limitation (see the discussion in Chapter 12 of IPCC (2013)) and thus downgrade the assessed likelihood of ensemble-derived ranges (see Thompson et al. 2016).

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 – referred to as a "non-discountable envelope" – that cannot be ruled out (Stainforth et al. 2007b). While less committal than the probability approach, also non-discountable envelopes raise questions. Why is it that the results in the envelope cannot be rules out?
Do results which 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 does not indicate the *complete* range of possibilities, making particular types of formalised decision-making procedures impossible. For a further discussion of these issues see Betz (2015).

The most prominent framework for communicating uncertainty to policy-makers is the IPCC's, which is used throughout the Fifth Assessment Report, is 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. A discussion of this framework can be found, for instance, in Wüthrich (2017).

4. Climate Decisions Under Uncertainty

What is the appropriate reaction to climate change? How much should we mitigate? To what extent should we adapt, and what form should it take? Should we build larger water reserves and other “fortifying” infrastructure? Should we adapt houses, and our social infrastructure more generally, to a higher frequency of extreme weather events like droughts, heavy rainfalls, tidal surges and heatwaves? The decisions that we make in response to these questions have consequences affecting people at different places and times. Moreover, the circumstances of many of these decisions involve severe uncertainty and disagreement, 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 what significance or value we should attach to these consequences. These considerations make climate decision-making both important and hard.

The standard model of decision-making under uncertainty, deriving from Savage (1954), 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 quantified by a single probability distribution 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.

Empirical uncertainty. There are two issues with confining all empirical uncertainty to a precise probability function over the state space. The first is that it is rather unnatural for complex decision problems. For instance, a mitigation decision problem might be usefully modelled with a state-space partition in terms of possible increases in average global temperature. In that case, the likelihood of the states would be conditional on the mitigation option taken; moreover, the consequence arising in each of the states would depend on further uncertain features of the world, not least the details of regional climates. The second issue is that using a precise probability function can misrepresent the severity of uncertainty. For instance, one may contend that the position of the scientific community is best represented by a precise probability distribution over average global temperature states, conditional on some mitigation option. But precise probabilities over the detailed outcomes associated with an option and each state, describing, ultimately, impacts on human welfare, is much less plausible. In response to this, many philosophers and statisticians advocate the use of *sets* of probability functions to represent uncertainty of varying severity, whether the uncertainty is due to evidential limitations or expert disagreement (see, e.g., Walley 1991). Roughly, the more severe the uncertainty, the more probability functions over the space of possibilities are needed to conjointly represent the epistemic situation.

Ethical Uncertainty. Decision makers face uncertainty not only about what will or could happen, but also about what value to attach to these possibilities. A major source of ethical uncertainty is how to distribute the costs and benefits of mitigation and adaptation amongst populations in different regions of the world. Such issues are “idealised away” in many policy-assessment models, including the global mitigation cost-benefit analyses of Stern (2007) and Nordhaus (2008). Other features of the value function used in these models have nevertheless been fiercely debated, in particular, the “pure rate of time preference” (see, Greaves 2017). This and other ethical uncertainty may be represented analogously to empirical uncertainty – by replacing the standard precise utility function with a set of possible utility functions.

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? One strategy is to reduce the severity of uncertainty by making further judgments of plausibility or confidence in candidate probability or utility functions. For instance, when these functions derive from different models or experts, the decision maker may regard some of these as more reliable than others, or else may simply discount some models or experts as insufficiently reliable. Variations on this approach can be found in Klibanoff et al. (2005), Gärdenfors and Sahlin (1988), Hill (2013), and Bradley et al. (2017). If pursuit of the first strategy does not fully resolve the decision maker’s uncertainty about what precise probability estimates to draw on, she can use an “imprecise” decision rule that generalises the idea of maximising expected utility. Numerous proposals for rules of this kind can be found in the decision theoretic literature (see Gilboa and Marinacci (2012) and Bradley (2017) for surveys). A prominent class of them recommend choosing cautiously in situations of severe uncertainty (cf., the “Precautionary Principle” discussed in, e.g., Steele (2006)). For instance, the influential Maxmin-EU rule recommends picking the action with greatest minimum expected utility (Gilboa and Schmeidler 1989). Other rules weigh caution against other considerations or instead achieve the same effect as caution by recommending actions that achieve satisfactory outcomes “robustly”, i.e., relative to every probability estimate in the set of those deserving consideration (see, e.g., Ben-Haim 2001).

