Climate Modeling: Comments on Coincidence, Conspiracy, and Climate Change Denial

Michael Goldsby
School of Politics, Philosophy, and Public Affairs
Washington State University

John Koolage
History and Philosophy Department
Eastern Michigan University

Climate Modeling: Comments on Coincidence, Conspiracy, and Climate Change Denial

Abstract

Despite overwhelming evidence that climate change is real and represents a serious challenge for human flourishing, many still hold that climate change is not a credible threat – including a surprising number of broadcast meteorologists. In this article, we look at the logic that underwrites such an attitude, which typically appeals to a distrust of climate models, natural variability, or the presence of a conspiracy. Using a model selection framework, championed by Elliott Sober and Malcolm Forster, we will show that appeals to such lines of reasoning do not provide sufficient warrant to dismiss the predictions of climate models.

1 Introduction[1]

There is overwhelming evidence for the existence of anthropogenic climate change. The evidence further suggests that the change, if left unchecked, will be for the worse (IPCC 2013; 2014; Oreskes 2004). Despite this fact, skeptics and deniers remain. For example, author Michael Crichton in an "Author's Message" at the end of his novel, *State of Fear*, claims that climate science lacks the requisite certainty to claim that meaningful climate change is happening (2005). In *A Perfect Moral Storm*, Stephen Gardiner (2011) briefly examines climate change denial by rationally reconstructing the skeptical position exemplified by Crichton's claims. At the heart of Gardiner's critique is a parity argument such that any epistemic standard that would endorse the findings of sciences such as evidence based medicine would likewise endorse the findings of climate science. His diagnosis of climate change denial is that deniers appeal to inappropriate epistemological standards. While Gardiner's analysis of a relatively naïve version of skepticism and denial is essentially correct, there is a brand of skepticism and denial that is more pernicious and promulgated by a better educated and more sophisticated group – namely broadcast meteorologists.

Recent news from George Mason University's Center for Climate Change Communication (4C) shows that "more than 9 in 10 TV weathercasters [surveyed] have concluded that climate change is happening" and that "nearly 9 in 10 think that human activity is at least partly responsible" (Maibach, et al. 2015). While these results indicate some shift in the attitudes of TV weathercasters when compared to earlier surveys (e.g., Maibach et al. (2010; 2011)),[2] there is reason to believe that such a shift does not represent a telling victory in the climate change debate, because Maibach et al. (2015) also reveals that few weathercasters acknowledge that such anthropogenic climate change is a serious threat. For example, when asked about various climate impacts that have occurred in the past 50 years, only around 3 in 10 believe that there have been significant impacts, and fewer than 3 in 10 hold that there will be significant impacts in the next 50 years (32-33). Furthermore, fewer than 50% of those who claim that climate change is happening hold that human activity is the *major driver* (Maibach et al. 2015, 31).

---

[1] Acknowledgements redacted for blind review.

[2] Maibach, Wilson and Witte (2010) discovered that nearly 1 in 2 TV weathercasters expressed skepticism or outright denied that global warming was happening.

These responses are consistent with Crichton's skeptical claims and very similar to the contrarian views expressed by Patrick J. Michaels and Robert C. Balling (2000).[3] Thus, if Crichton, Michaels, and Balling represent the skeptical position on climate change, then it appears that a sizable number of TV weathercasters are skeptics and deniers in the same sense. In light of that fact and the fact that most TV meteorologists are the primary science reporters for their respective stations with the ability to reach very large segments of the population, this skepticism and denial is deeply distressing.

In this paper, we will build on Gardiner's epistemic criticism of climate change skepticism/denial in order to extend that criticism to the more sophisticated brands held by some TV weathercasters and contrarian scientists. In the next section, we will draw a distinction between climate skepticism and climate denial, and we will identify three central claims the conjunction of which nearly every skeptic or denier denies is reasonable to believe. Then, we will rehearse Gardiner's argument against unsophisticated skeptics and deniers, but we will show how a more sophisticated denier or skeptic might evade his critique. We will also identify three main lines of reasoning that seem to motivate the skeptical position: (i) a distrust of climate models; (ii) the belief that apparent climate change is mere coincidence; and (iii) that apparent climate change is the result of a conspiracy. In Section 3, we introduce model selection theory, championed in the philosophical literature by Malcolm Forster and Elliott Sober (1994; 2010), to highlight the relevant similarities between weather modeling and climate modeling in order to extend Gardiner's parity argument to more sophisticated skeptics and deniers. We will argue that if one thinks the predictions of weather models represent our best bet with regard to weather issues, one should likewise trust the predictions of climate models with respect to climate issues.

Finally, in section 4, we will use the model selction framework introduced in section 3 to analyze the latter two skeptical positions. Our intent in this essay is not to show that the skeptical position is false, as plenty of effort has already been devoted to just such a project. We doubt that we have much to add in that regard, especially since climate change skepticism and denial has proven to be resilient. Instead, our goal is much more modest. We intend to show that the predictions derived from climate models, whether those models are true or false, are far more likely to provide better predictions than those suggested by the latter two skeptical hypotheses. Since the chief issue is what the state of the future climate will be, knowing which model is expected to have better predictions is important. Ultimately, we will argue that even the more sophisticated skeptics and deniers have much more work to do to make their case – work that has not been done.

2 Climate Skepticism and Denial

One basic tenet of inquiry is to remain willing to revise one's beliefs in response to evidence. That requires an open mind and the ability to recognize that the nearly ubiquitous uncertainty in one's investigations can sometimes lead to error. Given that potential for error, *some* doubt regarding the claims made by experts is warranted and even rational. That is to say one must be willing to recognize

[3] See Oreskes and Conway (2010) for a historical account of the work of such contrarians.

that even the best available science can be in error, and therefore subject to revision. To fail to do so would be to commit oneself to dogmatism.

Since dogmatic beliefs are unresponsive to evidence, dogmatism is, to be certain, an epistemic vice. Dogmatism is often mistakenly contrasted with skepticism. Since skepticism maintains a robust respect for doubt it is often seen as far more intellectually responsible than dogmatism. As such, those who deny climate change often bill themselves as skeptics, even if they outright deny that climate change is happening. Likewise, those who believe that climate change is a real threat tend to refer to both climate skeptics and deniers as "deniers." The idea that is operative here is that skeptical positions are non-dogmatic and non-skeptical positions are dogmatic. We think this is a mistake, because as Brian Skyrms (1984) points out, dogmatism is an equal opportunity intellectual vice. If one's skeptical position is unresponsive to the evidence, then one is still dogmatic. Furthermore, a dogmatist is still a dogmatist even if she willfully chooses to believe a claim that (through no fault or merit of her own) is currently supported by the best available evidence.

The assumption that skeptics are reasonable, and deniers are not, is itself prone to error. Some skepticism is unwarranted and some denial is reasonable. For example, if one were to hold a skeptical position with regard to tautologies or contradictions, then such a person's rationality would be open to criticism in most contexts that matter.[4] Similarly, while there is some uncertainty about whether there really is a United States of America, suspending judgment on whether the States do indeed exist seems to depart from the ordinary standards of epistemic rationality. On the other side of the coin, denying a contradictory claim (e.g., 2+2=5) or that the Earth is flat seems to be reasonable. We think it would be inappropriate to criticize such deniers. Our point here is simply that it is mistake to assume that skepticism is always reasonable and that denial is not. Any potential flaw of rationality has nothing to do with which doxastic attitude one takes but rather it has to do with whether such an attitude is supported by and responsive to the evidence. The issue that we are exploring is whether one can reasonably deny or be skeptical of climate change. In order to avoid begging the question, we will not assume that climate skeptics are reasonable while climate deniers are not. Thus, we will use the terms 'skeptic' and 'denier' as merely a description of one's doxastic attitude toward a claim or set of claims – where a denier denies the truth of the claim in question and a skeptic suspends judgment.

2.1 Varieties of Climate Skepticism/Denial

The team of scientists working on the Intergovernmental Panel on Climate Change (IPCC) Assessment Report 5 (AR5), make three central claims (among many others) in which they have high confidence:

[4] It may be worthwhile in a philosophy classroom to explore whether tautologies can be doubted – when discussing Cartesian skepticism, perhaps. However, we do not hold that the context of the climate debate is the same as the philosophy classroom. Those who deny or are skeptical of climate change are generally not global skeptics.

> (CC1) without mitigation, the global mean surface temperature will continue to increase, with the likely range of that increase being between 2.6 and 4.8°C, under a business as usual scenario;[5]
>
> (CC2) the major driver of this change is anthropogenic; and
>
> (CC3) increases in global mean surface temperature of 2°C or more will be bad for human civilization (IPCC 2013; 2014).

So this means that there at least three different claims that a skeptic or denier might doubt or deny. Denying one claim does not entail the denial of another, nor does denying one claim require that one assent to any of the others.

This suggests that there is wide variety of positions that a skeptic or denier might take with regard to CC1-3. Given this wide range of opinions, one might worry that we cannot meaningfully talk about climate skeptics and deniers as a group. Nonetheless, there is one thing that they do have in common. If one denies or is skeptical of one of the three claims, then one denies the following claim:

> (CC4) Given the evidence and ordinary standards of epistemic rationality, it is significantly more reasonable to believe the conjunction of CC1, CC2, and CC3 than to deny it.

