

TITLE: The Myopia of Imperfect Climate Models: The Case of UKCP09

AUTHORS AND AFFILIATIONS:

Roman Frigg

Department of Philosophy, Logic and Scientific Method, LSE
Centre for Philosophy of Natural and Social Science (CPNSS), LSE
Email: r.p.frigg@lse.ac.uk

Leonard A. Smith

Centre for the Analysis of Time Series (CATS), LSE
Pembroke College Oxford
Email: lenny@maths.ox.ac.uk

David A. Stainforth

The Grantham Research Institute on Climate Change and the Environment, LSE
Centre for the Analysis of Time Series (CATS), LSE
Email: d.a.stainforth@lse.ac.uk

ACKNOWLEDGMENTS

Work for this paper has been supported by the LSE's *Grantham Research Institute on Climate Change and the Environment* and the *Centre for Climate Change Economics and Policy* funded by the Economics and Social Science Research Council and Munich Re. Frigg further acknowledges financial support from the AHRC-funded 'Managing Severe Uncertainty' project and grant FF12012-37354 of the Spanish Ministry of Science and Innovation (MICINN).

ABSTRACT

The *United Kingdom Climate Impacts Program's* UKCP09 project makes high-resolution forecasts of climate during the 21st century using state of the art global climate models. The aim of this paper is to introduce and analyze the methodology used and then urge some caution. Given the acknowledged systematic errors in all current climate models, treating model outputs as decision relevant probabilistic forecasts can be seriously misleading. This casts doubt on our ability, today, to make trustworthy, high-resolution predictions out to the end of this century.

1. Introduction

There is now a widespread consensus that global warming is real and in large part due to human activities.¹ But knowing that the climate is getting warmer *on average* is of limited use in designing detailed adaptation strategies.² The impact of climate change on humans occurs at a local scale, and so ideally we would like to know what changes we have to expect in our immediate environment, and reliable answers would greatly aid decision makers (Oreskes et al. 2010; Sexton et al. 2012; Smith and Stern 2011; Tang and Dessai 2012).

The *United Kingdom Climate Impacts Program's* UKCP09 project aims to answer exactly such questions by making high-resolution forecasts of 21st century climate.³ It generates its predictions using state of the art global climate models. The IPCC has confidence that these models have some skill at continental scales and above. This leaves open the question whether decision relevant high resolution predictions could be constructed with today's models.

The aim of this paper is to introduce and analyze the methodology used by UKCP09 and then urge some caution. Given the acknowledged systematic errors in all current climate models, treating model outputs as the basis for decision relevant probabilistic forecasts can be seriously misleading (Stainforth et al. 2007). This casts doubt on our ability, today, to make trustworthy,⁴ high resolution predictions out to the end of this century.

Herein we introduce the aims of UKCP09 (Section 2), outline the method used to generate predictions (Section 3), discuss the project's handling of structural model error (Section 4), argue that crucial assumptions are untenable (Section 5), and then draw some conclusions (Section 6).

¹ The existence of a wide-spread a consensus is documented in (Oreskes 2007); the evidence for the warming being anthropogenic is documented in the last IPCC Report. Throughout 'IPCC' refers to (Solomon et al. 2007).

² It may well be enough for mitigation: knowing even roughly what is likely to happen may be reason enough not to go there.

³ 'UKCP' stands for 'United Kingdom Climate Projections' and '09' indicates that it was launched for public use in 2009. UKCP09 is documented in the Briefing Report (Jenkins et al. 2009), the Science Report (Murphy et al. 2010) and two recent papers, (Sexton et al. 2012) and (Sexton and Murphy 2012). The full set of predictions is at <http://ukclimateprojections.defra.gov.uk/>.

⁴ In this paper we use the word 'trustworthy' to denote probability forecasts which one might rationally employ for decision making purposes using probability theory in the standard way. Such probability forecasts are expected to be robust and reliable, the kind a good Bayesian would make. We wish to avoid the kind of analysis that inspired Rubin's remark that 'a good Bayesian does better than a non-Bayesian, but a bad Bayesian gets clobbered' (cited in Good 2009, 139). There may be many justifiable and interesting scientific reasons to construct probability forecasts, our criticism of them in this paper is only in regard to their direct use in decision support (as, for instance, illustrated in the worked examples of UKCP09).

