Studies in History and Philosophy of Science 49 (2015) 58–68

Contents lists available at ScienceDirect

Studies in History and Philosophy of Science

journal homepage: www.elsevier.com/locate/shpsa

Model robustness as a confirmatory virtue: The case of climate science

Elisabeth A. Lloyd

History and Philosophy of Science Department, Goodbody Hall 130, Indiana University, Bloomington, IN, 47405, USA

A R T I C L E I N F O

Article history:
Received 23 April 2014
Received in revised form
30 September 2014
Available online 13 January 2015

Keywords:
Robustness;
Confirmation;
Scientific inference;
Climate models;
Variety-of-evidence;
Causal attribution

A B S T R A C T

I propose a distinct type of robustness, which I suggest can support a confirmatory role in scientific reasoning, contrary to the usual philosophical claims. In *model robustness*, repeated production of the empirically successful model prediction or retrodiction against a background of independently-supported and varying model constructions, within a group of models containing a shared causal factor, may suggest how confident we can be in the causal factor and predictions/retrodictions, especially once supported by a variety of evidence framework. I present climate models of greenhouse gas global warming of the 20th Century as an example, and emphasize climate scientists' discussions of robust models and causal aspects. The account is intended as applicable to a broad array of sciences that use complex modeling techniques.

1. Introduction

Philosophers Zach Pirtle et al. documented the fact that climate scientists tend to be attracted to robustness and to think it boosts confirmation of models.[2] In a recent qualitative survey of the contents of six leading climate journals since 1990, they found 118 articles in which the authors relied on a rough concept of agreement between climate model predictions/retrodictions to inspire confidence in their results (Pirtle, Meyer, & Hamilton, 2010, p. 3). I, too, defended robustness as an empirical strength of the huge general circulation models,[1] GCMs, in earlier discussions of the variety of evidence supporting those models (Lloyd, 2009, 2010, 2012).

But philosophers of science usually do *not* consider robustness to be an empirical or confirmatory virtue, that is, a virtue that indicates that a model or models are more likely to be used to represent accurate or true claims about the observable world (e.g., Calcott, 2011; Houkes & Vaesen, 2012; Orzack & Sober, 1993). In philosopher Jim Woodward's recent exploration of four different types of robustness, including what he calls 'inferential robustness,' it is confirmatory only in a very narrow (and admittedly scientifically extremely unrealistic) range of circumstances: inference to the robust claim involves the assumption that a "complete" set of models under consideration includes a "true" model, and the parallel in probabilistic terms (2006, pp. 219–224). In Woodward's lovely understatement, "its range of applications looks rather limited" (2006, p. 222).

Here, I pursue a view related to that of Richard Levins (1966), William Wimsatt (1981, 2007), Michael Weisberg (2006), Weisberg and Reisman (2008), and Jay Odenbaugh (2011, ms), (a group henceforth abbreviated as 'LWWO'), and I expand arguments first made in (Lloyd, 2009). I describe a distinct type of informal inference using robustness, which I call 'model robustness.' It is based not only on the agreement or convergence of the empirically correct *outcomes* or *predictions/retrodictions* of a group of models, but *also* on the independent *empirical* support for the

E-mail address: ealloyd@indiana.edu.

[1] A retrodiction is a model result that describes phenomena that have already occurred. The advantages of modeling past phenomena are many, especially in that such models can be compared to any empirical measurements, data, or observations of such phenomena, as well as to observable proxies for any processes or phenomena that are claimed to have occurred. I will refer henceforth to "predictions/retrodictions" to remind the reader that the models to which I refer in this paper all relate to phenomena that have already occurred.

[2] In the context of the Pirtle et al. study, they refer to both predictions and retrodictions.

http://dx.doi.org/10.1016/j.shpsa.2014.12.002

variety of *assumptions and features* of a span of models that all share a common 'causal core.' The 'causal core' is a dependency among key variables or parameters of interest, common to all members of the model-type (a 'model-type' is where the models have in common a type of structure, sharing general characteristics, in which "certain parameters are left unspecified" (van Fraassen, 1980, 44)).

A model-type may be first equated, for our purposes, with the climate scientists' 'conceptual model,' in which key causal connections and processes are envisioned, but the details and/or parameters are not yet specified. Those causal ideas may be instantiated in an actual simulation model, (a GCM or simpler model), which specifies the previously-unspecified parameters, and which contains details we discuss in Section 3. The causal core of the model-type, i.e., the causal processes and explanations of interest, are endorsed directly and indirectly by both the successful predictions/retrodictions and the empirical support of assumptions of the models, and they are partly responsible for the predictions/retrodictions being correct. Thus, 'model robustness' involves all this direct and indirect empirical support for the *explanatory causal core* of the model-type, and by means of this causal core, the model prediction/retrodiction is also empirically supported. Note that this is very different from other philosophical meanings of 'robustness,' which are usually solely defined in terms of the convergent *predictions* of a group of possibly (or even preferably) unrelated models.

