Climate Models and the Irrelevance of Chaos

Corey Dethier

(Draft – please do not quote without permission)

Abstract

Philosophy of climate science has witnessed substantial recent debate over the existence of a dynamical or "structural" analogue of chaos, which is alleged to spell trouble for certain uses of climate models. In this paper, I argue that the debate over the analogy can and should be separated from its alleged epistemic implications: chaos-like behavior is neither necessary nor sufficient for small dynamical misrepresentations to generate erroneous results. I identify the relevant kind of kind of sensitivity with a kind of safety failure and argue that the resulting set of issues has different stakes than the extant debate would indicate.

1 Introduction

To make predictions about the future of a system, we need to know two things: the initial conditions, or, present state of the system; and the dynamics of the system, or, how it evolves with time. Chaotic systems present particular difficulties because small differences in initial conditions amplify into large differences in the end state of the system. Is there an analogous dynamical property of systems? Intuitively, it seems like there might be: small differences in the dynamics amplify into large differences in the end state of the system.

In a series of papers, a group of philosophers and scientists have argued this analogous dynamical property and that it spells epistemic trouble for certain hypotheses in climate science. Specifically, they argue that because most climate models heavily idealize the dynamics of the climate, the possibility that such models exhibit a dynamical phenomenon analogous to chaos should cause us to have low confidence in the accuracy of quantitative predictions that rest on them. An opposed group of critics have argued that the analogy breaks down and that the epistemic conclusions don't follow; the possibility of small dynamical errors doesn't undermine the general warrant for quantitative climate predictions.

While the mathematical question concerning the alleged analogy with chaos is interesting on its own terms, the focus on it is misleading from a purely epistemic perspective: a tight analogy to chaos is neither necessary nor sufficient for the kinds of epistemic error that motivate the debate. Chaotic behavior involves growth in physical distance with time; such growth is relevant to the accuracy of a given prediction only when (a) the starting distances are small relative to the desired level of precision, (b) the later distances large from the same perspective, and (c) the time frame covered by the prediction the same as that on which the system is chaotic. The type of *epistemic* sensitivity relevant to error is better captured by the failure of a kind of safety condition. And while it is true that there is good reason to worry about safety failures in climate science, the arguments in question are better seen as explaining known safety failures than as providing evidence for the existence of unknown ones.

A more detailed outline of my arguments is as follows. In sections 2 and 3, I briefly characterize the debate over dynamical analogues of chaos and argue that it has misfired insofar as it presupposes a connection between chaos-like behavior and (the probability of) error. There's no real connection here because chaos involves a type of interest-independent sensitivity, whereas the probability of error is inherently dependent on our interests. In section 4, I provide an alternative notion of sensitivity that is appropriately interest-dependent. This notion is best expressed in terms of the failure of a kind of safety principle—essentially, the safety principle fails when a hypothesis is only justified given an assumption that's uncertain or risky. Finally, in section 5, I offer a reinterpretation of the original arguments: what they motivate is low confidence in our ability to substantially increase the precision of model reports.

2   The debate over dynamical analogues of chaos

The debate over dynamical analogues of chaos has largely focused on a particular minimal condition on chaotic behavior, what's known as "sensitive dependence on initial conditions" (SDIC). Roughly speaking, a system exhibits SDIC if "even arbitrarily close initial conditions will follow very different trajectories" through the state space that characterizes the system (Frigg, Bradley, et al. 2014, 34). Trivially, SDIC implies that a model that slightly misrepresents the initial conditions of the system will misrepresent (some) later states of the system to a much larger degree. In a series of recent

papers, a group of philosophers and scientists associated with the London School of Economics—and thus termed the "LSE group" by their critics—have argued that "structural model error" (SME) presents epistemological problems similar to those presented by SDIC (Frigg, Bradley, et al. 2014; Frigg, L. A. Smith, and Stainforth 2013, 2015).[1]

To be precise, it's important to recognize that SME is *not* supposed to be directly analogous to SDIC. On the contrary, SME occurs whenever a model misrepresents the dynamics of the target system (Frigg, Bradley, et al. 2014, 32). The analogy, according to the LSE group, is between the behavior of non-linear systems relative to SME and to the behavior of SDIC-exhibiting systems relative to misrepresentation of initial conditions. That is: the existence of small amounts of SME will lead a model to dramatically misrepresent some later states of a non-linear system in much the same way that a small misrepresentation of initial conditions will lead a model to dramatically misrepresent some later states of a chaotic system. In effect, non-linearity induces a *sensitive dependence on dynamical equations*. The LSE group then draws the further conclusion that since climate models are (a) heavily idealized and (b) non-linear, we should have low confidence in "decision-relevant" quantitative climate predictions—though they acknowledge that just how sensitive the models are to SME is a question the requires further investigation (Frigg, Bradley, et al. 2014, 48).