2. UKCP: Aims and Results

The declared aim of UKCP09 is to provide decision-relevant forecasts on which industry and policy makers can base their future plans:

‘To adapt effectively, planners and decision-makers need as much good information as possible on how climate will evolve, and supplying this is the aim of [...] UKCP09. They are one part of a UK government programme of work to put in place a new statutory framework on, and provide practical support for, adaptation.

The projections have been designed as input to the difficult choices that planners and other decision-makers will need to make, in sectors such as transport, healthcare, water-resources and coastal defences, to ensure that UK is adapting well to the changes in climate that have already begun and are likely to grow in future.’ (Jenkins et al. 2009, 9)

In a system as complex as the world’s climate, it is absurd to produce point forecasts (i.e. forecasts saying that a particular event will happen at a particular time with certainty). UKCP09 produces what they dub Bayesian probability forecasts, which ‘assign a probability to different possible climate outcomes recognizing that [...] giving a range of possible climate change outcomes is better, and can help with robust adaptation decisions, but would be of limited use if we could not say which outcomes are more or less likely than others.’ (ibid., 23)

The challenges many decision makers have to address arise at a local level: flood barriers have to be built in a particular location and to a given height, and so on. For this reason, local user-relevant information about the impacts of climate change is the most useful, assuming of course that it is not mis-informative (Smith and Stern 2011).

UKCP09 strives to meet the demand for decision-relevant information at the local level by producing highly specific information (Jenkins et al. 2009, 6-7). Probabilities are given for events on a 25km grid. Forecasts are made for finely defined specific events such as changes in the temperature of the warmest day of a summer and the precipitation of the wettest day of a winter. It is predicted, 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 (ibid., 36).

3. The Architecture of UKCP09

These predictions are generated with a method involving both global climate models (GCM’s) and elaborate statistical techniques. In this section we outline the method, in five parts, with the aim of identifying key assumptions and making its architecture visible.⁵

⁵ Our account of the method is based on (Murphy et al. 2010, Ch. 3).

Part 1 – Modelling. The cornerstone of UKCP09 is HadCM3, a GCM developed at the Hadley Centre. The model consists of two coupled modules, one representing the earth's atmosphere (including land surface processes and surface-atmosphere exchanges) and one representing the oceans. Our best descriptions of these fluids come from nonlinear partial differential equations (PDE's), which define the evolution of continuous fields representing the atmosphere or ocean. It is neither possible to integrate PDE's exactly, nor to measure the continuous fields required to initialize them. Instead they are discretized onto a grid in space and in time. Today's computational constraints force climate models to use a relatively coarse grid; those used in UKCP09 have a typical resolution of around 300km.

The model includes 10,000s of dynamical variables and 100s of parameters specifying the value of physical magnitudes, representing physical constants or controlling small scale processes which are not resolved explicitly. To aid the discussion let us introduce some notation. Let $x(t) = \{x_1(t), x_2(t), \dots\}$ be the vector of all dynamical variables and $\alpha = \{\alpha_1, \alpha_2, \dots\}$ the vector of all parameters in the model; then let $\phi_t^C(x; \alpha)$ be time evolution of HadCM3, specifying the future value of the system's dynamical variables given certain initial conditions and certain parameter values.

Even state-of-the-art computers take a long time to make a run of $\phi_t^C(x; \alpha)$, and so a simpler model is needed for most calculations (a 'run' is the calculation of the future value of x given a particular initial condition and a set of parameter values). To this end the entire ocean module is eliminated and replaced by a so-called *slab ocean*. With no currents and a uniform effective depth this slab ocean is defined via simpler equations. The result is HadSM3.⁶ We write $\phi_t^S(x; \alpha)$ to denote the time evolution of this model, where we take it as understood that the vectors x and α vary with the model structure (HadSM3 having fewer variables and parameters than HadCM3).