It should be clear that if one thinks that CC1 or CC2 or CC3 is false, then one would likewise deny that is reasonable to believe the conjunction of all three claims. Similarly, if one is skeptical of CC1-3, then one must think that the evidence is insufficient to assent to CC1-3. If that is the case, then such a skeptic would deny that it is reasonable (given the evidence available) to believe the conjunction of CC1, CC2, and CC3. So, if one denies or is skeptical of any one of CC1-3, then one must deny CC4.

Before moving on, we should say something about what we mean by "ordinary standards of epistemic rationality." We take a rough-grained evidentialist approach to what it means for an agent to be epistemically rational in this context, but we also remain neutral with regard to the exact model of epistemic rationality. One can easily replace "ordinary standards of rationality" with a more specific model without loss of generality.

Nonetheless, one might object to any evidentialist model of rationality, by maintaining that one can reasonably believe a claim even when one does not have sufficient evidence to believe that claim. In other words, there might be pragmatic reasons for believing a claim, even when the evidence is insufficient. However, we think such considerations are not salient within the context of the climate

[5] Strictly speaking the AR5 provides four likely ranges for the increase of global mean surface temperature with each fitted to a particular Representative Concentration Pathway (RCP). These RCPs range from low concentrations of greenhouse gases (GHGs) (RCP 2.6) to high concentrations of GHGs (RCP 8.5) and might be seen as correlating with scenarios of high mitigation and no mitigation respectively – although there are many scenarios under which a particular RCP might obtain. RCP 8.5 can be realized by the "business as usual" scenario, where the likely increase is expected to be between 2.6°C and 4.8°C.

change debate. First of all, climate change skeptics and deniers rarely if ever marshal pragmatic arguments to defend their denial or skepticism, and they do not accept pragmatic reasons for assenting to CC1-3. Instead the claim on offer is that once one takes all the evidence on balance, one cannot reasonably maintain that the conjunction of CC1, CC2, and CC3 is true. Furthermore, even if they did think that it was pragmatically reasonable for some to believe that CC1-3 were true (perhaps due to financial incentives), they seem to maintain, nonetheless, that the evidence alone does not support such a position. Thus, the issue at hand is whether the evidence supports the claim that climate change is happening not whether it is rational to believe so in a broader sense. [6]

2.2 Gardiner, Foundationalism, and CC4 Deniers

Michael Crichton (2005) provides a bullet-pointed summary of his actual position on the global climate change debate. His conclusions include:

- We are … in the midst of a natural warming trend that began about 1850, as we emerged from a four-hundred-year old cold spell known as the "Little Ice Age."
- Nobody knows how much of the present warming trend might be a natural phenomenon.
- Nobody knows how much of the present warming trend might be man-made.
- Nobody knows how much warming will occur in the next century. … If I had to guess – the only thing anybody is doing, really – I would guess the increase will be 0.812436 degrees C. There is no evidence that my guess about the state of the world one hundred years from now is any better or worse than anyone else's. (We can't "assess" the future, nor can we "predict" it. These are euphemisms. We can only guess. An informed guess is just a guess. (571-3)

It should be clear from Crichton's statements that he denies CC4, as he is skeptical of CC1 and CC2. We can also see three lines of reasoning that emerge from his statements: (i) there is insufficient certainty in climate science; (ii) the evidence does not sufficiently rule out that the warming trend is a product of natural climate variability; and (iii) the climate models, used by climate scientists, cannot be trusted because "We can only guess."

Still despite those claims, there is a puzzle for Crichton. Why do climate scientists, nevertheless, maintain that anthropogenic climate change is happening? Crichton does offer a diagnosis for why he thinks climate scientists insist that CC4 is true: (iv) any apparent evidence in favor of CC4 is the result of a conspiracy. This constitutes a fourth line of reasoning behind his denial.

Crichton does not suggest that the conspirators are the nefarious sorts like the villains of his novel. After all, such cartoonish villainy is what makes a thriller so thrilling, and poetic license is more at home in fiction than in a considered opinion about why climate scientists believe as they do. However, he does suggest that a much more mundane and decentralized conspiracy might be operative:

[6] We would like to thank an anonymous reviewer for pushing us to be clearer on this point.

- We desperately need a nonpartisan, blinded funding mechanism to conduct research to determine appropriate policy. Scientists are only too aware whom they are working for. Those who fund research – whether a drug company, a government agency, or an environmental organization – always have a particular outcome in mind. Research funding is almost never open-ended or open-minded. Scientists know that continued funding depends on delivering the results the funders desire. As a result, environmental organization "studies" are every bit as biased and suspect as industry "studies." Government "studies" are similarly biased according to who is running the department or administration at the time. No faction should be given a free pass. (573)

From the character of the quote above, it seems that Crichton does hold that scientific consensus with regard to climate change is motivated by and directed toward maintaining funding – a much less exciting and a much more decentralized conspiracy, but a conspiracy nonetheless.

Stephen Gardiner (2011) addresses the denial of CC4 by way of addressing Crichton's claims. Gardiner only considers two of the four lines of reasoning that we have identified: (i) the presence of a conspiracy, and (ii) insufficient scientific certainty. Presumably, Gardiner assumes that the distrust of climate models and the assumption of the natural variability of the climate can be adequately answered by addressing these two lines of reasoning alone. He eschews the conspiracy line by pointing out that there is insufficient grounds to say whether it is a "Left" or "Right" conspiracy that is aimed at the public in this dispute, and that conspiracy theories are difficult to rebut (460). Gardiner spends more time on (ii), by arguing that deniers of CC4 attempt to make their case by appealing to the idea that climate change predictions include uncertainties. These uncertainties provide just enough of a wedge to open the door for Cartesian style doubt, says Gardiner.

Cartesian style doubt rests on a theory of justification, known as foundationalism, which suggests that one is justified in believing p if and only if either (i) p is self-justified or (ii) p can be inferred through a chain of arguments that originate with all and only self-justified premises. Since self-justified premises are few and far between (if there are any), and providing an entire argument chain would be tedious (if not impossible), it is easy to generate some doubt for nearly any p.

This problem is made worse, Gardiner thinks, if one finds Hume's Problem of Induction tempting. On this account, reasoning beyond the data, particularly to the future, is always unjustified. As a result, the argument chains would have to be deductive to be justified on this view. Since most of science makes extensive use of non-deductive reasoning, such a requirement when combined with the strictures of foundationalism would lead us to doubt virtually every scientific claim.

The upshot is that if one holds that Cartesian and Humean standards must be met in order for a scientific finding to be used to inform a policy decision, then one will never be able to rely on any scientific finding to make a decision. However, rationality merely requires that we base our decisions on the best available information, even if that information falls short of certainty. Most of those who deny CC4 generally feel comfortable relying on scientific findings that fall noticeably short of eliminating all doubt. Gardiner rightly points out that climate skeptics and deniers are cheating by leveling these concerns at climate science, and not at other sciences (e.g., physics, chemistry, economics, or even evidence-based medicine). Thus, Gardiner insists that if one takes science seriously at all, one must likewise take climate science seriously.

This is a good first pass at the issue. However, as we have already indicated, Gardiner may have been a tad too quick, especially with regard to more sophisticated CC4 deniers. Nonetheless, there is an important lesson that we ought to draw from Gardiner's discussion – *ceteris paribus*, one is not allowed to apply different epistemic standards in one area that one would not apply in another.

While we intend to expand Gardiner's project, we will part ways with Gardiner here for three reasons. First, for any sophisticated denier, rebutting Gardiner's proposal is easy: reject foundationalism and endorse any epistemology that does not require returning all the way to self-justified beliefs for justification. For example, one could endorse a kind of externalism or any sort of Bayesian epistemology, which makes sense of confirmation, disconfirmation, evidence, and even credence, *without* foundational, self-justified beliefs. Second, there is a group of sophisticated deniers who subscribe to a non-foundationalist epistemology, endorse non-deductive reasoning, and recognize that a fallibilist/probabilistic science is still a good guide to action – namely TV weathercasters. Furthermore, Gardiner assumed that the reason that CC4-deniers, like Crichton, distrust climate models is due to their conviction that climate science lacks a sufficient amount of certainty. Similarly, he seems to assume that the belief that the current warming trend is the result of natural variation is likewise motivated by the perceived lack of certainty in climate science. However, with more sophisticated deniers of CC4, it is the other way around. They distrust the climate science *because* they distrust the models, and they doubt the predictions of climate scientists *because* they hold that climate science has not sufficiently ruled out that the warming trend is due to natural variation. We intend to address these issues in the sections that follow.

Finally, we believe that more can be said in response to the claims of conspiracy often cited as a reason to deny CC4. It should already be clear that the conspiracy line of reasoning plays an important part in the denial of CC4. As such, we think it is important not to dismiss this considerations out of hand, as that may encourage an equally dismissive response from deniers of CC4. As a result we think it is important to address their denial of CC4 head on. While it is true that conspiracy theories are difficult to rebut, using sophisticated epistemic tools endorsed by TV weathercasters we will show that positing a conspiracy does not provide sufficient grounds to doubt the *predictions* of climate scientists.