Biography and Further Reading

Ben-Haim, Y. (2001) *Information-Gap Theory: Decisions Under Severe Uncertainty*, London: Academic Press. (This book presents a new theory of decision making under severe uncertainty.)

Betz, G. (2015) ‘Are Climate Models Credible Worlds? Prospects and Limitations of Possibilistic Climate Prediction’, *European Journal for Philosophy of Science* 5 (2): 191–215. (This article explores three different arguments justifying that climate models are imperfect credible worlds.)

Bradley, R. (2017) *Decision Theory with a Human Face*, Cambridge: Cambridge University Press. (This book aims to develop a decision theory that is tailored for “real” agents.)

Bradley, R., Helgeson, C., & Hill, B. (2017) ‘Climate Change Assessments: Confidence, Probability and Decision,’ *Philosophy of Science* 84 (3): 500-522. (This article applies a formal decision-making perspective to an investigation of how the uncertainty framework of the IPCC might inform decisions.)

Daron, J. D., & Stainforth, D. A. (2013) ‘On predicting climate under climate change’, *Environmental Research Letters* 8: 1-8. (This article explores implications of ensemble size for predicting climate change, arguing that small ensembles can be misleading and that ensembles of several hundred members may be required).

Frigg, R., Bradley, S., Du, H., & Smith, L. A. (2014) ‘The adventures of Laplace's demon and his apprentices’, *Philosophy of Science* 81 (1): 31–59. (The paper discusses the notion of structural model error and argues that it can limit a model's ability to generate decision-relevant pre-dictions.)

Frigg, R., Smith, L. A., & Stainforth, D. A. (2015) 'An Assessment of the Foundational Assumptions in High-Resolution Climate Projections: The Case of UKCP09', *Synthese* 192: 3979–4008. (This paper describes and analyses the methodology used in the United Kingdom Climate Impacts Programme's UKCP09 project and urges some caution about its results.)

Frisch, M. (2015) 'Modeling Climate Policies: A Critical Look at Integrated Assessment Models', *Philosophy and Technology* 26: 117-137. (This paper critically examines neo-classical economic cost-benefit analysis and so-called integrated assessment models that were developed to tackle questions concerning intergenerational justice.)

Gärdenfors, P., & Sahlin, N.-E. (1988) 'Unreliable probabilities, risk taking, and decision making. In P. Gärdenfors, & N.-E. Sahlin (Eds.), *Decision, Probability and Utility*, Cambridge: Cambridge University Press, 313-334. (This paper presents a decision theory which is essentially Bayesian in its approach but which takes epistemic reliability of decision situations into consideration.)

Gilboa, I., & Marinacci, M. (2012) 'Ambiguity and the Bayesian Paradigm. In D. Acemoglu, M. Arellano, & E. Dekel (Eds.), *Advances in Economics and Econometrics: Theory and Applications (Tenth World Congress of the Econometric Society)*, Cambridge: Cambridge University Press, 179-242. (This chapter provides a survey of some of the recent decision-theoretic literature involving beliefs that cannot be quantified by a Bayesian prior.)

Gilboa, I., & Schmeidler, D. (1989) 'Maxmin expected utility with non-unique prior', *Journal of Mathematical Economics* 18: 141-153. (This paper provides a mathematical analysis of maximim expected utility when there is no unique prior.)

Greaves, H. (2017) 'Discounting for public policy: A survey', *Economics and Philosophy* 33 (3): 391-439. (This article surveys the debate on the social discount rate, with a focus on the economics literature.)

Hill, B. (2013) 'Confidence and Decision', *Games and Economic Behavior* 82: 675–692. (The article proposes a model of confidence in beliefs and uses it to develop a rule for making decisions under uncertainty.)

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 [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)] (This is the 2013 report prepared by the Intergovernmental Panel on Climate Change (IPCC). The report provides a comprehensive assessment of knowledge on climate change, its causes, potential impacts and response options.)

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. (This paper argues that Parker's adequacy-for-purpose approach to assessing climate models typically cannot be implemented. As an alternative, a possibilist approach to assessing climate models is defended.)

Klibanoff, P., Marinacci, M., & Mukerji, S. (2005) 'A smooth model of decision making under ambiguity. ', *Econometrica* 73: 1849-1892. (This paper develops a model for making decisions under ambiguity.)

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., & Meehl, G. A. (2010) ‘Challenges in combining projections from multiple climate models’, *Journal of Climate* 23: 2739-2758. (This paper discusses the various challenges arising when combining projections from multiple climate models.)