The arguments of the LSE group have spawned a series of responses (Goodwin and Winsberg 2016; Nabergall, Navas, and Winsberg 2019; Winsberg 2018; Winsberg and Goodwin 2016) from a group of philosophers and scientists associated with the University of South Florida (who I'll term the "USF group" for parallel's sake). The main contention of the USF group is that the analogy between systems that exhibit SDIC and what I above termed sensitive dependence on dynamical equations cannot be made precise for two reasons.[2] First, the space of dynamical equations is topological but not metrical, meaning that there's no general way to say what it means to be an "arbitrarily close" equation (Winsberg and Goodwin 2016, 14). Second, and for similar reasons, the mathematically well-defined property closest to sensitive dependence on dynamical equations shows only that small dynamical misrepresentations *can* amplify into large errors in the representation

---

[1] Also worth noting are Mayo-Wilson (2015), L. A. Smith (2007), and Thompson (2013), who explore the possibility of a dynamical analogue of chaos without drawing the same epistemic conclusions vis-a-vis climate science.

[2] They have also advanced a number of other objections, most notably that the central motivating example of the LSE group involves methods that are distinct from those used in much of climate modeling (see, e.g., Winsberg and Goodwin 2016, 12,15). The point is well taken—the example employed by the LSE group is not sufficient to establish general conclusions—but, as we'll see, the relevant epistemic worries have nothing to do with the features specific to that example.

of later states of the target system not that it *will* (Nabergall, Navas, and Winsberg 2019, 11-12). They conclude that there's no general threat to quantitative climate predictions stemming from "infinitesimally small" dynamical misrepresentations (Nabergall, Navas, and Winsberg 2019, 21). Acknowledging, of course, that dynamical misrepresentations do pose epistemic problems in some cases, they argue that the failure of the analogy means that we should resist the LSE group's general conclusions; the epistemic implications for decision-relevant quantitative climate predictions must be evaluated individually (Nabergall, Navas, and Winsberg 2019, 20; Winsberg and Goodwin 2016, 16).

3   Chaos and error

Both the LSE and USF groups appear to consider the (alleged) epistemic problem to be one of *error*.[3] In their central thought experiment, for instance, the LSE group present the problem associated with SME as one of erroneous probabilistic predictions: the agent facing SME-related problems "regards events that do not happen as very likely, while he regards what actually happens as very unlikely" (Frigg, Bradley, et al. 2014, 39). Similarly, in their discussion of the primary motivating case study—a project involving generating "decision-relevant" probabilistic predictions about the future climate in Great Britain—they worry that:

> Trying to predict the true climate with structurally wrong models is like trying to predict the trajectory of Mercury with Newtonian models. These models will invariably make misleading (and likely maladaptive) projections beyond some lead time, and these errors cannot be removed by adding a linear discrepancy term derived [solely] from other Newtonian models. (Frigg, L. A. Smith, and Stainforth 2015, 3997)

And the USF group is no different. They echo the language of the LSE group in their own discussion of the motivating example (Goodwin and Winsberg 2016, 1125), and more recently, they've stated that "only strong versions [of chaos] are usually taken to have strong epistemological consequences, since they are *likely* to produce error" (Nabergall, Navas, and Winsberg 2019, 7, note 13).

---

[3]That said, in their more careful moments, at least, the LSE group can be read as primarily concerned with particular sorts of tradeoffs between precision and certainty (see, e.g., Frigg, Bradley, et al. 2014, 50). In light of the arguments presented in this section, I think that the most charitable interpretation is likely to emphasize this concern over the concern for error. See section 5 for more details.

