

Climate Models and the Irrelevance of Chaos

Corey Dethier*†

Philosophy of science has witnessed substantial recent debate over the existence of a structural analogue of chaos, which is alleged to spell trouble for certain uses of climate models. 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. The kind of sensitivity that matters in epistemology is one that induces unsafe beliefs, and the existence of a structural analogue to chaos is better seen as an explanation for known safety failures than as providing evidence for unknown ones.

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 exists 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 some of the quantitative

*To contact the author, please write to: Leibniz Universität Hannover; e-mail: corey.dethier@gmail.com.

†I would like to thank Roman Frigg for valuable comments on an earlier version of this article. The article also benefited greatly from conversations with Anjan Chakravartty and Wendy Parker.

Philosophy of Science, 88 (December 2021) pp. 997–1007. 0031-8248/2021/8805-0021\$10.00
Copyright 2021 by the Philosophy of Science Association. All rights reserved.

predictions that rest on them. An opposed group of critics have argued that the analogy breaks down and that the epistemic conclusions do not follow; the possibility of small dynamical errors does not 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 are large from the same perspective, and (c) the time frame covered by the prediction is 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 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.

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. In section 4, I provide an alternative notion of sensitivity that is best expressed in terms of the failure of a safety principle 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 is 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 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, Smith, and Stainforth 2013, 2015; Frigg et al. 2014).

It is 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 et al. 2014, 32). The analogy, if

there is one, is between the behavior of complex systems relative to SME and 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 complex 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, complex systems exhibit *sensitive dependence on dynamical equations*. The LSE group then draws the further conclusion that this sensitivity should lead us to have low confidence in “decision-relevant” probabilistic climate predictions in at least some cases—although they acknowledge that just how sensitive the models are to SME is a question the requires further investigation (48).¹

The arguments of the LSE group have spawned a series of responses (Goodwin and Winsberg 2016; Winsberg and Goodwin 2016; Winsberg 2018; Nabergall, Navas, and Winsberg 2019) from a group of philosophers and scientists associated with the University of South Florida (who I will term the “USF group”). 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. First, the space of dynamical equations is topological but not metrical, meaning that there is 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 of later states of the target system not that they *will* (Nabergall et al. 2019, 11–12). The USF group concludes that there is no general threat to quantitative climate predictions stemming from “infinitesimally small” dynamical misrepresentations (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 (Winsberg and Goodwin 2016, 16; Nabergall et al. 2019, 20).

3. Chaos and Error. Both the LSE and USF groups appear to consider the (alleged) epistemic problem to be one of *error*.² In their central thought experiment, for instance, the LSE group presents the problem associated with

1. This presentation of the LSE group undersells the detail found in the actual discussion of climate models in Frigg et al. (2013, 2015).

2. That said, in their more careful moments, the LSE group can be read as primarily concerned with the trade-off between precision and certainty (see, e.g., Frigg et al. 2014, 50).

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 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 et al. 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 stated that “only strong versions [of chaos] are usually taken to have strong epistemological consequences, since they are *likely* to produce error” (Nabergall et al. 2019, 7 n. 13).

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 is 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 do not imply that there is a connection between chaos (or chaos-like behavior) and error in general. If there were such a connection, the consequence would be that we cannot make accurate or precise predictions about chaotic systems—or (more weakly) that the behavior of such systems is generally harder to predict than that of nonchaotic systems. But this simply is not 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 arc seconds per century—that is, a prediction of where Mercury will appear in the sky a hundred years out will exhibit an error roughly one-fortieth of the apparent width of the moon. It is hard to argue that such small errors are genuinely maladaptive. And the solar system as a whole is chaotic: eventually—that is, approximately 5 million years from now—small differences between present conditions will have grown exponentially larger. And yet we are nevertheless able to make astoundingly precise (and decision-relevant) predictions about the locations of various stellar bodies, for the simple reason that the 5 million year timescale is totally irrelevant for predictions in the here and now.³ While

3. 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).

there are cases in which chaotic behavior creates genuine problems for predictive accuracy, in other words, it is simply illegitimate to draw inferences from either chaotic behavior or the lack thereof to 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 $f : X \rightarrow X$ such that $x_t = f(x_{t-1})$. This system exhibits SDIC if and only if there is some y “arbitrarily” close to x_0 ,

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

where λ , the “Lyapunov exponent,” is positive. Essentially, trajectories that are currently “nearby” will grow exponentially farther apart. If $d(x_0, y)$ represents the present uncertainty, then SDIC entails that uncertainty will grow exponentially with time. It is 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 is no guarantee that the timescale relevant to our predictions will be the same one that is 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

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

Essentially, the probabilistic 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 does not 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 do not 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 does not 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 are not 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 are 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.⁴ 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 in which 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 more or less safe to the extent that sufficiently small changes to the assumptions of the model do not (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 is a relatively high chance that these assumptions do not 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 is some degree of uncertainty about the quality of the evidence. Since humans are not ideal reasoners, we are often in situations in which we do not 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 either whether Q is true or what our confidence in P should be given relatively likely alternatives to Q . 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

4. This notion of safety is essentially the one found in Reed (2000) and Staley (2004) and is tightly connected to Smith’s (2002) discussion of “*quam proxime*” reasoning. I take it that a more precise definition is unnecessary for present purposes.

cannot know a priori whether 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.⁵ In other words, we do not 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 is not, 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.⁶ 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

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

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 the square of the distance ($a \propto d^{-2}$). It is 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 does not): the difference between an ellipse with the sun at a focus and an ellipse with the sun at the center is virtually undetectable with seventeenth-century tools. The first problem with the derivation, then, is that we are 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 is 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

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

5. Is this not a case in which we know Q to be false? No, in this context, it is a mistake to read Q as a claim about the truth of the idealized model rather than as a claim about its "adequacy for purpose" (Parker 2020).