Part 2 – PPE. The problem in determining the future values of x is that 'the available information is seldom precise enough to allow the appropriate value of a given parameter to be accurately known' (Murphy et al. 2010, 37). Not knowing what value of α to use in our calculations, assuming there is one, 'gives rise to the parameter component of model error' (ibid.).⁷

The technique of a *perturbed physics ensemble* (PPE) is designed to address this difficulty.⁸ The idea of a PPE is to calculate future values of x for a number of different values of α . If, for instance, modelers are uncertain about the reasonable value of parameter α_2 but believe that it lies between $a_{2,\min}$ and $a_{2,\max}$, they carry out

⁶ Going from HadCM3 to HadSM3 roughly doubles the speed of the model.

⁷ This assumption is controversial. Smith (Smith 2006) argues that for imperfect models appropriate values (leading to trustworthy forecasts) may not exist. For want of space we set these worries aside; for more on this point see (Smith and Stern 2011).

⁸ We note in passing the lack of unanimity on whether the second 'P' of PPE stands for 'parameter', 'parameterization', or 'physics'. For more on PPE's see (Allen and Stainforth 2002).

calculations of x for as many values in the interval $[a_{2,\min}, a_{2,\max}]$ as they can afford. The variability of the outcomes then gives them a sense of the sensitivity of the model. Calculating future x 's for a number of different parameter values amounts to constructing a PPE because the variation of the parameter values amounts to perturbing the physics yet without changing the mathematical structure of the model (because all equations remain functionally unchanged).

In a complex model like HadCM3 a single research center can only make a relatively small number of runs due to the limitation of computational resources. The question then is how to construct a PPE for a model with 100s of parameters if only a small number of runs can be made. UKCP09 solves this problem by first restricting attention to atmospheric parameters and then soliciting parametrisation experts to identify those parameters that control the crucial processes in the system and on which the future values of x depend most sensitively. This process led to the identification of 31 crucial parameters and the definition of associated plausible intervals for them.

To explore the uncertainty of future values of x brought about by the variation in these 31 parameters, 280 runs were made with HadSM3. Information from 17 HadCM3 runs was added later.

Part 3 – Emulator. Unfortunately this number remains too small to provide a good understanding of the diversity of outcomes. An *emulator* is therefore built to provide values of x corresponding to values of α for which no runs were made. In other words, the emulator ‘fills the gaps’ between the 280 points obtained in Part 2.

Part 4 – Probability. What is the uncertainty of future values of x given the diversity in α ? Uncertainty is quantified by giving a probability distribution over the interval associated with α . The emulator correlates every value of α with an outcome x , and these distributions are translated into probabilities for x . UKCP09 assumes that each value of α within the middle 75% of the interval is equally likely and that the probability linearly drops to zero at the minimum and maximum values. These probabilities are then adjusted by assigning relative weights to all values of α according to the emulator-implied ability of the model to represent observations when simulating a period similar to the past.

Part 5 – Downscaling. The model calculations are done with a resolution near 300km but predictions are sought at 25km scale. To generate predictions at that level of detail the results are downscaled using simulations of a limited area regional climate model configured from HadCM3 and run at 25 km horizontal resolution.

The endeavors of these five parts taken together produce the predictions we have seen in the last section.

4. Structural Model Error

Each of these steps raises potentially significant conceptual and methodological questions. For want of space we can only deal with what we see as one central problem in Part 1: *structural model error* (SME). Like every model, HadCM3 has its imperfections. In order to specify $\phi_t^C(x; \alpha)$ a number of strongly idealising assumptions are made. These include distortion of the topography of the earth (mountain ranges like the Andes are systematically too smooth and too short, small volcanic islands chains with visible impacts on atmospheric circulation do not exist), and approximations of the effects of cloud fields which cannot be simulated realistically at the available resolution. Furthermore, solutions of the discretized PDE differ from those of the original PDE, and the PDE itself differs from what the true equations of the world would be (assuming such equations exist at all). In addition, there are limitations to our scientific understanding of the climate system and there may be relevant factors and processes about which we are simply unaware which would lead us to alter the equations of the model even under our current computational constraints.