2.3 TV Meteorologists and Skepticism/Denial

In a survey conducted by the Center for Climate Change Communication, Maibach et al. (2011) provide a more nuanced picture of TV weathercasters that deny CC1:

- About half of those, who are "undecided" or "unconvinced" about climate change, are skeptical of historic climate and proxy data (tree rings, ice cores, etc.). Further, between two-thirds and three-quarters of the members of these groups also distrust the models climate scientists employ in modeling the climate of the past (let alone the future).[7]

- 100% of those unconvinced that there is climate change assert that "the weather is always changing" and "any current weather variability is simply a part of natural climate variability."
- 64-70% of those unconvinced that there is climate change and 41-59% of those who are undecided said that claims of anthropogenic climate change are due to ulterior motives.
- The unconvinced and undecided, when they do trust the data, distrust what the scientists themselves do with the data, reporting a highest average trust in scientists of 28%.

Four lines of reasoning emerge that roughly parallel those provided by Crichton. First of all, there is an expressed lack of confidence in climate models among TV weathercasters, and secondly they also have worries about the proxy data. This, as Gardiner points out, is oddly arbitrary. If one doubts the models and the admittedly fallible evidence provided by proxy data, then it is hard to see why the models and fallible evidence in meteorology fare much better. We will address this concern briefly, siding with Gardiner that this kind of skepticism is arbitrary or circular.

The third reason that TV weathercasters might deny CC4 is that they maintain that the climate exhibits a natural variability. In other words, CC4 deniers might insist that the warming trend is just a part of an ordinary stochastic process that governs the climate. According to this line of reasoning, climate science can be doubted because climate science has not ruled out the possibility that observed warming trends are merely coincidence.

Finally, CC4-deniers, whether TV weathercasters or otherwise, might claim that there is some sort of conspiracy being perpetrated to mislead or abuse the public. This line of reasoning does not impugn the science per se, but rather accuses mainstream climate scientists of things such as tampering with the data and suppressing contrary results in order to further some less-than-noble end. In what follows, we will show that a single framework will shed some light on all four of these skeptical lines of reasoning. While it should not be too surprising that the framework will address the first two, it is surprising, however, that it can be used to address the other two as well.

3 Probabilities, Forecasting, and Model Selection Theory

Good weather forecasting is a complicated endeavor. There are a number of well-understood sources of weather, such as atmospheric circulation – which is largely a result of the Earth's size, rotation rate, heating, and atmospheric depth (Wallace and Hobbs 2006). However, these global features of weather and climate are poor predictors of local weather, which is affected by all sorts of more fine grained variables – such as local altitudes, location of large heat masses (such as cities), and so on. As a result, the meteorologist must have some way of getting the well understood, largely deterministic, features of

[7] 80% of the "undecided" distrust computer models of climate change, and 72% distrust the ability of such models to replicate past climate events; whereas 87% of those who are "unconvinced" distrust the computer models and 69% distrust the ability of those models to replicate past climate events.

atmospheric circulation *and* the salient local features to produce predictions about the weather on a local scale. This requires the use of models, which is the epistemic feature upon which a number of our arguments will turn.

Anyone can forecast the weather. For example, observations of changes in barometric pressure and visual confirmation of cloud cover can be used to predict precipitation. Good forecasting, however, depends on a much better understanding of which variables are most predictive, under what conditions they are most predictive, and a collection of atmospheric facts. Furthermore, since most predictions are based on changes in conditions over time, weather forecasters use a stepwise procedure (an iterative process of data entry) to make their predictions. In actual practice, real time data is fed into a computer program, which uses models to give an estimate of the chance of precipitation.

3.1 Disentangling Probabilities

The *probability of precipitation* is a measure of the confidence that precipitation will occur given current conditions (atmospheric and local) and a particular area, known as a station.[8] The feature of interest is the "confidence that precipitation will occur," which is expressed as a probability.

It is important to note that there are several confidences expressed as probabilities at stake with weather forecasting. The first is the probability there will be rain, for example, at time t somewhere in area S. This is determined by developing and "training" a model with previous data, and then inserting current local values for the variables of interest into the fitted, or "trained," model to make the prediction. The resultant probability is just the output of a fitted model. This is the probability that is reported in standard weather reports. There is another confidence expressed by a probability that is operative in weather forecasting – the confidence one has in the particular fitted model itself. That is, how confident one should be that the model's projected probability of precipitation will be correct. It is the second probability, or confidence, that is doing the epistemic heavy lifting in the weather forecasting context, and consequently that probability will be the focus of the remainder of this paper.

However, let us briefly return to the probability of precipitation. Let A_i be the set of atmospheric conditions ($a_1, a_2, \ldots, a_n$) present at some time i, and let V_i, be the set of all the local variables ($v_1, v_2, \ldots, v_n$) for i. It is reasonable to conclude that probability of precipitation at time t is a function of A_t and V_t. A naïve approach for determining the probability of precipitation at time t would be to look at the weather record for those times where the atmospheric conditions and the local variables are identical to those at time t, and then conclude that the probability of precipitation is simply that proportion of those days with identical conditions in which precipitation had actually occurred. For example, suppose that for every i prior to t, where A_i and V_i is identical to A_t and V_t, respectively, that it rains 7 out of 10 times, the naïve conclusion would be that the probability of precipitation is likewise 70%. In other words, according to the naïve approach, the evidence we have for precipitation is merely the relation between current atmospheric and local conditions, on the one hand, and past rain events under those same

[8] Henceforth, we will omit the "station" qualifier for purposes of concision. However, when we speak of the probability of precipitation, it will be safe to assume that we are talking about the probability of precipitation at a station.

conditions on the other. There is a reason that forecasters do not do this, however, and there is a reason they should not do this.

The naïve approach suggested above fails to take into account two things known to most weathercasters. First of all, the weather changes rapidly, so the naïve approach lacks some complexity that would render a more precise estimate of the probability of precipitation that forecasters wish to make. For example, hourly forecasts would not be served well by the naïve approach, because much of the weather record is not fine-grained enough to provide the necessary information. Secondly and more importantly, the sheer volume of variables that might influence the weather is immense. Forecasters at the National Weather Service, for example, consider over 175 potential variables to generate their forecasts (Krzysztofowicz & Maranzano 2006). These variables include items such as total precipitation in the last 3 hours, relative pressure and humidity, and so on. Given the sheer number of known variables that might affect the weather, it is unlikely that any two time periods, x and y, are such that the ordered pairs, $\langle A_x , V_x \rangle$ and $\langle A_y , V_y \rangle$, are identical. As such, it is just as unlikely that meteorologists will get a sample size large enough to get a good relative frequency of precipitation upon which they might base their estimates, as the naïve approach would have it.

One obvious solution would be to use a subset of the known atmospheric and local conditions that affect the weather to form a sample. If the subset of conditions is smaller, then it is more likely that a meteorologist will have more instances that could be used to infer the probability of precipitation. The current benchmark computer program used by the National Weather Service, the Model Output Statistics (MOS) technique does truncate the number of variables in order to make a prediction. It selects a maximum of 15 variables from over 175 candidate variables to generate the model it uses for forecasting (Krzysztofowicz & Maranzano 2006).

This is how a precipitation prediction is made; however, there is yet another issue that complicates matters. In order to make the truncated models work, forecasters would need a good understanding of which conditions matter most. That means the real question about confidence is hiding just behind this story – how does one know that one is using the right model? That is, the story about how a model is developed to generate the probability of precipitation is pretty straightforward, but it is the confidence one has in the model that is the epistemic workhorse. The confidence one has in the model is a result of *evidence*, unlike the confidence value used in the probability of precipitation, which is the result of a properly employed model.

There are a number of ways to assign a level of confidence to one's model (i.e. to properly select a model). According to Sargent (2005), two widely used methods in weather modeling, are what he calls the *Face Validity* (FV) method and the *Comparison to Other Models* (CTOM) method. With regard to FV, the basic idea is that an expert recognizes a good model when she sees one. That is to say that she makes an *a priori* assessment of whether the model is good enough to do the job based upon her past experience with other models. Certainly, if the model lacks FV, that might count as some (defeasible) reason to doubt the model. However, the mere fact that the model "looks good" to an expert ought not count for much in raising one's confidence in the model's predictive abilities. A similar worry can be leveled against the second method. As Sargent describes it, the CTOM method looks to how well the model agrees with other models developed by other experts. Rough coherence with other models might indeed be a sign of a good model in some contexts, but it is not clear that mere coherence ought to increase one's confidence in the model's predictions in every context. For example, it may be that the extant models with which one might compare a newly developed model are awful predictors. In which case, failure to cohere with the previous models might actually be a good thing. Finally, and more importantly, both methods seem to have little or no contact with data in assigning a confidence to the model.