Lloyd, E. A. (2010) ‘Confirmation and Robustness of Climate Models’, *Philosophy of Science* 77 (5): 971-984. (This paper reviews various ways of how climate models are confirmed, including model fit, variety of evidence, independent support for aspects of the model, and robustness.)

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. (This paper argues that robustness is a confirmatory virtue: climate models gain support from robust results.)

Lloyd, E. A., & Oreskes, N. (2019) ‘Climate Change Attribution: When Does it Make Sense to Add Methods?’, *Epistemology & Philosophy of Science* 56 (1): 185-201. (This paper sheds light on the question of what methods should be used to attribute climate change.)

Lorenz, E. (1995) ‘Climate is what you expect.’, *Prepared for publication by NCAR. Unpublished*: 1-33. (This paper provides a classic discussion of various possible ways to define climate and climate change.)

Lusk, G. (2017) ‘The Social Utility of Event Attribution: Liability, Adaptation, and Justice-Based Loss and Damage’, *Climatic Change* 143: 201-212. (Probabilistic event attribution aims to attribute weather events to anthropogenic forcings. This paper investigates to what extent the development of this methodology is motivated by social utility rather than scientific understanding.)

Mastrandrea, M. D., Mach, K. J., Plattner, G.-K., Edenhofer, O., Stocker, T. F., Field, C. B., et al. (2011) ‘The IPCC AR5 guidance note on consistent treatment of uncertainties: a common approach across the working groups’, *Climatic Change* 108: 675-691. (This paper discusses the main features of the uncertainty framework for communicating climate results proposed by the IPCC.)

Nordhaus, W. C. (2008) *A Question of Balance*, New Haven: Yale University Press. (In this influential book the 2018 nobel laureate for economics aims to integrate the entire spectrum of economic and scientific research to weigh the costs of reducing emissions against the benefits of reducing long-run damages from global warming.)

Oreskes, N., Shrader-Frechette, K., & Belitz, K. (1994) ‘Verification, validation, and confirmation of numerical models in the Earth Science’, *Science New Series* 263 (5147): 641-646. (This paper provides an overview over philosophical questions that arise in climate science, such as verification, validation and confirmation.)

Parker, W. (2009) ‘Confirmation and adequacy-for-purpose in climate modeling’, *Aristotelian Society Supplementary Volume* 83 (1): 233–249. (This article argues that what matters is that climate models are adequate for their purpose and then discsusses in detail what this involves.)

Parker, W. (2010) ‘Comparative Process Tracing and Climate Change Fingerprints’, *Philosophy of Science* 77 (5): 1083-1095. (This article examines climate change fingerprint studies that investigate the causes of climate change. It is argued that these studies have much in common with Steel's comparative process tracing).

Parker, W. (2011) ‘When Climate Models Agree: The Significance of Robust Model Predictions’, *Philosophy of Science* 78 (4): 579-600. (This article examines conditions relating to truth, evidence and security under which robust model predictions in climate science have epistemic significance).

Savage, L. J. (1954) *The Foundations of Statistics*, New York: John Wiley & Sons. (A classic book providing an introduction into Bayesian statistics.)

Sexton, D. M. H., Murphy, J. M., Collins, M., & Webb, M. J. (2012) ‘Multivariate probabilistic projections using imperfect climate models part I: Outline of methodology’, *Climate Dynamics* 38: 2513-2542. (This is the first part of two-part paper explicating the conceptual foundations of the United Kingdom Climate Impacts Programme’s UKCP09 project.)

Sexton, D. M. H., & Murphy, J. M. (2012) ‘Multivariate probabilistic projections using imperfect climate models part II: Robustness of methodological choices and consequences for climate sensitivity’, *Climate Dynamics* 38: 2543-2558. (This is the second part of two-part paper explicating the conceptual foundations of the United Kingdom Climate Impacts Programme’s UKCP09 project.)

Shepherd, T. G. (2019) ‘Storyline approach to the construction of regional climate change information’, *Proceedings of the Royal Society A* 475 (2225): 20190013. (This paper argues that for regional climate decisions one should focus on the question of the impacts of particular actions under uncertain regional climate change.)

Smith, L. A., & Stern, N. (2011) ‘Uncertainty in science and its role in climate policy’, *Philosophical Transactions of the Royal Society A* 369: 1-24

Spiegelhalter, D. J., & Riesch, H. (2011) ‘Don’t know, can’t know: embracing deeper uncertainties when analysing risks’, *Phil. Trans. R. Soc. A* 369: 4730-4750. (This paper developes a five-level structure for assessing and communicating uncertainties.)