Inasmuch as SME is due to shortcomings in the equations of the model, the challenges it poses to forecasting cannot be resolved by varying the model's parameters. If a model has SME this means that the time evolution of an ensemble will, eventually, differ from that of a better model and indeed reality itself, if a relevant distribution can be associated with reality. No adjustment of the parameters can remove this difference. The crucial question is: how soon do dramatic effects of SME manifest themselves in a given situation? And to what extent can a model with SME still be informative about the target system? On what timescales does the science (which underlies the model) suggest that a decision maker should 'expect' a big surprise if he took the model outputs as trustworthy?

UKCP09 acknowledges the presence of SME and proposes a way to deal with it. The message is that the uncertainties due to SME can be estimated and taken into account in projections.⁹ In this section we outline their approach, and in the next we ask whether its use for the provision of quantitative decision support is justified.

UKCP09 aims to capture the difference between the model and the real world with a so-called *discrepancy term*, which

'represents how informative the climate model is about the true climate, and it measures the difference between the climate model and the real climate that cannot be resolved by varying the model parameters. Such differences could arise from processes which are entirely missing from the climate model, or from fundamental deficiencies in the representation of processes which are included, through (say) limited resolution or the adoption of an erroneous assumption in the parameterisation scheme.' (Sexton et al. 2012, 2515)¹⁰

⁹ The UKCP09 science report calls the proposed method 'an appropriate means of quantifying uncertainties in projected future changes' (Murphy et al. 2010, 66).

¹⁰ See also (Murphy et al. 2010, 63-64).

Assume, then, that we are interested in ‘the true climate’¹¹ at a particular future instant of time t^* (for instance August 2080) and let c be the *true* value of x at t^* (hereafter ‘target’). The relation between the model output and the target then is:¹²

$$c = \phi_{t^*}^S(x_0; \alpha^*) + d,$$

where α^* the set of parameter values that best simulates the target. The discrepancy d is a vector in the system’s state space, and it can be interpreted as telling us ‘what the model output would be if all the inadequacies in the climate model were removed, without prior knowledge of the observed outcome’ (ibid.).

Now follow two crucial assumptions. The first assumption is ‘that the climate model *is* informative about the real system and the discrepancy term can be seen as a measure of how informative our climate model is about the real world’ (ibid., original emphasis). The idea is that informativeness comes in degrees and is indirectly proportional to the length of d : the smaller d the more informative the model. That the model is informative then amounts to assuming that d is small. We call this the *informativeness assumption*. The second assumption concerns the discrepancy. While d is defined as a vector, in practice one cannot know the exact vector and so it is assumed that there is probability distribution ε over d . This distribution is then assumed to be Gaussian (ibid.). We refer to the package of the two as the *Core Assumption*.

With this assumption in place UKCP09 sets out to estimate the parameters of ε . Not being omniscient, one cannot just compare model outputs with the truth. The crucial move in UKCP09 is to use a multi model ensemble (MME) as a proxy for the truth: ‘Our key assumption is that sampling the effects of structural differences between the model chosen for the PPE and alternative models provides a reasonable proxy for the effects of structural errors in the chosen model relative to the real world.’ (Sexton et al. 2012, 2516)¹³

The MME in question contains 12 models (Sexton et al. 2012, 2519). The claim then is that measuring the average distance of HadSM3 to a set of different models yields a similar result as measuring its distance to the real world – hence, d can be determined by measuring by how much HadSM3 diverges from those other models. We call the view that an MME is a trustworthy proxy for the real world the *proxy assumption*.

¹¹ There is, obviously, a serious confusion that we cannot clarify here: climate is a distribution, the state of the atmosphere at a given time (‘weather’) is a point. Our uncertainty in the ‘true’ (we would prefer ‘target’) value of x at time t does not correspond to the climate at time t .

¹² See (Sexton et al. 2012, 2521). Throughout we use our own notation, which differs from Sexton’s.

¹³ See also (Sexton et al. 2012, 2526) and (Murphy et al. 2010, 64).

For each model in the ensemble α^* , the best HadSM3 analogue, is determined. Having found the best analogue, the prediction error b is calculated; essentially the difference between the two model outputs. With these b 's the mean and variance of ϵ is determined.¹⁴

Under the proxy assumption, this procedure quantifies the additional uncertainty due to model error. This uncertainty is now added to the uncertainty about values of x obtained in Part 4, yielding the total uncertainty. The uncertainty is expressed as a probability distribution; it is this distribution that is presented as guidance for decision makers.