A better method would be to use the data from the weather record as evidence. To do so, we could treat the model and its suggested results as a hypothesis, which can then be tested in terms of fit-to-data to the historical record. This is a more evidence-based solution to the problem. This method *is* used in both meteorology and climate science. That is, just like all good sciences, weather forecasting relies on the same sort of evidence as climate science when deciding which model to use. The content may be different, but the methods are relevantly similar.

3.2 Model Selection Theory, Weather Forecasting, and Climate Science

A model is, essentially, a representation of a target system that can be used to predict the behavior of that system. Models can take a variety of forms, but in the context of climate modeling and numerical weather prediction a model is an *interpreted* mathematical formula, which specifies some number of variables and a number of adjustable parameters. Consider the following toy models:

$$y = a + bx_1 \qquad \text{(VAR1)}$$

$$y = a + bx_1 + cx_2 \qquad \text{(VAR2)}$$

The variables of these models are x_1, x_2, and y, though no interpretation has yet been given; a, b, and c, on the other hand, are the parameters of the models. For a model to make a prediction about the current or future state of some interpreted variable (y in this instance), it needs to be *fitted* or "trained" by some available data. This is how the parameters a, b, and c are assigned values, the result of which is a *fitted model* or curve. Once we have a fitted model, values can be inserted for various variables, and predictions of the dependent variable can be made.

The MOS technique, mentioned above, determines which variables to use in a particular weather model by looking at how "tightly" the variable predicts the weather – precipitation in our example. The number of variables the model employs (and thus number of adjustable parameters) is determined by an arbitrary selection of how "tight" (in terms of variance) the variables are with respect to past predictive successes under the conditions at hand (Krzysztofowicz & Maranzano 2006). In the case of the MOS technique, the number of variables is arbitrarily limited to 15 variables maximum. The upshot is that the MOS technique takes account of two things – how well the model might fit the data and the complexity of the model.

Due to computing demands, this technique may be a practical necessity given the sheer number of candidate variables, but it can be epistemically worrisome. First of all, this sort of technique requires that a computer program already have access to salient variables from which to draw, but given our knowledge of weather processes, this may not be too much of a concern. Secondly, one might worry that to select an arbitrary cut off for the number of variables used in prediction is, well, arbitrary. Nonetheless, limiting the number of variables is in some sense on the right track for dealing with some well-known issues in model selection.

There are two general problems that face model selection relevant to our discussion: (i) more complex models always do as well or better than simpler models when it comes to fit-to-data; and (ii) when making predictions about future events (like the probability of precipitation), maximizing fit-to-data can result in overfitting. In order to illustrate the first difficulty (i), reconsider our toy models. VAR1 is just a special case of VAR2, where $c = 0$. Additionally, VAR1 is simpler than VAR2 as it has fewer adjustable parameters. Since a model's fit-to-data can be determined by the likelihood of its best fitting fitted model, VAR1 cannot fit the data better than VAR2.[9] That is because the best fitting member of VAR1 is *contained in* the set of all possible fitted models of VAR2. As a result, more complex models always fit the data as well *or better* than simpler models.

While more complex models are always able to fit the data better, simpler models are sometimes better at prediction – especially when the data set used for training the model is small (Forster and Sober 1994; Goldsby 2013). This is known as the problem of overfit, and working scientists are all too aware of this phenomenon. Overfit is a problem because the world is noisy due to the fact errors in observation are often unavoidable.[10] As a result, it is possible to fit the training data *too well*, capturing not just signal but also noise. Good model selection frameworks, however, balance the tradeoff between fit-to-data and simplicity by assigning a penalty for the number of parameters used in the model. The confidence we have that a model is good predictor goes beyond how well the model fits the training data; it includes a notion of predictive accuracy. The mathematician Hirotagu Akaike (1973) demonstrated that predictive accuracy[11] can be maximized by optimizing the tradeoff between increasing fit-to-data and minimizing the number of parameters.[12] This has been widely discussed in the philosophical literature (Forster and Sober 1994; 2010; Forster 2000; 2001; Burnham and Anderson 2002).

---

[9] Karl Popper (1959) made a similar point when considering the relation between simplicity and probability.
[10] Additionally, if it turns out that the process is fundamentally stochastic (i.e., not deterministic) or even if the it is fundamentally deterministic, but we have no epistemic access to the fundamental laws governing the process, then the process might be even noisier.
[11] A model's predictive accuracy is simply how well the model predicts new data when trained or fitted to old data. The purpose of model selection theory is to provide estimates of a model's predictive accuracy. If a model is determined to be more predictively accurate than another model using a model selection framework, that fact constitutes evidence that model will provide better predictions. See Forster and Sober (1994; 2010) as well as Forster (2000).
[12] Akaike's theorem provides a framework from which we can compare the estimated predictive accuracy of two or more models. This framework uses the Akaike information criterion, also known as AIC, where the *AIC score of a model M*, $\text{AIC}(M) =_{\text{def}} \log\{Pr[\text{data}|L(M)]\} - k$. The term, $L(M)$, refers to the fitted model of M that has the highest likelihood of all possible fitted models of M when compared to the training data, and $Pr[\text{data}|L(M)]$ is that likelihood. The complexity of the model in terms of adjustable parameters is represented by k. The log-likelihood of $L(M)$ is a measure of how well the model fits the data and k represents a correction for complexity. What is important about $\text{AIC}(M)$ is that it gives us an unbiased estimate of M's predictive accuracy (Forster and Sober

Likewise the MOS technique manages the tradeoff between fit-to-data and complexity by arbitrarily setting a maximum number of parameters, and thus variables, to stave off over-fitting the training data. This approach may be suboptimal when compared to more sophisticated model selection techniques, like that suggested by Akaike's theorem, but given the sheer size of the training data used in weather forecasting, the 15 variable maximum instituted by the MOS technique is likely good enough. Thus, it appears that weather forecasting makes some use of a model selection framework.

A model selection framework also has a place in climate modeling. The climate system is a highly complex system governed internally by the current and evolving states of the atmosphere, the hydrosphere, the cryosphere, the lithosphere, and the biosphere, as well as external forcings such as volcanic eruptions, solar variations, and anthropogenic forcings (IPCC 2013). As such, climate models are far more complex than their weather forecasting cousins. Climate modeling is a technical and highly specialized field. As a result, climate modelers are a sophisticated lot that use a wide variety of tools. Due to space considerations, we cannot provide an in-depth and exhaustive account of climate modeling. Instead, what follows is a brief sketch of climate modeling that will illustrate how model selection tools are often used in the development of climate models.[13]

Most climate models in use today are *coupled models* (e.g., Atmosphere-Ocean General Circulation Models (AOGCMs)). A coupled model is essentially a set of models, with each model in that set meant to represent a sub-system of the larger system being represented. Those sub-models are 'linked' or coupled together via a *coupler,* which is a bit of code that exchanges information of interest between coupled sub-models. In the case of climate models, the information of interest that is exchanged by a coupler usually includes information about heat and moisture (Gent 2012).

There are several reasons why climate scientists and modelers use coupled climate models as opposed to a more unified model. We will mention a few. First of all, unified climate modeling is a relatively new sub-discipline. Rather than reinvent the wheel, climate modelers prefer to use models already developed independently in related sub-disciplines (e.g., models from atmospheric and oceanic sciences, as well as other fields). Secondly, the best performing models for each of the sub-systems often run at differing spatial and temporal scales (Gent 2012). As such, coupling models allows each sub-model to make use of the spatiotemporal scale that is optimal for that sub-system. A third reason is that the interactions between the subsystems are exceedingly complex. Introducing that complexity into a unified climate model would introduce a level of complexity that is not likely to be accompanied by a concurrent increase in predictive accuracy (given the data currently available).

Model selection, using tools like the Akaike Information Criterion (AIC), is used in a variety of ways in climate modeling. First of all, the sub-models that are used in coupled models are typically chosen for

---

1994; Forster 2000, 2001; Sober 2008). The absolute value of the score is not important; what is important is that we can use AIC to compare the estimated predictive accuracy of unfitted models, such that we would expect that the model with the higher AIC score under this formulation would be better at predicting new data (Goldsby 2013).