To be sure, in the context of the examples employed by the LSE group, introducing chaos *while holding the predictions of the agent fixed at a given level of precision* does increase the probability of error. Introducing sensitive dependence to dynamical equations has the same effect. The USF group are also right that the effect is only significant for stronger versions of chaos and (we could add) only significant if the time frames line up in the right way. But these facts don't imply that there's a connection between chaos (or chaos-like behavior) and error *in general*. If there were such a connection, the consequence would be that we can't make accurate and/or precise predictions about chaotic systems—or (more weakly) that the behavior of such systems are generally harder to predict than that of non-chaotic systems. But this simply isn't the case.

The LSE group's own analogy illustrates the point nicely. The errors in Newtonian predictions of the trajectory of Mercury are on the order of mere arcseconds per century—that is, a prediction of where Mercury will appear in the sky a hundred years out will exhibit an error roughly $1/40^{\text{th}}$ the apparent width of the moon. It's hard to argue that such small errors are genuinely maladaptive. And the solar system as a whole is chaotic: eventually—that is, approximately five million years from now—small differences between present conditions will have grown exponentially larger.[4] And yet we're nevertheless able to make astoundingly precise (and "decision-relevant"!) predictions about the locations of various stellar bodies, for the simple reason that the five-million year timescale is totally irrelevant for predictions in the here-and-now.[5] While there are cases in which chaotic behavior creates genuine problems for predictive accuracy, in other words, it's simply illegitimate to draw inferences from either chaotic behavior or the lack thereof to the the existence of such problems without further information.

The same conclusion is suggested by close attention to more precise definitions of chaotic behavior. Consider the common definition of SDIC in terms of Lyapunov exponents. Suppose that there is a system characterized by a state space X and dynamical mapping $f : X \to X$ such that $x_t = f(x_{t-1})$. This system exhibits SDIC if and only if for all y "arbitrarily" close to x_0,

$$d(x_t, y_t) > e^{\lambda t} d(x_0, y)$$

---

[4]That is, the "Lyapunov time" of the solar system—the time it takes for distances to grow by a factor of e—is approximately five million years.

[5]Well, not totally irrelevant, because the same physical properties that engender the chaotic behavior of the solar system generate attractors that can affect satellite trajectories; see Wilhelm (2019).

where λ, the "Lyapunov exponent," is positive. Essentially: trajectories that are currently "nearby" will grow exponentially farther apart. If we interpret $d(x_0, y)$ as the present uncertainty, then SDIC entails that uncertainty will grow exponentially with time. It's a significant step from uncertainty growing exponentially with time to either a high probability of error at a given time or some sort of guarantee of inaccuracy. To make this step there needs to be a tight relationship between the relevant timescales and (as we saw above) there's no guarantee that the timescale relevant to our predictions will be the same one that's relevant to chaos. Similar comments apply to other technical definitions of chaos. Werndl (2009), for example, shows that a system is "topologically-mixing" if and only if

$$\lim_{n \to \infty} Pr(x_0|x_{-n}) - Pr(x_0) = 0$$

which, in English, says that the probablistic relevance of past events to future events eventually approaches zero. Chaotic systems "lose" information over time, but the mere fact that information is guaranteed to be lost *eventually* doesn't implicate our ability to make precise or accurate predictions now.

The explanation for the disconnect between chaotic behavior on the one hand and predictive inaccuracy on the other is that SDIC defines a notion of physical sensitivity that is *independent of human interests*. Until we specify a timeframe and desired level of precision for a hypothesis, we cannot know what implications SDIC will have for said hypothesis. Since our interests don't map onto physical distances in any consistent way—a few centimeters of error is a disaster in a surgical setting but incredible in astrophysics—SDIC doesn't have any general implications for either error or the probability of error. Similarly, we should expect that the failure of a system to exhibit SDIC—or an SDIC-like property—also has no general implications for error. The contested claims about the analogy between SDIC and sensitive dependence on dynamical equations therefore has no clear or direct implications for the epistemology of climate modeling; like the solar system, climate models could exhibit exponential growth in the distances between alternative trajectories over time frames on the scale of millions of years. Or they could fail to exhibit any growth in distances between trajectories but the starting uncertainty could be too substantial to license "decision-relevant" predictions. Insofar as our concern is something like the probability of error in general,

chaos and chaos-like behavior simply aren't relevant.

4   Failures of safety

To determine whether the presence of dynamical misrepresentations renders "decision-relevant" quantitative climate predictions untrustworthy, we need a different, interest-relative, concept of sensitivity to small errors. My view is that the relevant concept is given by the failure of a kind of safety condition.