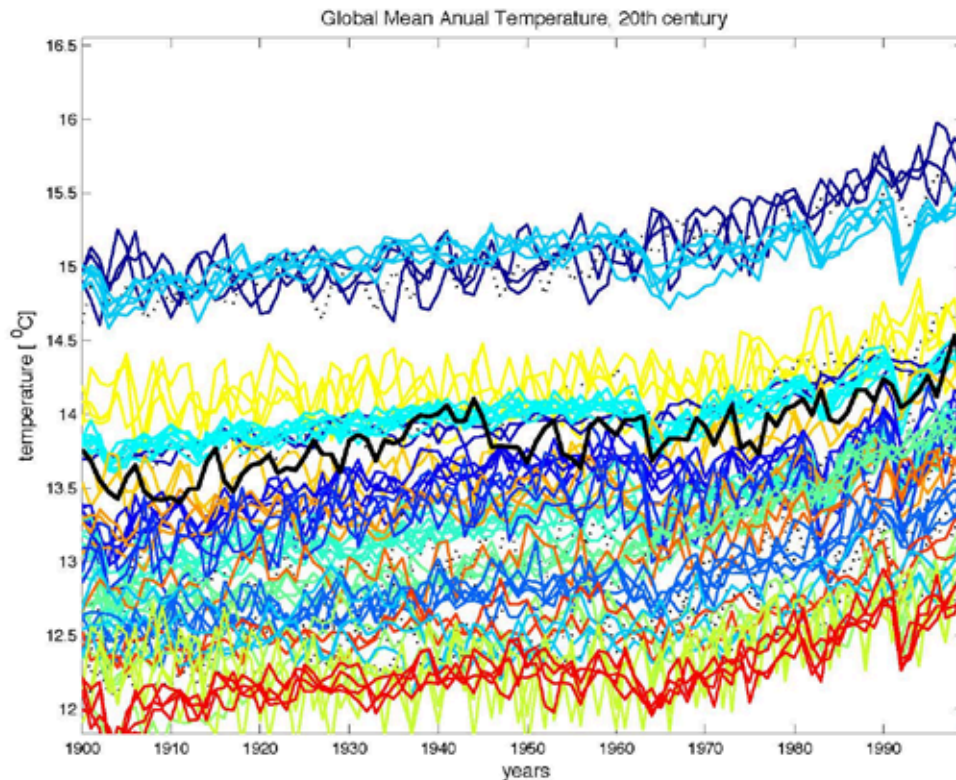
In the following section we argue that neither the core assumption nor the proxy assumption are well-founded. The restriction to these two assumptions is due to lack of space; other aspects of the approach also raise serious questions.

5. The Assumptions Scrutinized

Our discussion of the core assumption focuses on the informativeness. It is an undisputed fact that systematic errors in the models in question lead to non-trivial macroscopic errors of simulation, of the past and of the future. Seager et al. (2008) have noted their inability to reproduce the dust bowl of the 1930's even given the observed sea-surface temperatures. This is not a small inadequacy when one is focused on the resolution offered by UKCP09. Given these systematic errors, there are lead times at which the failure of the model to simulate realistic weather cause the climate of the model to differ from that of the planet (Smith and Stern 2011). Inasmuch as the models used are not close to the target, the informativeness assumption fails. The figure below shows model global mean temperatures over the last century of the 24 CMIP3 models.¹⁵ Note that while all models show warming between 1900 and 2000, their average temperatures vary tremendously. The magnitude of the error in the global mean in a hindcast of the last century casts significant doubt on the viability of the informativeness assumption on a 25 km forecast to the end of this century.

¹⁴ See (Sexton et al. 2012, 2521-27). To be precise, what is determined is the covariance matrix.

¹⁵ Thanks to Ana Lopez for producing the figure.