[13] For a more detailed and highly technical overview of the modeling process in climate science, see IPCC (2013) especially chapter 9. For a less technical overview, one could turn to Maslin (2014).

their ability to optimize the tradeoff between simplicity and fit-to-data relative to that sub-system.[14] Secondly, the choice of spatiotemporal scale used in the sub models and the entire coupled model is itself an exercise in model selection, as one must balance providing fine-grained detail with the ability to make good predictions. Thirdly, model selection is a consideration that is used in the practice of *parameterization* (Delsole and Shukla 2006). Parameterization, in the climate modeling context, is the process of taking a sub-process within a model (one common example is the modeling of clouds) that is more realistically represented by n parameters and replacing it with a representation that uses m parameters (where $n > m$) in order to improve the model's estimated predictive accuracy (IPCC 2013; Gent 2012; Maslin 2014).[15]

So, model selection plays an important role in climate modeling. Similarly, TV weathercasters, who make use of MOS, are endorsing a crude kind of model selection framework, since they think they *should* use the models when they are making their reports. The important point is this: despite the worries we raised with regard to the ham-handed approach to model selection that MOS uses, it nonetheless does use a model selection framework that is relevantly similar to that used by climate scientists – i.e., maximizing fit-to-data constrained by a "size" rule to determine which model is then trained and used to make a specific prediction. Thus there is a basic parity of methods between climate modeling and weather modeling, at least insofar as both practices make use of a model selection framework. So, if a TV weathercaster has a beef with the basic climate modeling methodology, then he or she is confused.[16]

Nonetheless, one who denies CC4 might still maintain either that (i) the models that climate scientists use are suboptimal because they either overfit the data or they don't fit the data well enough for predictive purposes,[17] or (ii) that the data upon which the models are trained are somehow flawed. In

---

[14] Some examples of the explicit use of AIC in the development of sub-models include Jones et al. (2004) in the case of atmospheric models, Vermeer and Rahmstorf (2009) with oceanic models, and Boeckli et al. (2012) with models of the cryosphere.

[15] Parameterization may be used in the development of sub-models, but it might also be used in the 'tuning' of a coupled model at the end of the development process (IPCC 2013, 749-750; Mauritsen et al. 2012). The goal of tuning is to improve the fit between model's simulated values and the observed values of the target system in terms of predictive accuracy. Features targeted in the tuning process not only include the main output value of interest but other simulated features of the coupled model that might be compared to observational data (Mauritsen et al. 2012; Parker 2011).

[16] One potential wrinkle is the use of multiple model ensembles (MME) in climate science. Nearly all of the predictions made by the AR5 use the results from the Coupled Model Intercomparison Project Phase 5 (CMIP5). Strictly speaking, the CMIP5 is actually a set of experiments to test the validity of using MMEs in climate science, but the projections are made as if the CMIP5 is an MME. To make a prediction, an MME uses the results from several models to develop a picture of what the future state of the target system will be. Ordinary weather forecasting does not routinely make use of MMEs. One might point to this difference in technique as a relevant difference between climate modeling on the one hand and weather modeling on the other. We set aside this potential objection for two reasons. First of all, weather forecasting *does* make use of MMEs when tracking winter storms and tropical cyclones (National Hurricane Center 2009). Since TV weathercasters do presumably place their faith in such models when reporting on these weather events, it would be strange to cite this practice as a reason to distrust climate models. Secondly, the results of the CMIP5 are largely consistent with predictions of individual (coupled) models.

[17] For example, such a CC4 denier might maintain that either some irrelevant process is included in the climate model or that some relevant process is left out of the climate model.

order to show that (i) is true one must provide a sophisticated analysis of climate science. That is to say that one must show that an alternative model better optimizes the tradeoff between simplicity and fit-to-data – in short, that an alternative model is more likely to predict new data accurately. A brief survey of climate change denial literature reveals that few actually take that tack.[18] Still, in the next section we will seriously consider the claim that apparent climate change is the result of natural variation and that climate scientists have overfit the data. Additionally we will take a look at the second line of reasoning - nanmely (ii) that the assumptions and the data used by the models are somehow flawed. According to such a view, the data are flawed either because the proxy data are not reliable, or because the data are intentionally skewed as the result of some conspiracy.

4 Model Selection Theory and the Other Skeptical Hypotheses

As we have seen, when building predictive models, the forecaster uses training data as evidence to manage the tradeoff between fit-to-data and complexity, in order to deal with the problems of over- and under-fitting the data. To avoid the first problem, modelers use a size rule, and then maximize the fit to avoid the latter problem. Some methods are more optimal than others, but the basic idea remains the same. With this common epistemic ground in place, we can make some headway into evaluating the other three reasons for denying climate change.

4.1 Data, Evidence, and Proxies

The more sophisticated objection that is hiding here is one about proxies for historical climate. Since climate science extends well beyond the 135 year record we have of the weather, climate scientists must use other proxies for identifying past climate events, such as yearly rainfall, average temperature, sea levels, the amount of carbon in the atmosphere, and so on. These proxies include tree rings, ice cores, and the like. Recall that climate models will be "trained" and tested against this data; so, if someone were to find fault with these proxies, they would be well within their epistemic rights to lower their confidence in the model's predictions.

Our model selection framework suggests that even this more sophisticated skepticism requires greater sophistication. Notice that a proxy is just a hypothesis (or model) about a relationship between the proxy and the event in question. Ice cores are proxies for historic temperatures, for example. But, we can test how well ice cores predict temperatures *in the same way* we can test whether a change in air pressure under certain conditions will produce rain in a certain area. Further, the model selection framework suggests ways to improve our proxies. For example, using a model with more than one parameter to select proxies might in some cases be a good idea. The model selection framework can provide advice for whether adding more parameters will improve our predictions about what the

[18] Consider Patrick J. Michaels and his collaborators. Michaels has made a career out of CC4 denial (Conway and Oreskes 2010). He and his collaborators hold that climate models run too hot (Michaels, Knappenberger, and Davis 2000; Michaels et al. 2002; Michaels and Knappenberger 2013). Their main evidence for this claim is that there are chunks of the record where climate models do not seem to fit well. However, they have yet to use anything like model selection theory to show that the increase in fit will be enough to overwhelm the danger of overfitting for the purposes of prediction.

climate was like in the past.[19] At any rate, CC4 deniers will have to do better than merely suggest that we do not have a written record of temperatures beyond 135 years or so to demonstrate a real problem with proxies.

4.2 Coincidence

Should we buy the claim, advanced by some TV weathercasters, that observed climate events are merely coincidence? To understand why we should be skeptical of this claim, let us consider another problem faced by model selectors.

First of all, consider one of the biggest sucker bets in history – the birthday bet. In a room of about 40 people, if you were to bet even odds that two people in the room have the same birthday, then in the long run you would come out ahead. That is because the event is very likely (about 89%) in that circumstance.[20] So given the probability and the odds generally accepted, one can expect to make a bit of money in the long run. However, why are people willing to give such odds – i.e., what is it about the birthday bet that makes it a sucker bet? Well, when most people consider the bet, they are thinking about the probability that two people will have a particular date for their birthday, which is quite low – for example the chance that two people will have May 15th as a birthday in the same room is about 1%. What's interesting about the birthday bet is that both the hustler and the sucker are correct in their assessments of the probabilities: (i) it is true that it is far more likely than not that at least two people out of 40 will have the same birthday; and (ii) it is highly unlikely that two people out of 40 will both have May 15th as their shared birthday. The sucker's first mistake is to think that (ii) properly describes the terms of the bet rather than (i).

Taking a page from Elliott Sober's (2012) discussion of coincidence, let us now consider the case of Evelyn Adams, who won the New Jersey Lottery not once, but twice. What are the odds of such an occurrence? Well, almost none. So what should we say about the case of Evelyn Adams? Sober considers a couple of typical responses. The naïve, as Sober calls them, reason that the lottery must have been fixed. Sober contrasts the naïve with those he calls the sophisticates, who say that the double win was a virtual certainty, and that it's being Ms. Adams is a mere coincidence. Who's right? The difference between the naïve and sophisticated reasoners in this case seems to be how they describe the event in question. The naïve specify who it was that won, the number of times they won, and the lotteries in question. The sophisticates point out that if the event were described thusly, "someone wins some lottery more than once," and when there is a large number of people playing a large number of lotteries, a double win is a virtual certainty by chance alone. Diaconis and Mosteller (1989) go so far as to say, "with a large enough sample, any outrageous thing is likely to happen" (859). So, is it the case that the naïve reasoners have made the same error as the sucker who agrees to wager on the birthday bet?

---

[19] It might sound strange to some readers to talk about predicting past events. However, there is a predictive element insofar as we are inferring from the observed to the unobserved.

[20] The probability that n people have different birthdays can be calculated by solving for $365!/365^n\,(365 - n)!$. When n = 40, the resulting probability is about 11%. That means that the probability that at least two people (out of 40) having the same birthday is about 89%.

The case of Evelyn Adams might seem at first blush to be very similar to the birthday bet – a mere difference in description – but it isn't. With the birthday bet, the hustler and the sucker differ over their description of the terms of the bet. As a result of that difference, the sucker and the hustler are considering two distinct explananda – whether some two persons have some birthday in common (high probability) vs. whether two persons have a particular birthday in common (low probability). With the case of Evelyn Adams, however, the explanandum is that Evelyn Adams won two lotteries. Where the naïve and sophisticated reasoners differ is over the appropriate explanans: for pure naïve reasoners, *everything* (including the case of Ms. Adams) is causally connected and for pure sophisticated reasoners, *everything* (including the case of Ms. Adams) is a coincidence. This is a substantive difference. So, should we be naïve or sophisticated reasoners? Well, that depends on the case and the size of the data set one uses as evidence.

A model selection framework, such as AIC, can provide some guidance (Sober 2012). We can treat the ideas expressed by the naïve and the sophisticates as hypotheses. The naïve argue that the double win suggests a causal connection between the two wins (i.e., the lottery was rigged), and the sophisticates argue that the double win is a mere coincidence, completely expected if there have been enough plays, players, and lotteries.