Speaking abstractly, when we're concerned with whether we should believe some hypothesis, one relevant desideratum is that the justification for the hypothesis should be *safe*: the degree of support for the hypothesis should be (nearly) the same given nearby alternative background assumptions, where a background assumption is "nearby" to the extent that it has a relatively high probability on the total evidence available.[6] So, for instance: if my evidence for the fact that it is freezing outside is the reading of my thermometer, then the hypothesis is safer in the situation where the thermometer reads -5°C than when it reads -1°C; the former allows for more leeway in the background assumptions concerning the accuracy of the thermometer. When the evidence for a hypothesis rests either fully or partially on a model, the hypothesis is going to more or less safe to the extent that sufficiently small changes to the assumptions of the model don't (substantially) affect the results or outputs of the model. The reasoning here is the same. If the hypothesis is only supported by the model given precise and risky assumptions, then there's a relatively high chance that these assumptions don't hold. By contrast, if the hypothesis is supported regardless of whether we use the specific assumptions in question or any one of a number of nearby assumptions, then the hypothesis is safe.

Intuitively, safety is going to be related to the probability of error at least under conditions in which there's some degree of uncertainty about the quality of the evidence. Since humans are not ideal reasoners, we're often in situations in which we don't know how likely some hypothesis is on our evidence. So, for instance, we might know that we should be confident in P if Q is true, but not

---

[6]This notion of safety is essentially the one found in Reed (2000) and Staley (2004), and is tightly connected to G. E. Smith (2002, 2014)'s discussion of "*quam proxime*" reasoning. Like other safety conditions, the best way to define this one precisely is in terms of possible worlds and a distance measure between them, though what we want is a graded measure that allows for higher and lower degrees of safety. I take it that how this all works intuitively is clear enough for present purposes.

either whether Q is true or what our confidence in P should be given relatively likely alternatives to Q.[7] Why might we be in this situation? One common and relevant reason is that our evidence relies on an idealized model; since the inner workings of models are often "opaque" (Humphreys 2004), we can't know *a priori* whether or not the idealizations in question merely serve to simplify the problem in a harmless manner or, by contrast, whether they substantially affect the output of the model.[8] In other words, we don't know whether P is *safe*, whether it would still be justified given small changes to the background assumptions. If it is safe, then the evidence is trustworthy and provides good reason to believe that P; if it isn't, then the total evidence does not provide reason to believe that P. If we accept P, therefore, safety and error will be inversely correlated: the safer the hypothesis, the higher its overall justification, and thus the lower the chance that it is has been accepted in error.

The foregoing is highly abstracted from the practices of science. Consider, therefore, the derivation of inverse-square gravity from Kepler's first law.[9] Suppose that Kepler's first law holds *exactly*, meaning that sun is at the focus of each planet's elliptical orbit and that the distance function between planet and sun is

$$d = A \frac{(1 - \epsilon^2)}{1 - \epsilon \cos \theta}$$

where A is the long arm of the ellipse and ϵ the eccentricity. In combination with some other information about the nature of ellipses, this equation entails that the acceleration of the planet is proportional to the inverse of square of the distance ($a \propto d^{-2}$). It's thus possible to derive the inverse-square law from Kepler's first law. In the context of the present discussion, however, this derivation faces two problems. First, there was little evidence available that Kepler's first law held *precisely* (and, in fact, it doesn't): the difference between an ellipse with the sun at a focus and

---

[7]Epistemologists term cases like these instances of "higher-order uncertainty." There's disagreement concerning whether we can *rationally* have higher-order uncertainty (see, e.g., Dorst 2019; Titlebaum 2015); I won't take a stance on that here. My concern is that as non-ideal agents we frequently are uncertain when we would rather not be.

[8]Isn't this a case in which we know Q to be false? No, though defending this point adequately here would take us too far afield. The central idea is that it's a mistake to read Q as a claim about the truth of the idealized model rather than as a claim about its (non-literal) accuracy (Frigg and Nguyen 2019), adequacy-for-purpose (Parker 2009), or reliability (Dethier 2019).