Even if one were to discard the above as undue pessimism and uphold the informativeness assumption, there is a further problem. An argument to uphold the informativeness assumption *now* must be based on the trustworthiness of the modeling assumptions, typically taking the form that if the model assumptions are close to the truth, then the model outputs must be close to the truth too. While this inference works in some specific applications, it is generally false for nonlinear models *even if their SME is extremely small* (Frigg et al. 2013). Since the relevant climate models are nonlinear, it follows that even if the model assumptions were close to the truth this would not automatically warrant trust in the model outputs. In fact, the outputs for relevant lead times fifty years from now could still be seriously misleading.

The discussion of the proxy assumption is complicated by the fact that the literature on the subject exhibits a certain degree of schizophrenia. On the one hand the method is illustrated and advertised as delivering trustworthy results; on the other hand disclaimers that effectively undermine the crucial assumptions are also included, sometimes parenthetically, obscurely or deep within technical discussions.¹⁶ The documentation gives with one hand and takes back with the other. We now review the activities of both hands and conclude that the hand that takes back voids the trustworthiness of the forecasts for quantitative decision support.

¹⁶ An example is (Murphy et al. 2010, 63-69).

The first reason cited in support of the proxy assumption is that multi model averages give a better representation of climate than any individual model: ‘Indeed, the multimodel ensemble mean has been shown to be a more skillful representation of the present-day climate than any individual member’ (Sexton et al. 2012, 2526). Yet it is also acknowledged that ‘systematic errors to all current climate models persist’ (ibid.) so even if one were to accept that such a multi-model mean were *more* skillful at representing the present day than individual models, is ‘more skillful’ close to being ‘skillful’? Unfortunately there appears to be no evidence that ‘more skillful’ can be equated with ‘skillful’ for many variables of importance for future climate change.

The second reason mentioned in support of the proxy principle is that ‘the structural errors in different models can be taken to be independent’ (Murphy et al. 2010, 66) and that therefore the ensemble samples uncertainty well. However, immediately after we are warned that

‘Whilst there is evidence for a degree of independence [...], there is also evidence that some errors are common to all models [...], due to shared limitations such as insufficient resolution or the widespread adoption of an imperfect parameterisation scheme. From this perspective, our estimates of discrepancy can be viewed as a likely lower bound to the true level of uncertainty associated with structural model errors.’ (Murphy et al. 2010, 66)

And then the conclusion is drawn that: ‘The main (and inevitable) limitation, however, is that it [the proxy assumption] does not account for the potential impacts of errors common to all climate models used in the prediction (Sexton et al. 2012, 2516).

One cannot have ones cake and eat it too. If there are common errors the proxy assumption fails. Indeed such common errors have been widely acknowledged (see, for instance, Parker 2011). Furthermore, the mathematical space of all possible climate models (if there is some such thing) is huge, and there is no reason to believe that the 12 models we *de facto* work with provide a representative sample.

For these reasons, the assumption that the use of an MME will accurately quantify the distance to our true target is unjustified. It produces a distribution that is more consistent with the diversity of current models but need not reflect the uncertainty in our future. It is important to note that the fear is not so much that the width of the uncertainty distribution is too narrow, but rather that the distribution is simply in the wrong place; that the mean of the distribution will shift significantly if the model simulations become realistic.

Echoing Murphy *et al.*, we note that ‘[i]t is important to stress that our approach to the specification of discrepancy can only be expected to capture a subset of possible structural modelling errors and should be regarded as a lower bound’ (Murphy et al. 2007, 2011). A lower bound need neither yield trustworthy forecasts nor provide a suitable basis for quantitative decision support.

6. Conclusion

We have argued that there is little evidence for interpreting UKCP09's predictions as trustworthy forecasts for quantitative decision support. Questioning the evidence, however, does not amount to proving it wrong. Our point is that the premises of the argument do not warrant trust in the results, and for decision support in the face of climate change this is the crucial aspect.

To be fair to the scientists who worked very hard to make UKCP09 the best it could be, several points should be noted. First the deliverables of the project were defined before any viable approach to meet them was available in the peer-reviewed literature. Second, the *United Kingdom Climate Impacts Program*, which is much broader than UKCP09, faced the dilemma of motivating users to engage with the real challenges and risks posed by climate change in the face of deep uncertainty: the challenge of keeping users interested when the information they most desire lies beyond the reach of today's science. And lastly, pointers to the fact that a naïve interpretation of UKCP09 probability distributions is untenable can indeed be found within the UKCP09 material.