Here are the two hypotheses about Ms. Adams' double-win:

> (FAIR) The fair lottery hypothesis suggests that each individual lottery is fair and independent. Thus, for a ticket t purchased for lottery i (where $1 \leq i \leq r$), the Pr(t wins | t was purchased in lottery i) = αi. Here there is one adjustable parameter for each lottery 1 through r. The model is a conjunction of all lotteries, and the correct fitted model determines αi, the chance of anyone winning that lottery, independently for each lottery i. So, this model has r adjustable parameters – one for each lottery.
>
> (RIGGED) If we think that the lottery is fixed for some player to win a particular lottery, then the model becomes much more complicated. Now the hypothesis must suggest that for a ticket t purchased by player j for lottery i, the Pr (t wins | t was purchased by player j for lottery i) = α_{ij}. This model is a conjunction of all lottery-player pairs, and the correct fitted model determines αij, the chance of a ticket belonging to a winning lottery-player pair. So, this model has a much larger number of parameters – one for each lottery-player pair!

Like VAR1 and VAR2 from Section 3.2, FAIR is actually a special case of RIGGED. If we set j to 0 in RIGGED, we get FAIR. So, FAIR can never do better than RIGGED in terms of its fit-to-data. This seems to suggest that the naïve reasoner's causal connection hypothesis is in good shape. However, a good model selection framework suggests that one should accept the more complex model over the simpler model only if the difference in terms of relative fit-to-data overcomes the penalty for additional parameters. Since RIGGED has *exponentially* more adjustable parameters than FAIR, RIGGED would have to fit the data *far* better than FAIR to be selected. Since the training data only contains two data points (both of Ms. Adams' wins), it is highly unlikely that RIGGED's fit-to-data will overcome the penalty of adding additional parameters. If on the other hand, Ms. Adams had won several more lotteries, there would be some n at which RIGGED might do so. Nonetheless, she did not win another lottery nor are there reports of several other multi-wins, so from a model selection standpoint, FAIR might be preferred to RIGGED, given the current data set.

This should sound familiar. Climate scientists suggest that the observed rise in global temperatures year after year must be causally connected (CONNECTED) as opposed to CC4 deniers who might suggest that such rises in global temperatures are nothing more than mere coincidence (COINCIDENCE) resulting from a huge and ever changing system. Could CC4 deniers make use of the model selection framework to make a similar simplicity argument against climate change, insisting that COINCIDENCE is the better model? It is certainly not a slam-dunk. There is far more data with respect to climate change than there is with respect to New Jersey lotteries. As a data set grows arbitrarily large, more complex models (like CONNECTED) tend to fit the data better and better. As a result, there will be some point (if we continue to add data) at which the more complex model's fit-to-data will overcome the penalty attributed due to additional parameters (Goldsby 2013). Simpler models typically do better only when the data set is small relative to the number of parameters used in the models being compared. Nonetheless, 2014 was the warmest year on record and the ten warmest years have all occurred since 1998 (NOAA National Climactic Data Center 2014). Furthermore, the first quarter of 2015 is the warmest first quarter on record (NOAA National Climactic Data Center 2015). In light of these facts, the COINCIDENCE model is at least strained.

However, we have taken great pains to insist that model selection ought not be an *a priori* enterprise – i.e., which model we use ought to be determined at least in part by the historical data. That means that it is not necessarily a slam-dunk for CONNECTED either. The answer is to be determined by the data – meaning that one must do the climate science to distinguish between the two. As we've pointed out in the previous section this is what is missing from the climate change denial literature. Still, there is a presumption in favor of simpler models when considering how well the model will predict new data. More complex models must demonstrate a much greater fit to justify their use in predictive enterprises according to the model selection framework. Nonetheless, given the warming trend that has been observed, it seems possible (if not likely) that the CONNECTED has met this benchmark but we leave it here as an open question.

All that being said, there is an epistemically noteworthy consideration at issue here. Model selection methods are about estimating predictive success, not necessarily about truth (Forster and Sober 1994; 2010). In fact, models that are known to be false might be great predictors. So, winning in the model selection game is not decisive about whether the model or accompanying theory is true; it only tells us how much confidence we should have that the model will make good predictions. This is not to take back everything we have said about a model selection argument against CC4 deniers: using model selection theory we ***can*** determine whether the CC4 deniers, and their mere coincidence model is a better predictor – if we do the science and the statistical analysis. Much of the climate change debate is about prediction – e.g., what will the climate be like in 2100? However, there may be more to the fight than mere prediction. For example, suppose (contrary to the data and statistics) that the model selection theory favored RIGGED over FAIR. We might still prefer FAIR over RIGGED when we consider the fact that Evelyn Adams is currently destitute, which seems inconsistent with the claim that lottery was rigged in her favor. In other words, we may impose a higher standard on the climate change issue –

one of truth. If we are to hold ourselves to this higher standard, then our models must *at least* cohere with other known facts, such as the fact that there are anthropogenic effects on the carbon cycle. COINCIDENCE seems to imply that there is no anthropogenic effect on the carbon cycle and that carbon has no meaningful effect on the climate. In this respect, CONNECTED certainly outperforms COINCIDENCE.

4.3 Conspiracy

Many CC4 deniers are also conspiracy theorists – like Crichton (2005), Inhofe (2012), and even Michaels (2015). Conspiracy theories are notoriously difficult to rebut, as Gardiner points out. That is because all well-designed conspiracy theories adopt something like the following claim:

> (CS) The conspirators have concocted a cover story designed to mislead non-conspirators from the truth, and any "evidence" that seems inconsistent with the conspiracy theory is actually a part of the cover story.

The general problem with rebutting such well-designed conspiracy theories is that no piece of evidence seems to count against the theory. To see why, suppose one holds that there is some conspiracy X, and a new piece of evidence *E* is discovered. If *E* is at least consistent with X, then it is accepted as "real" evidence. If, on the other hand, *E* is inconsistent with X, then according to CS it is considered to be "misleading" evidence designed to conceal the "truth." It should be clear, given CS, that well-designed conspiracy theories are untestable because there is no piece of evidence that might distinguish the theory from its rival. In fact, evidence that seems to support a rival theory over the conspiracy theory is often seen as further "evidence" supporting CS and by extension the conspiracy theory itself.

Given this feature of conspiracy theories, one might think that attempting to rebut a conspiracy theory is a fool's errand. Of course, it is reasonable to criticize such a claim on the grounds that it makes the theory untestable, which should give pause to anyone with a respect for science. However, CS dismisses such criticisms as being intentionally or unintentionally complicit with the conspiracy. Consequently, CS seems to deny any chance of achieving common epistemic ground between someone who is a conspiracy theorist and someone who is not. Without that common ground, it is hard to imagine how progress toward the truth can be made. For this reason, we make no attempt to disprove the conspiracy theories that might motivate CC4 denial, for such rebuttals either will be preaching to the choir or will fall on deaf ears. Instead, our project is more oblique. Using a model selection framework, we will show that even conspiracy theorists should be cautious about the predictions of their theories and models. Since the disagreement largely involves what the future state of the climate will be, the predictive ability of one's models is important. Thus, even a conspiracy theorist must take our criticisms seriously.

According to climate conspiracy theories, the reports of the IPCC and the data provided by climate scientists are part of the cover story. The "real" story, according to these conspiracy theories is that the scientists have concocted the whole thing. This means that climate conspiracy theories are not just one theory, but a conjunction of two theories: (a) a theory about the behavior of scientists (to include their methods and motives) and (b) a theory about what is "really" going on with the climate. Before moving on we would like to say a few words about the (b) part of those theories.

Consistency is perhaps the only epistemic virtue to which conspiracy theorist and non-conspiracy theorists agree. To even have a shot at consistency, climate conspiracy theories must account for the recent climate events that have been observed – such as the record-setting global temperatures of the last 15 years or so (NOAA National Climactic Data Center 2014; 2015). There are two options available to the conspiracy theorist. The first option is to accept that the observed climate events are real, but deny that said events are indicative of significant anthropogenic climate change – either because the change is minimal, or it is natural, or that it is merely coincidence. If the conspiracy theorists were to avail themselves of this option, they would still have to address the concerns that we've raised in the previous sections – i.e. they would have to show that alternative models are more likely to be better predictors than CONNECTED. That is to say that they must show why their alternative model would be more likely to provide accurate climate predictions, independently of how well the (a) part of the theory fares. On the other hand, a conspiracy theorist could deny that the events have even occurred (or in a less plausible case, they might maintain that they are real but were manufactured by the conspirators). Availing oneself of this option would be tantamount to maintaining that the observed climate events were merely apparent. In which case, the conspiracy theorist would account for the "apparent" climate events by appealing to the behavior of the scientists. Interestingly enough, this option does avoid many of the issues we raised earlier, but at the cost of adopting a decidedly skeptical stance regarding empirical observations. Such a move would depend exclusively on how well the (a) part of the conspiracy theory fares.