[9]I'm borrowing this example from G. E. Smith (2002). Smith's point, which is worth emphasizing, is that the safety failure present in this example provides the best explanation for why Newton himself didn't derive the inverse-square law in this manner, preferring instead the evidence provided by the apsides of the planets (Newton 1727/1999, 802), relative to which the hypothesis is *extremely* safe.

an ellipse with the sun at the center is virtually undetectable with 17th century tools.[10] The first problem with the derivation, then, is that we're uncertain whether the assumptions built into it hold precisely.

The second problem is that the derivation is extremely sensitive to small deviations from Kepler's first law. As just noted, at low eccentricity, there's very little difference between an ellipse with the sun at the focus and one with the sun at the center. The distance function for the latter is given by

$$d = A\sqrt{1 - \epsilon^2 \sin^2 \theta}$$

And this function, in combination the same assumptions about the nature of ellipses, entails that the acceleration of the planet is proportional to the distance *directly* ($a \propto d$). The derivation from Kepler's first law therefore provides extremely poor evidence for the conclusion in the sense that it relies on a particular assumption holding precisely when the best evidence available only indicates that the assumption holds approximately.

This case provides an exemplar of a safety failure in a number of respects. Recall: safety failures arise because the quality of the evidence varies dramatically with small changes in background assumptions. Here the changes to background assumptions are small not because the two equations are nearby in any mathematical sense but because the evidence makes both assumptions relatively likely. And the difference in the quality of the evidence is dramatic because of our particular choice of how to divide up the hypothesis space: what matters is that d and d^{-2} make for extremely different theories of gravitation. If our hypothesis was simply that there is *some* relationship between distance and acceleration, there would not be a safety failure to be found. It is also exemplary with regards to effects: the safety failure makes it likely *on our evidence* that if we accept the hypothesis, we're going to do so erroneously—which is just to say that when the hypothesis fails to be safe, the evidence doesn't give us much reason to believe it. (Though, of course, and as evidenced by this example, other evidence might; see note 9.)

The definition of safety given in this section provides a notion of sensitivity that is appropriately

---

[10]Kepler argued for his first law by showing that it held to a high approximation with respect to Mars (Miyake 2015); taking it to hold to a high approximation with respect to the other planets is a relatively risky inductive move—and one that Newton knew his own theory would show to be invalid (Newton 1727/1999, 817-18).

dependent on human interests. The class of objects that there is sensitivity to are defined or identified according to our epistemic abilities: in the modeling context, it's the class of assumptions that are empirically adequate by our standards. Similarly, whether or not there is sensitivity to the differences between these representations depends on our interests and concerns insofar as those affect how precise we want or need our hypotheses to be. There's a *rough* analogy to SDIC or chaos here in that a safety failure involves a "growth" in "distances" in an interest-dependent sense: the initial distance is small relative to our ability to distinguish between different scenarios and the latter one large relative to our desire for precision. But this is connection is not mathematically precise. In particular, safety failures are not analogous to SDIC in the ways that the USF group argues present problems for the LSE group.[11] There's no interest-independent distance measure to be placed on either the different starting characterizations of the system or the resulting equations. The different distance equations are similar just in the sense that they're both empirically adequate in the given situation; the different relationships between acceleration and distance are dissimilar in the sense that their broader fit in the theory is dramatically different. Further, we haven't shown that *any* nearby deviation from Kepler's laws (or even Kepler's first law) will lead to an arbitrarily different relation between acceleration and distance. All that we've shown that there is a particularly salient alternative that has this effect.

This section has provided an appropriate notion of sensitivity to employ in getting clearer about the debate over chaos. In the next section, I'll argue for a reinterpretation of the LSE group's arguments in terms of safety failures.

5   Reinterpreting the LSE group

I think that the arguments presented by the LSE group are important, but they don't show that the possibility of small dynamical errors should cause us to lower our confidence in various claims supported by climate models. Instead, they should be interpreted as offering an explanation of (empirically-ascertained) levels of model precision in terms of small dynamical errors—an explanation that, if true, has important implications for which projects in climate science are likely to be successful.

---

[11]There are other disanalogies as well, besides those at issue in the debate surveyed in §2. For instance, neither initial conditions nor time has any role in this case, though both are essential to the understanding of SDIC.