6. I am borrowing this example from Smith (2002). Smith's point is that the safety failure found in this example provides the best explanation for why Newton himself did not derive the inverse-square law in this manner, preferring instead the evidence provided by the apses of the planets, relative to which the hypothesis is extremely safe.

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 that 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 regard to effects: the safety failure makes it likely on our evidence that if we accept the hypothesis, we are going to do so erroneously—which is just to say that when the hypothesis fails to be safe, the evidence does not give us much reason to believe it. (Although, as evidenced by this example, other evidence might; see n. 6.)

The definition of safety given in this section provides a notion of sensitivity that is appropriately dependent on human interests. The class of objects that there is sensitivity to is defined or identified according to our epistemic abilities: in the modeling context, it is the class of assumptions that are empirically adequate by our standards. Similarly, whether 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 is a rough analogy to SDIC 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 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.⁷ There is 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 are both empirically adequate in the given situation. The different relationships between acceleration and distance are dissimilar in the sense that

7. There are other disanalogies as well. For instance, neither initial conditions nor time has any role in this case, although both are essential to the understanding of SDIC.

their broader fit in the theory is dramatically different. Further, we have not 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 we have shown is 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 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 do not 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.

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 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 (Stainforth et al. 2007; Oreskes, Stainforth, and Smith 2010). My suggestion is that we should read their arguments concerning chaos through the lens of this earlier work.⁸ 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

8. Here I am following a suggestion made by Greg Lusk and Mike Goldsby in their talk—titled “The Decision-Relevancy of Climate Model Results: Idle Arguments or Idle Dreams?”—at Models and Simulations 8.

and accuracy. In giving this explanation, the LSE group is stressing that our inability to draw conclusions about local policy from climate models is not 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 are not likely to reach a point any time in the near future when we have the ability to determine which of these assumptions is true. That is just the nature of the system—a conclusion, I will note, that is widely shared among climate scientists (see, e.g., Knutti and Sedláček 2013). In other words, hypotheses that we know are unsafe on the basis of 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 in which 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 analogy to SDIC is largely irrelevant to whether the present and future levels of uncertainty about the dynamics are likely to undermine the evidence for future climate hypotheses. Insofar as the LSE group puts forward this analogy, therefore, their arguments are misleading—a point the USF group has stressed, arguing that the presentation and rhetoric of the arguments does not always align with the more limited conclusions that the LSE group wants to draw (see Goodwin and Winsberg 2016; Winsberg and Goodwin 2016). As we have just seen, however, whatever the disconnect between the rhetoric and the arguments in the LSE group's papers, there is not a genuine worry that their arguments might—even if successful—undermine much more than they intend. On this interpretation, the arguments simply do not motivate changing our confidence in any particular results of the models; they motivate “only” changing our confidence that we will be able to get well-justified decision-relevant predictions out of the models any time soon.

To be clear, I am not stating that the argument just sketched is correct. Nevertheless, the conclusion is interesting and the argument 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 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 article, I 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 the failure of a kind of safety condition: what is 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

- Frigg, Roman, Seamus Bradley, Hailiang Du, and Leonard A. Smith. 2014. "Laplace's Demon and the Adventures of His Apprentices." *Journal of Philosophy* 81 (1): 31–59.
- 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 M., 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, Claudia Tebaldi, Jan Cermak, and Gerald A. Meehl. 2010. "Challenges in Combining Projections from Multiple Climate Models." *Journal of Climate* 25 (10): 2739–58.
- Knutti, Reto, and Jan Sedláček. 2013. "Robustness and Uncertainties in the New CMIP5 Climate Model Projections." *Nature Climate Change* 3:369–73.
- 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.
- 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. 2020. "Model Evaluation: An Adequacy-for-Purpose View." *Philosophy of Science* 87 (3): 457–77.
- Reed, Baron. 2000. "Accidental Truth and Accidental Justification." *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*, ed. David Malament, 31–70. La Salle, IL: Open Court.
- Stainforth, David A., Myles R. Allen, Edward R. Tredger, and Leonard A. Smith. 2007. "Confidence, Uncertainty and Decision-Support Relevance in Climate Predictions." *Philosophical Transactions of the Royal Society A* 365 (1857): 2145–61.
- Staley, Kent W. 2004. "Robust Evidence and Secure Evidence Claims." *Philosophy of Science* 71 (4): 467–88.
- Werndl, Charlotte. 2009. "What Are the New Implications of Chaos for Unpredictability?" *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 B* 65:97–102.
- Winsberg, Eric. 2018. *Philosophy and Climate Science*. Cambridge: Cambridge University Press.
- Winsberg, Eric, and William M. Goodwin. 2016. "The Adventures of Climate Science in the Sweet Land of Idle Arguments." *Studies in History and Philosophy of Science B* 54:9–17.