In order to evaluate the various conspiracy theories about the behavior of scientists, let us first consider the contrasting non-conspiracy model:

> (HONEST) The observed climate events are real, and scientists earnestly present their most evidentially justified understanding of those events.

The HONEST model assumes that scientists are as they present themselves. Each scientist gathers evidence and constructs her theories in order to better understand the world. It is consistent with the HONEST model that scientists sometimes make mistakes, but it assumes that said mistakes are due to the limitations of the evidence and the cognitive abilities of the scientists and not some hidden agenda designed to mislead. It further assumes that when scientists provide their reports that they honestly present the theory that to their minds is best supported by the evidence. From a model selection perspective, HONEST has adjustable parameters for each individual scientist's relation to the world and an error term. Admittedly that is a lot of parameters, but the data set is large, and as we will discover, HONEST is the simplest model by far.

Let us now turn to the conspiracy models. Here are three possibilities:

> (PRODUCE) The observed climate events are real, but scientists have actually produced those events to make it seem, contrary to the facts, that significant anthropogenic climate change is happening in order to promote some agenda.

> (TAMPER) The observed climate events are either (i) merely apparent, (ii) insignificant, (iii) natural, or (iv) the result of mere coincidence, but scientists have tampered with the data to

make it seem like significant anthropogenic climate change is happening in order to promote some agenda.

(MISREPRESENT) The observed climate events are either (i) insignificant, (ii) natural, or (iii) the result of mere coincidence, but the scientists misrepresent those events to make it seem like significant anthropogenic climate change is happening in order to promote some agenda.

Starting with the first conspiracy model, PRODUCE is a little bizarre. We doubt that many CC4 deniers – especially TV weathercasters – hold this view. Nonetheless, suppose some group of people is out to convince the world there is climate change – just like the plot of Michael Crichton's novel. To flirt with science fiction, let us suppose that this nefarious cabal might orchestrate this by actually causing some severe climate event, and that this can be done on an isolated basis, without causing permanent climate change.

Of the three conspiracy models PRODUCE is the only one that holds the observed climate events are *the result* of the scientists' behavior. As such, this model directly competes with both COINCIDENCE and CONNECTED rather than HONEST. How will such a model fare relative to COINCIDENCE? As we have already seen, it will face a tremendous uphill battle. Notice that this hypothesis will be far more complicated in terms of the number of adjustable parameters than even CONNECTED, as there will have to be additional parameters for all the additional variables (such as the individuals planning the chaos, the means of affecting the chaos in question, the cover ups required to keep people from knowing about these climate altering events, and so on). While this may fit the data well, it will have to fit the data *exponentially* better than either COINCIDENCE or CONNECTED to overcome the penalty for additional parameters. As such, we should have little faith in its predictive abilities. This is a common problem with many conspiracy theories; they typically do a good job *accommodating* the old data, but they fare poorly at *predicting* new data. Model selection theory provides an account of why that is the case.

A similar point can be made against TAMPER. Of course, TAMPER says that climate scientists intentionally cook the data in order to make it look like climate change is happening. Unlike PRODUCE, TAMPER models the behavior of the scientists and not the causes of climate events. Thus, we should compare TAMPER to HONEST. Recall that HONEST has an adjustable parameter for each individual scientist's relation to the world, as does TAMPER; so in that regard it is a wash. However, TAMPER also adds parameters for the scientist's motives, the conspirator's cover story, and the methods used to tamper with the data. Additionally, TAMPER must add parameters for the ignorance, indifference, or collusion of institutions (such as peer review, hiring practices, etc.) designed to hold scientists to a particular level of rigor in data collection. The list goes on and on. As such, TAMPER is many times more complex than HONEST. Consequently, that means that TAMPER would have to fit the training data astronomically better to overcome the penalty for complexity. It is plausible to think that HONEST already fits the data quite well. If it didn't, then the conspirators wouldn't stand much of a chance of exploiting our misplaced trust. This means that there is not much room for TAMPER to fit the data better. As a result, this means that it is highly unlikely that TAMPER's fit-to-data will overcome the penalty assessed for complexity. Thus, model selection theory would likely rule that HONEST is a better predictor of the behavior of scientists than TAMPER.

MISREPRESENT is a bit trickier. According to this model, climate scientists have presented us with a sucker's bet – much like the birthday bet. The idea is that by describing the climate events in a certain way, it will make anthropogenic climate change seem very likely, but if the events were "properly" described, then we would see that anthropogenic climate change is far less likely. It is easy to grasp the idea behind this model, but it is far more difficult to cash it out, because it is unclear as to exactly how climate scientists are supposed to be doing this. Furthermore, unlike the birthday bet, the explanandum or object of inquiry is the same, regardless of description. So MISREPRESENT cannot be analyzed in the same way.

One possibility, for understanding MISREPRESENT is that it might be saying that climate scientists are "spinning" the data in their descriptions to make it appear that assenting to CC4 is more reasonable than it actually is. Presumably, according to this line of reasoning, that spin is little more than rhetoric designed to exploit our cognitive biases in order to make it appear that anthropogenic climate change is actually happening. It is hardly controversial that rhetoric can be used to mislead people. Plato and Socrates cautioned against it and relatively recent cognitive science has shown just how deep our biases run (Kahneman and Tversky 1974; Kahneman 2011). However, various statistical tools, including model selection theory, can help us overcome these biases in order to see through such spin. Here's the upshot: model selection theory looks at the raw data and how well the model fits that data. How that data is described when reported makes no difference to model selection tools. It's not how we "feel" about the data or even what we *believe* to be true that determines whether a model is more likely to be a better predictor. What determines whether a model is likely to make good predictions is how well the model fits the data and how many adjustable parameters it has relative to the data set. As such, spin will gain little traction when we use the models selection framework to evaluate our models. Therefore, the task for CC4 deniers who hold this view, *once again*, is to show how an alternative model is more likely to be a better predictor using model selection tools.

MISREPRESENT might also be saying that the description of events provided by climate scientists misleads us by omitting relevant data. This is a serious epistemic concern, as it would constitute a violation of the *principle of total evidence*. It is a truism that if one cherry picks the data one can make it look like it supports nearly any proposition. Thus, such practices ought not be a part of scientific inquiry. We have two things we would like to say against this charge. First of all, if we understand MISREPRESENT in this way, it begins to look a lot like TAMPER in terms of complexity, because such a model have to include parameters for how the conspirators could avoid the various review institutions and practices designed to prevent cherry-picking the data. As such, it is highly unlikely that MISREPRESENT will do better than HONEST in a model selection context. Secondly, the claim is false. Nonetheless, this perception that climate scientists are ignoring or omitting data is strong in the popular imagination of CC4 denial. In fact, one climate change denial book, *The Mad, Mad, Mad World of*

Climatism, published by the Heartland Institute,[21] goes so far as to suggest that climate scientists are trying to rewrite history by ignoring the Medieval Climate Anomaly (MCA) and the Little Ice Age (LIA). Since the MCA happened between 850 and 1250 CE and the LIA happened between 1450 and 1850 CE, it should be noted that both events occurred prior to the weather record. That means it is hard to push the claim and criticize proxy data as CC4 deniers are fond of doing. Additionally, it is demonstrably false that climate scientists ignore these periods, as a chapter of the IPCC AR 5 is dedicated to dealing with it (2013, Ch. 5).

What makes these claims ironic is that it is contrarian climate science (the 3%) that has demonstrated cherry-picking practices. Michaels and Knappenberger (2013) is one example. They look at the most recent 15 years to make their claim that climate models are running too hot. While the IPCC (2013) does not directly cite the work of Michaels and Knappenberger, it addresses this claim by pointing out that the conclusion only follows if one restricts the training data set to the last 15 years (61-3). So, while we agree that cherry-picking data is bad, since it violates the principle of total evidence, it is far from clear that main-stream climate scientists are doing so. Furthermore, MISREPRESENT scores poorly against HONEST using model selection theory, and as such its predictions are untrustworthy.

9. Conclusion

While we think CC4 is true, and best supported by the evidence, we have not made that case here. Our conclusion is more guarded. We merely claim that CC4 deniers have not given us enough evidence to deny the truth of CC4. That is not to say that CC1-3 are unassailable. Like all scientific claims they are based on fallible evidence, and if evidence were to be presented contrary to those claims, one must adjust one's beliefs accordingly. Thus, it is possible that a case could be made to show that CC1-3 are unreasonable. In fact, we have provided a road map for anyone who wishes to do so. However, we have shown that no steps have been taken in that direction, and that the task ahead for the CC4 denier is difficult to say the least. Furthermore, since CC4 depends upon how well our models project the state of the future climate, it seems reasonable to accept the results of climate models as our best available predictions – even if some parts of the models are false.

We have taken on this project to engage a more sophisticated brand of CC4 denier. It is certainly reasonable to suggest that much of our argument does not apply to the more general problem of how to shift public opinion on climate change. In fact, it would be of great interest to see how our argument might be messaged more boradly. Still, our project has been an extension of Gardiner's arguments against CC4 denial in the broader population. Thus, we are content to rest on his argument with respect to less-sophisticated CC4 deniers.