2. Introduction to robustness

The key insight comes from biology. Theoretical ecologist Richard Levins described robustness in a landmark book concerning theoretical biological methodology in 1968. There, he noted that when there are multiple, varying models of the same phenomenon in nature, the scientist often focuses on a common causal structure in the models, represented in Fig. 1 by the rough-edged bursting central node.

This causal core[3] reliably relates to a common outcome, **T**, regardless of the differing idealizations or assumptions, represented by the varying arrows, made in the various models. Ultimately, in the hands of philosophers Wimsatt, Weisberg, and Ken Reisman, Levins' insight is translated into a claim that a common structure in the models, the shared bursting node, represents a real world phenomenon or *cause* (Calcott, 2011, p. 284; Levins, 1966, p. 431; Weisberg, 2006, p. 737; Weisberg & Reisman, 2008, pp. 114–115; Wimsatt, 2007, p. 60).

Biologist Steve Orzack and philosopher Elliott Sober argue against Levins' view, saying that since his proposed robustness inference does not involve examining data, it is a distinct and non-empirical form of confirmation, one that they reject as ineffective for making predictive inferences (1993, pp. 541–544). Levins, on the other hand, argues that Orzack and Sober have mischaracterized robustness analysis, and insists that there are, contrary to their claims, central, *empirical* aspects of robustness. Specifically, Levins emphasizes the *empirical support present for the common core* in the models, as well as *for the various assumptions* appearing in the variety of models under investigation (Levins, 1993, p. 554; see Fig. 2).

In sum, Levins-style robustness analysis does indeed involve empirical evidence, but that observational and experimental evidence focuses on the model's *assumptions and core structure* (Fig. 2), not its prediction/retrodiction (see Fig. 3).

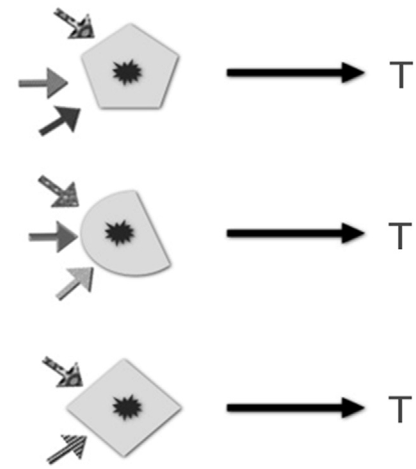

Fig. 1. Varying nodes with differing assumptions of various models, all predicting T.

Thus, we can see that Orzack and Sober had a different target, namely predictive inference to the model's outcome, as shown in Fig. 3, about which they were likely correct, although that is a topic for a different paper. Levins, in contrast, emphasized the key empirical evidence for the model structure under consideration from the other side of the arrow to the model outcome, as shown in Fig. 2. Thus, we can see that they were talking past one another on this point.

Still, that does little to establish a positive claim on Levins' behalf. Although the LWWO line of analysis has been extremely helpful by its frequent insistence on empirical support for assumptions of the model, (but see Odenbaugh & Alexandrova, 2011), they have not adequately described how or why the inference works to increase the confidence of the investigator in the causal core. Philosopher Brett Calcott, commenting on the Levins-Wimsatt approach, writes that although a series of models might be seen as robust, "by itself this is not enough to confirm anything. The models must be connected to the world, and this relies on making good on the resemblance they are meant to have with the phenomena in question" (Calcott, 2011, p. 287; Houkes & Vaesen, 2012).

In this paper I describe how 'model robustness,' in the context of climate science, provides—against the usual philosophical claims, e.g. Woodward, 2006—a *confirmatory virtue*, through discussing the case of greenhouse gas models of Twentieth Century warming.[4] When we start with a LWWO-type approach, ensure the independent empirical support of model assumptions, in addition to the predictive success of the models, and add a bit of reasoning about variety of evidence, we can help supply a philosophical confirmatory framework for the reasoning about robustness being done by the climate scientists. Philosophers Wybo Houkes and Krist Vaesen agree with my previous set up and conclusion that, contrary to the traditional philosophical view (e.g. Woodward, 2006, etc.), confirmation through robustness *may* occur (2012; Lloyd, 2009, 2010), but offer no explanation regarding *how*, as I do here. The result is intended to apply to many scientific cases, where the structure of complex model types and causal foci appears.

An interpretive note regarding my treatment of models and confirmation: When discussing models and modeling, I assume that the models (and climate simulations, which I treat as large models, although they may, under different circumstances, be treated as distinct (Edwards, 2010)) are indicated as similar to, and intended to represent particular aspects of the real world climate

[3] More correctly, such a structure is a 'causal focus,' as it can represent parameterizations, parameter values, etc. But I will call it a 'causal core' here, as that is a common use.

[4] Weisberg and Reisman, in contrast, in their very useful discussion of the Lotka Volterra models, are not arguing for a confirmatory virtue (2008, p. 108).