That said, the aim of UKCP09 was to provide trustworthy forecasts now, and this, we have argued, they fail to do. Kelly's 1979 plea holds today: climate prediction experiments remain essential, and they must 'be conducted with scientific rigour and presented with an honest assessment of the uncertainties involved' (Kelly 1979, 182)

References

- Allen, Myles R. and David A. Stainforth. 2002. "Towards Objective Probabilistic Climate Forecasting." *Nature* 419:228-228.
- Frigg, Roman, Seamus Bradley, Hailiang Du and Leonard A. Smith. 2013. "The Adventures of Laplace's Demon and His Apprentices." *Philosophy of Science* forthcoming.
- Good, Irving J. 2009. *Good Thinking*. Mineola/NY: Dover Publications.
- Jenkins, Geoff, James Murphy, David Sexton, Jason Lowe and Phil Jones. 2009. *UK Climate Projections*. Briefing Report. DEFRA.
- Kelly, P. Mick. 1979. "Towards the prediction of climate." *Endeavour* 3 176-182.
- Murphy, James, David Sexton, Geoff Jenkins, Penny Boorman, Ben Booth, Kate Brown, Robin Clark, Mat Collins, Glen Harris and Lizzie Kendon. 2010. *UK Climate Projections Science Report: Climate Change Projections*. Version 3, Updated December 2010. Met Office Hadley Centre.
- Murphy, James M., Ben B. Booth, Mat Collins, Glen R. Harris, David M. H. Sexton and Mark J. Webb. 2007. "A Methodology for Probabilistic Predictions of Regional Climate Change for Perturbed Physics Ensembles." *Philosophical Transactions of the Royal Society A* 365:1993-2028.
- Oreskes, Naomi. 2007. "The Scientific Consensus on Climate Change: How Do We Know We're Not Wrong?" In *Climate Change: What It Means for Us, Our Children, and Our Grandchildren*, ed. Joseph F. C. DiMento and Pamela Doughman, 65-99. Boston: MIT Press.
- Oreskes, Naomi, David A. Stainforth and Leonard A. Smith. 2010. "Adaptation to Global Warming: Do Climate Models Tell Us What We Need to Know?" *Philosophy of Science* 77:1012-1028.

- Parker, Wendy. 2011. "When Climate Models Agree: The Significance of Robust Model Predictions." *Philosophy of Science* 78:579-600.
- Seager, Richard, Yochanan Kushnir, Mingfang F. Ting, Mark Cane, Naomi Naik and Jennifer Miller. 2008. "Would Advance Knowledge of 1930s SSTs Have Allowed Prediction of the Dust Bowl Drought?" *Journal of Climate* 21:3261-3281.
- Sexton, David M. H. and James 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.
- Sexton, David M. H., James M. Murphy, Mat Collins and Mark J. Webb. 2012. "Multivariate Probabilistic Projections Using Imperfect Climate Models Part I: Outline of Methodology." *Climate Dynamics* 38:2513-2542.
- Smith, Leonard A. 2006. "Predictability Past Predictability Present." In *Predictability of Weather and Climate*, ed. Tim Palmer and Renate Hagedorn, 217-250. Cambridge: Cambridge University Press.
- Smith, Leonard A. and Nicholas Stern. 2011. "Uncertainty in Science and its Role in Climate Policy." *Philosophical Transactions of the Royal Society A* 369:1-24
- Solomon, Susan, Dahe Qin and Martin Manning. 2007. "Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change." Cambridge University Press
- Stainforth, David A., Myles R. Allen, Edward R. Tredger and Leonard A. Smith. 2007. "Confidence, Uncertainty and Decision-Support Relevance in Climate Predictions." *Philosophical Transaction of the Royal Society A* 365:2145-2161.
- Tang, Samuel and Suraje Dessai. 2012. "Usable Science? The UK Climate Projections 2009 and Decision Support for Adaptation Planning." forthcoming in *Weather, Climate, and Society*