Stainforth, D. A., Allen, M. R., Tredger, E. R., & Smith, L. A. (2007a) ‘Confidence, uncertainty and decision-dupport relevance in climate predictions’, *Philosophical Transaction of the Royal Socity A* 365 (1857): 2145-2161. (This paper focuses on complex climate models as predictive tools on decadal and longer time scales and argues that the role of these models needs to be reassessed; sources of uncertainty and experimental strategies to quantify uncertainties are discussed.)

Stainforth, D. A., Downing, T. E., Washington, R., Lopez, A., & New, M. (2007b) ‘Issues in the interpretation of climate model ensembles to inform decisions’, *Philosophical Transactions of the Royal Society A* 365: 2163-2177. (This paper discusses how recent developments of large ensembles of climate model simulations can be interpreted and how they can inform societal decisions concerning mitigation and adaption.)

Steele, K. (2006) ‘The precautionary principle: a new approach to public decision-making?’, *Law Probability and Risk* 5: 19-31. (This article argues that the value of the precautionary principle lies in emphasizing aspects of good decision-making that go beyond the scope of formal decision theory and that it is best conceived as providing guidelines for formulating a decision problem.)

Steele, K., & Werndl, C. (2013) ‘Climate Models, Confirmation and Calibration’, *The British Journal for the Philosophy of Science* 64: 609-635. (This paper analyses double-counting – i.e. using data to both calibrate and confirm a model – in climate science and argues that it is legitimate from a Bayesian perspective.)

Steele, K., & Werndl, C. (2018) ‘The Need for a More Nuanced Picture on Use-Novelty and Double-Counting’, *The British Journal for the Philosophy of Science* 69 (2): 351–375. (This paper analyses double-counting and use-novelty in model selection theory, and it argues that a more nuanced picture is needed to make sense of the practices in model-selection theory.)

Stern, N. (2007) *The Economics of Climate Change: The Stern Review*, Cambridge: Cambridge University Press. (This review discusses the effect of global warming on the world economy. Although not the first economic report on climate change, it is significant as the largest and most widely known and discussed report of its kind).

Thompson, E., Frigg, R., & Helgeson, C. (2016) 'Expert Judgment for Climate Change Adaptation', *Philosophy of Science* 83 (6): 1110–1121. (The paper argues that expert judgment should play a central role in the provisionof local climate change projections intended to support decision-making.)

van Oldenborgh, G. J., van der Wiel, K., Kew, S., & al., e. (2021) 'Pathways and pitfalls in extreme event attribution', *Climatic Change* 166 (13). (This paper is about extreme event attribution – i.e. estimating which factors lead to an observed extreme weather or climate event – and assesses the methodology developed by the World Weather Attribution collaboration to find those factors).

Walley, P. (1991) *Statistical Reasoning with Imprecise Probabilities*, London: Chapman and Hall. (This book provides an examination of topics involved in statistical reasoning with imprecise probabilities).

Werndl, C. (2016) 'On defining climate and climate change', *The British Journal for the Philosophy of Science* 67 (2): 337-364. (This article discusses how to best define climate and climate change and proposes a new definition.)

Werndl, C. (2018) 'Initial Conditions Dependence and Initial Conditions Uncertainty in Climate Science', *The British Journal for the Philosophy of Science* https://doi.org/10.1093/bjps/axy021. (This article discusses initial conditions dependence and initial condition uncertainty of predictions and projections in climate science).

Winsberg, E. (2018) *Philosophy and Climate Science*, Cambridge: Cambridge University Press. (This book provides an introduction to debates in the philosophy of climate science.)

Winsberg, E., & Goodwin, W. M. (2016) 'The adventures of climate science in the sweet land of idle arguments', *Studies in History and Philosophy of Modern Physics* 54: 9–17. (This paper offers a criticial discussion of the argument in Frigg et al. 2014.)

Worrall, J. (2010) 'Error, Tests, and Theory Confirmation. In D. G. Mayo, & A. Spanos (Eds.), *Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science*, Cambridge: Cambridge University Press, 125-154.(This article discusses the use of data, in particula, use-novelty and double-counting in science.)

Wüthrich, N. (2017) 'Conceptualizing Uncertainty: An Assessment of the Uncertainty Framework of the Intergovernmental Panel on Climate Change. In M. Massimi, J.-W. Romeijn, & G. Schurz (Eds.), *EPSA15 Selected Papers*, Cham: Springer, 95-107. (This paper discusses the uncertainty framework proposed by the IPCC.)