Studies in History and Philosophy of Science 56 (2016) 95–102

Contents lists available at ScienceDirect

Studies in History and Philosophy of Science

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

Computer models and the evidence of anthropogenic climate change: An epistemology of variety-of-evidence inferences and robustness analysis

Martin A. Vezér

The Network for Sustainable Climate Risk Management (SCRiM), Earth and Environmental Systems Institute, Pennsylvania State University, United States

ARTICLE INFO

Article history:
Received 27 September 2015
Received in revised form
31 December 2015
Available online xxx

Keywords:
Variety of evidence;
Robustness;
Computer models;
Epistemology;
Climate change;
Global warming

ABSTRACT

To study climate change, scientists employ computer models, which approximate target systems with various levels of skill. Given the imperfection of climate models, how do scientists use simulations to generate knowledge about the causes of observed climate change? Addressing a similar question in the context of biological modelling, Levins (1966) proposed an account grounded in robustness analysis. Recent philosophical discussions dispute the confirmatory power of robustness, raising the question of how the results of computer modelling studies contribute to the body of evidence supporting hypotheses about climate change. Expanding on Staley's (2004) distinction between evidential strength and security, and Lloyd's (2015) argument connecting variety-of-evidence inferences and robustness analysis, I address this question with respect to recent challenges to the epistemology robustness analysis. Applying this epistemology to case studies of climate change, I argue that, despite imperfections in climate models, and epistemic constraints on variety-of-evidence reasoning and robustness analysis, this framework accounts for the strength and security of evidence supporting climatological inferences, including the finding that global warming is occurring and its primary causes are anthropogenic.

When citing this paper, please use the full journal title *Studies in History and Philosophy of Science*

1. Introduction

The global climate system is large and complex, with many causal factors interacting to produce all sorts of climatic phenomena. To study climate change, scientists employ computer models, which are imperfect representations of target systems. The most detailed, high-resolution models omit representations of major features of the planet that affect the climate (e.g., mountain ranges) and contain parameterizations that simplify complex climatic processes (e.g., cloud formation). Knowing whether a given climate model provides insight into questions about a target system at various scales and about its responses to different perturbations is, therefore, often difficult to determine.

Among the findings in the Fifth Assessment Report (AR5) of the Intergovernmental Panel on Climate Change (IPCC) is the conclusion that "[i]t is *extremely likely* that human activities caused more than half of the observed increase in global mean surface temperature [GMST] from 1951 [to] 2010" (Bindoff et al., 2013, p. 869). Important sources of evidence for this conclusion are the results of computer model simulations. Given that each of the modelled worlds used to study the climate is substantially different from the earth, how do scientists use computer simulations to generate knowledge about the causes of observed climate change?

I address this question by employing a distinction between evidential strength and security (Staley, 2004), and ideas of variety-of-evidence reasoning and robustness analysis (Lloyd, 2015), focussing on the epistemic advantages of drawing on a range of observations, experiments and models, expanding upon related philosophical enquiries into this field of study (Edwards, 2010; Katzav and Parker, 2015; Lloyd, 2009, 2010; Norton & Suppe, 2001; Oreskes, 2007; Parker, 2006, 2010; Weisberg, 2006). Although other accounts also emphasize the importance of multiple sources of data, further analysis of the relationship between

E-mail addresses: vezer@psu.edu, martinvezer@gmail.com.

http://dx.doi.org/10.1016/j.shpsa.2016.01.004
0039-3681/

varieties-of-evidence and robustness arguments can both clarify the epistemology of current climate modelling research and advance philosophical debates about the structure of reliable scientific methodology. Applying this epistemology to case studies documented by the IPCC, I argue that varieties-of-evidence reasoning and robustness analysis account for the strength and security of evidence supporting important climatological inferences that make use of imperfect computer models.

The approach of the paper is as follows. Section 2 illustrates the logic of variety-of-evidence inferences by describing a set of evidential pathways that converge in support of the global warming hypothesis. This example presents a clear case of strong evidential reasoning, which is a precursor to the more complex case of addressing the role of computer model simulations in climate change attribution studies. Section 3 explains how multiple imperfect models contribute to the security of scientific knowledge because multiple independent approximations of a target system provide alternative evidential pathways to support particular hypotheses. I explain how model agreement in studies of anthropogenic climate change exemplify the confirmatory roles of varieties-of-evidence reasoning and robustness analysis and I address some of Parker's (2011) concerns about evidential security, while also acknowledging important limitations of modelling methodology. Section 4 concludes the paper by summarizing the main points of the argument.