The main motivation behind this interpretation of the LSE group is that the arguments that they offer are neither necessary nor sufficient to establish that we should be less confident in the claims supported by climate models in general. They are not sufficient because they would need to show that climate scientists have generally been overconfident in modeling results—but climate scientists are well aware that climate models can be highly misleading, even in the aggregate (Knutti, Furrer, et al. 2010). They are not necessary because general considerations about safety failures provide much less powerful (and precise) evidence for caution about specific climate hypotheses than is provided by the empirical evaluation of climate models. Evaluation studies provide evidence not just about the degree of confidence licensed by a given model, but also about where the models excel, where they struggle, and what assumptions account for these struggles. Any general considerations about safety failures are likely to simply be swamped by the empirical evidence from this domain.

Of course, the LSE group is well aware of this empirical literature—as evidenced by their prior work drawing out the implications of it for decision-making (Oreskes, Stainforth, and L. A. Smith 2010; Stainforth et al. 2007). My suggestion is that we should read their arguments concerning chaos through the lens of this earlier work.[12] Specifically, we should view the combination of small dynamical errors and system complexity as providing an explanation for why climate models are only able to achieve certain levels of precision and accuracy. In giving this explanation, the LSE group is stressing that our inability to draw conclusions about local policy from climate models isn't a temporary defect of these models. On the contrary, hypotheses about how climate change is going to affect a town, region, or (small) country are simply too sensitive to small changes in modeling assumptions, and we're not likely to reach a point any time in the near future where we have the ability to determine which of these assumptions are true. That's just the nature of the system (a conclusion, I'll note, that is widely shared among climate scientists; see Knutti and Sedlàček 2013). In other words: hypotheses that we know are unsafe based on our empirical evaluations of the models are likely to remain unsafe—and thus, as the LSE group explicitly suggests, we need methods for determining how to make decisions under conditions where the quality of our evidence is uncertain in precisely this way.

If this is the correct interpretation *and* the arguments given in prior sections are correct, then the way that the LSE group has presented their arguments for this conclusion is misleading; the

---

[12]Here I'm following Mike Goldsby and Greg Lusk.

analogy to SDIC is largely irrelevant to whether or not the present and future levels of uncertainty about the dynamics are likely to undermine the evidence for future climate hypotheses. My experience is that they're doubly misleading to those unfamiliar with precise definitions of chaos—many of those presented with the arguments seem to automatically assume that chaotic behavior means that *anything* goes. Many of the USF group's criticisms (particularly those not discussed above) are aimed directly at this point: the presentation and rhetoric of the analogy to chaos, they contend, doesn't align with the more limited conclusions that the LSE group wants to draw (see Goodwin and Winsberg 2016; Winsberg and Goodwin 2016). As we've just seen, however, whatever the disconnect between the rhetoric and the arguments in the LSE group's papers, there isn't a genuine worry that their arguments might—even if successful—undermine much more than they intend: on this interpretation, the arguments simply don't motivate changing our confidence in any particular results of the models; they motivate "only" changing our confidence that we'll be able to get well-justified decision-relevant predictions out of the models any time soon.

To be clear, I am not arguing that the argument just sketched is correct. Nevertheless, the conclusion is interesting and the arguments itself has the advantages of fitting nicely with the prior work of the LSE group, not relying on mistakes concerning the relationship between chaos and error, and not—if successful—implicating far more of climate science than can be plausibly be justified on the evidence appealed to. We thus have good reason to interpret the LSE group in this manner, even if the argument ends up being unsound.

6   Conclusion

In this paper, I've argued that the details of dynamical analogues of chaos are largely irrelevant to the epistemological questions raised in the recent debate over them. Mathematically interesting as the alleged analogy may be, a tight analogy to chaos is neither necessary nor sufficient for the kinds of epistemic error that motivate the debate. The type of *epistemic* sensitivity relevant to error is better captured by failure of a kind of safety condition: what's worrying about dynamical misrepresentations is that they undermine the evidence provided by the model. Once the irrelevance of chaos is recognized, it becomes clear that the upshot of the debate is not whether models are likely to be erroneous but an explanation for why models are not more precise than they in fact

are.

References

Dethier, Corey (2019). How to Do Things with Theory: The Instrumental Role of Auxiliary Hypotheses in Testing. *Erkenntnis* (online first).

Dorst, Kevin (2019). Higher-Order Uncertainty. In: *Higher-Order Evidence: New Essays.* Ed. by Mattias Skipper and Asbjørn Steglich-Petersen. Oxford: Oxford University Press: 35–61.