2. Evidence of global warming

An evidential pathway is a collection of information that supports an inference. An evidence claim is a proposition stating that some information is evidence for a hypothesis (*H*). Kent Staley (2004) suggests that attaining multiple lines of evidence can increase the plausibility of *H* by enhancing its evidential strength and security.[1] While he draws a distinction between evidential strength (i.e., the degree to which evidence indicates *H*) and security (i.e., the insensitivity of an evidence claim to changes in some evidential pathway), he sets aside the question of how one can increases the strength of an inference (p. 468). As suggested by the analysis of Elisabeth Lloyd (2015), variety-of-evidence reasoning can, however, account for the way in which multiple lines of evidence strengthen an inference. According to this perspective, evidential strength is attained when multiple independent evidential pathways indicate *H* such that it would be unlikely that these various lines of evidence would agree, if *H* were incorrect. In this context, it is not the improbability of the individual lines of evidence occurring that increases evidential strength but the improbability of their agreement, if *H* were incorrect.

Variety-of-evidence reasoning is essentially the severe test criterion advocated Mayo (1996), though Staley applies it somewhat differently, as a constraint on appeals to robustness as "second order evidence."[2] The question regarding this particular use of the severity criterion as a measure of evidential strength is how exactly its satisfaction is established. On the severity criterion, to show that the convergence of different modelling results is improbable, supposing *H* to be false requires considering the alternatives to *H* and the probabilities those alternatives confer on the convergence of evidence. This criterion is stronger than the formulation of William Whewell's "consilience of inductions" as the argument that "[n]o accident could give rise to such an extraordinary coincidence" (Whewell, 1858, p. 88), which only establishes that *H* has passed a severe test if an "accident" is the only alternative to *H*. In the current context, alternative hypotheses to account for global warming which have been established as improbable include the claims that natural variability or other forcing mechanisms, such as variations in solar input, volcanic activity, and orbital cycles, are the main drivers of recent warming (Bindoff et al., 2013).

Climate science contains many examples of variety-of-evidence reasoning, the case of global warming providing a vivid illustration. The IPCC exemplifies this point in its evaluation of the Global Warming hypothesis (*GW*)—that is, the proposition that "Global Mean Surface Temperature has increased since the late 19th century" (Hartmann et al., 2013, p. 161). The evidence converging on *GW* includes observations of different interconnected components of the climate system, the collection of which would be unlikely to occur if *GW* were incorrect. Land-surface weather stations provide the most direct evidential pathway supporting *GW*, but other climate indicators include measured changes in atmospheric and oceanic temperatures at various heights and depths; in glacier mass, snow coverage, and sea ice extent; in sea level; and in atmospheric water vapour content.

Fig. 1 contains 10 graphs depicting the consilience of evidence supporting this hypothesis. Since the atmosphere and hydrosphere are interconnected fluid bodies, a warming at the earth's surface produces detectable effects at different levels of the atmosphere and ocean. Some of the energy absorbed by the climate system is stored in the oceans, and this energy uptake is detectable in global ocean heat content records going back to the 1950s. Another line of support is the change in the amount of water vapour in the atmosphere, i.e., its specific humidity, measurements of which show a positive change both over the land and the oceans. Observed sea-level rise is another line of support; warming oceans result in water expansion, leading to rising sea levels, which are further heightened by additional water input from melting glaciers and ice sheets and changes to the storage and usage of water on land. The cryosphere (i.e., the frozen parts of the planet) is also affected by changing temperatures. Snow cover, particularly during the spring, is sensitive to temperature changes. Since the 1950s, Northern Hemisphere spring snow cover has declined. Similarly, Arctic sea-ice losses are detectable in satellite records, particularly at the end of the annual melt in September, which is the time of its minimum extent. For at least the last 20 years, the amount of ice contained in glaciers globally has declined (Hartmann et al., 2013).

Since observations of these diverse phenomena are both consistent with *GW* and would be improbable if *GW* were incorrect, these findings comprise a case of variety-of-evidence reasoning. With the detection of *GW* thus well established, the next question to consider is: What has been causing this warming?

3. Climate modelling, robustness analysis, and anthropogenic global warming

While a diversity of evidence increases the strength of *H* by drawing on multiple sources of information, this mode of reasoning is distinct from the idea of evidential robustness. The literature on this subject distinguishes several notions of robust reasoning pertaining to theorems, phenomena, modes of detection (Calcott, 2011; Levins, 1966, 1993; Orzack and Sober, 1993; Wimsatt, 2001), inferences, measurements, derivations, causal relationships (Woodward, 2006), parameter values, mathematical structures, representation frameworks (Weisberg & Reisman, 2008), computer models and simulations (Houkes & Vaesen, 2012; Lloyd, 2015; Muldoon, 2007; Parker, 2011). Although distinctions among these ideas are philosophically interesting, for the purposes of this paper I will focus on a general sense of robust evidential reasoning, which

[1] Also see Staley (2011, 2012).

[2] Staley also links this criterion to Campbell and Fiske's (1959) criterion of "discriminant validation."

[3] For related historical accounts, see Fleming (1998) and Weart (2011).