Frigg, Roman, Seamus Bradley, et al. (2014). Laplace's Demon and the Adventures of His Apprentices. *The Journal of Philosophy* 81.1: 31–59.

Frigg, Roman and James Nguyen (2019). Mirrors Without Warnings. *Synthese* (online first).

Frigg, Roman, Leonard A. Smith, and David A. Stainforth (2013). The Myopia of Imperfect Climate Models: The Case of UKCP09. *Philosophy of Science* 80.5: 886–97.

— (2015). An Assessment of the Foundational Assumptions in High-Resolution Climate Projections: The Case of UKCP09. *Synthese* 192.12: 3919–4008.

Goodwin, William Marc and Eric Winsberg (2016). Missing the Forest for the Fish: How Much Does the 'Hawkmoth Effect' Threaten the Viability of Climate Projections? *Philosophy of Science* 83.5: 1122–32.

Humphreys, Paul (2004). *Extending Ourselves: Computational Science, Empiricism, and Scientific Method.* Oxford: Oxford University Press.

Knutti, Reto, Reinhard Furrer, et al. (2010). Challenges in Combining Projections from Multiple Climate Models. *Journal of Climate* 25.10: 2739–58.

Knutti, Reto and Jan Sedlàček (2013). Robustness and Uncertainties in the New CMIP5 Climate Model Projections. *Nature Climate Change* 3: 369–73.

Mayo-Wilson, Conor (2015). Structural Chaos. *Philosophy of Science* 82.5: 1236–47.

Miyake, Teru (2015). Underdetermination and Decomposition in Kepler's *Astronomia Nova. Studies in the History and Philosophy of Science Part A* 50: 20–27.

Nabergall, Lukas, Alejandro Navas, and Eric Winsberg (2019). An Antidote for Hawkmoths: On the Prevalence of Structural Chaos in Non-linear Modeling. *European Journal for Philosophy of Science* 9.21: 1–28.

Newton, Isaac (1727/1999). *Mathematical Principles of Natural Philosophy*. Trans. by I. Bernard Cohen and Anne Whitman. 3rd edition. Berkeley: University of California 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* 77.5: 1012–28.

Parker, Wendy S. (2009). Confirmation and Adequacy-for-Purpose in Climate Modelling. *Aristotelian Society Supplementary Volume* 83.1: 233–49.

Reed, Baron (2000). Accidental Truth and Accidental Justification. *The Philosophical Quarterly* 50.198: 57–67.

Smith, George E. (2002). From the Phenomenon of the Ellipse to an Inverse-Square Force: Why Not? In: *Reading Natural Philosophy: Essays in the History and Philosophy of Science and Mathematics*. Ed. by David Malament. La Salle: Open Court: 31–70.

— (2014). Closing the Loop: Testing Newtonian Gravity, Then and Now. In: *Newton and Empiricism*. Ed. by Zvi Beiner and Eric Schliesser. Oxford: Oxford University Press: 262–351.

Smith, Leonard A. (2007). *Chaos: A Very Short Introduction*. Oxford: Oxford University Press.

Stainforth, David A. et al. (2007). Confidence, Uncertainty and Decision-Support Relevance in Climate Predictions. *Philosophical Transactions of the Royal Society Series A* 365.1857: 2145–61.

Staley, Kent W. (2004). Robust Evidence and Secure Evidence Claims. *Philosophy of Science* 71.4: 467–88.

Thompson, Erica (2013). Modelling North Atlantic Storms in a Changing Climate. PhD dissertation. Imperial College, London.

Titlebaum, Michael (2015). Rationality's Fixed Point (or: In Defense of Right Reason). *Oxford Studies in Epistemology* 5: 253–294.

Werndl, Charlotte (2009). What Are the New Implications of Chaos for Unpredictability? *The British Journal for the Philosophy of Science* 60.1: 195–20.

Wilhelm, Isaac (2019). Celestial Chaos: The New Logics of Theory-Testing in Orbital Dynamics. *Studies in History and Philosophy of Science Part B* 65: 97–102.

Winsberg, Eric (2018). *Philosophy and Climate Science*. Cambridge: Cambridge University Press.

Winsberg, Eric and William Marc Goodwin (2016). The Adventures of Climate Science in the Sweet Land of Idle Arguments. *Studies in History and Philosophy of Science Part B* 54: 9–17.