Nonetheless, our argument parallels Gardiner's in many ways. We argue that TV weathercasters must endorse a model selection framework, as their own predictive tools are models. With this common epistemic ground, we have demonstrated that their lines of reasoning in favor of CC4 denial suffer some

[21] The copyright restrictions of the book explicitly forbid direct quotation except for purposes of review, unless the rights holder permits it. Since this article is not a review of the book, and we have not asked for permission, we will not do so.

difficulties. First, denying proxy data without evaluating the evidence for the proxy is unmotivated. Second, the idea that the climate events that climate “alarmists” cite as evidence for anthropogenic climate change are mere coincidence has, to its credit, a simple model to support it. However, that model would have to be estimated to be extremely predicatively successful in light of the fact that there is settled science that demonstrates there are some causal connections between emissions and climate. Further, none have developed COINCIDENCE-like models that have been demonstrated to have a higher estimated predictive accuracy than the models used by climate scientists. Finally, we consider the conspiracy theories forwarded by CC4 deniers. Again, the model selection framework provides the tools for an argument against the conspiracy theorist – not by showing the conspiracy theory is false – which is notoriously difficult to do given the unfalsifiability of such theories – but by demonstrating that such models are not very likely to predict new data well.

References

Akaike, Hirotugu. 1973. "Information Theory as an Extension of the Maximum Likelihood Principle." In B. Petrov and F. Caski (eds.), *Second International Symposium on Information Theory*. Budapest: Akademiai Kiado, 267-281.

Boeckli, L., A. Brenning, S. Gruber, and J. Noetzli. 2012. "A Statistical Approach to Modelling Permafrost Distribution in the European Alps or Similar Mountain Ranges." *The Cryosphere* 6 (1): 125–40.

Burnham, Kenneth P., and David R. Anderson. 2002. *Model selection and multi-model inference: a practical information-theoretic approach*. New York: Springer.

Cangialose, John and James Franklin. 2012. "2012 National Hurricane Center Forecast Verification Report." *National Hurricane Center*.

Crichton, Michael. 2005. *State of Fear*. New York: Harper Collins.

DelSole, Timothy, and J. Shukla. 2006. "Specification of Wintertime North American Surface Temperature." *Journal of Climate* 19 (12): 2691–2716.

Diaconis, P. and Mosteller, F. (1989): "Methods of Studying Coincidences." J. Amer. Statist. Assoc. 84: 853-861.

Forster, Malcolm. 2000. "Key Concepts in Models Selection: Performance and Generalizability." *Journal of Mathematical Psychology* 44: 205-231.

---. 2001. "The New Science of Simplicity." In *Simplicity, Inference, and Modelling*, ed. Arnold Zellner, H. A. Keuzenkamp, and Michael McAleer, 83-119. Cambridge: Cambridge University Press.

Forster, Malcolm and Elliott Sober. 1994. "How to Tell when Simpler, More Unified, or Less *Ad Hoc* Theories will Provide More Accurate Predictions." *British Journal for the Philosophy of Science*, 45: 1-35.

--- 2010. "AIC Scores as Evidence – a Bayesian Interpretation." In Malcolm Forster and Prasanta Bandyopadhyay (eds.), *The Philosophy of Statistics.* Dordrecht: Kluwer.

Gardiner, Stephen M. 2011. *A Perfect Moral Storm: the Ethical Tragedy of Climate Change*. Oxford: Oxford University Press.

Gent, Peter. 2012. "Coupled Climate and Earth System Models" In *Climate Change Modeling Methodology*, ed. Phillip J. Rasch, 5-30. New York: Springer Science + Business Media.

Goldsby, Michael. 2013. "The 'Structure' of the 'Strategy': Looking at the Matthewson-Weisberg Tradeoff and its Justificatory Role for the Multiple-Models Approach." *Philosophy of Science,* 80(5): 862-873.

Inhofe, James. 2012. *The Greatest Hoax: How the Global Warming Conspiracy Threatens Your Future*. 1 edition. Washington, D.C: WND Books.

IPCC. 2013. Summary for Policymakers. In T. Stocker, D Qin, G. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & B Midgley (Eds.), *Climate Change 2013: The Physical Science Basis – Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.

IPCC. 2014. Summary for Policymakers. In O. Edenhoffer, R. Pichs-Madruga, Y. Sokona, j. Minx, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schloemer, C. von Stechow, & T. Zwickel (Eds.), *Climate Change 2014: Mitigation of Climate Change – Working Group III Contribution to the Fifth Assessment Report of the Intergovernmental Panel on climate Change*. Cambridge: Cambridge University Press.

Jones, Charles, Leila M. V. Carvalho, R. Wayne Higgins, Duane E. Waliser, and K. E. Schemm. 2004. "A Statistical Forecast Model of Tropical Intraseasonal Convective Anomalies." *Journal of Climate* 17 (11): 2078–95.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus, and Giroux.

Kahneman, Daniel and Amos Tversky. 1974. "Judgment Under Uncertainty: Heuristics and Biases." *Science* 185: 1124-1131.

Maibach, Edward, Kristopher Wilson and James Witte. 2010. "A National Survey of Television Meteorologists about Climate Change: Preliminary Findings." George Mason University. Fairfax, VA: Center for Climate Change Communication.

Maibach, Edward, Sara Cobb, Anthony Leiserowitz, Erin Peters, Vanessa Schweizer, Carole Mandryk, Jim Witte, David Straus, Dan Sklarew, Lee Clarke, Jagadish Thaker, Keith Seitter, Steve Harned, Dennis Todey, Eugene Bierly, Martin Hoerling, David Herring, Rick Bonney, and Heidi Cullen. 2011. "A National Survey of Television Meteorologists about Climate Change: Education." George Mason University. Fairfax, VA: Center for Climate Change Communication.

Maibach, Edward, Raphael Mazzone, Robert Drost, Teresa Myers, Bernadette Woods Placky, Heidi Cullen, Bud Ward, Joe Witte, Keith Seitter, Steve Harned, Katherine Hayhoe, Susan Hassol, Jeff Lazo, Michael Slater, David Herring, Frank Neipold, and Ned Gardner. 2015. "A National Survey of Broadcast Meteorologists About Climate Change: Initial Findings. George Mason University. Fairfax, VA: Center for Climate Change Communication.

Maranzano, Coire J., and R. Krzysztofowicz. 2006. "Bayesian Processor of Output for Probabilistic Quantitative Precipitation Forecasting." In *18th Conference on Probability and Statistics in the Atmospheric Sciences*.

Maslin, Mark. 2014. *Climate Change: A Very Short Introduction*. Oxford: Oxford University Press.

Mauritsen, Thorsten, Bjorn Stevens, Erich Roeckner, Traute Crueger, Monika Esch, Marco Giorgetta, Helmuth Haak, et al. 2012. "Tuning the Climate of a Global Model." *Journal of Advances in Modeling Earth Systems* 4 (3).

Michaels, Patrick J. 2015. "Global Warming Scientists Scrap Real Science, Bow before President Obama Instead." *Cato Institute*. Accessed April 19. http://www.cato.org/publications/commentary/global-warming-scientists-scrap-real-science-bow-president-obama-instead.

Michaels, Patrick J., and Robert C. Balling. 2000. *The Satanic Gases: Clearing the Air About Global Warming*. Cato Institute.

Michaels, Patrick J., and Paul C. Knappenberger. 2013. "Climate Data vs. Climate Models; Why Do Climate Change Assessments Overlook the Difference between the Climate Models They Use and the Empirical Data." *Regulation* 36: 32-36.

Michaels, Patrick J., Paul C. Knappenberger, and Robert E. Davis. 2000. "The Way of Warming." *Regulation* 23 (3): 10-15.

Michaels, Patrick J., Paul C. Knappenberger, Oliver W. Frauenfeld, and Robert E. Davis. 2002. "Revised 21st Century Temperature Projections." *Climate Research* 23: 1-9.

National Hurricane Center. 2009. "Technical Model Summary." *National Oceanic and Atmospheric Administration*.

NOAA National Climatic Data Center. 2014. "State of the Climate: Global Analysis for Annual 2014." published online December 2014, retrieved on April 20, 2015 from http://www.ncdc.noaa.gov/sotc/global/2014/13.

NOAA National Climatic Data Center. 2015. "State of the Climate: Global Analysis for March 2015." published online April 2015, retrieved on April 20, 2015 from http://www.ncdc.noaa.gov/sotc/global/2015/3.

Oreskes, Naomi. 2004. "The Scientific Consensus on Climate Change." *Science* 306 (5702): 1686–1686.

Oreskes, Naomi and Erik Conway. 2010. *Merchants of Doubt: How a Handful of Scientists obscured the Truth on Issues from Tobacco smoke to Global Warming*. New York: Bloomsbury Press.

Parker, Wendy. 2011. "When Climate Models Agree: The Significance of Robust Model Predictions." *Philosophy of Science*, 78 (4): 579-600.

Pew Research Center. 2015. "Trends in Views About Climate Change." *Pew Research Center's Internet & American Life Project*. Accessed April 19. http://www.pewinternet.org/2015/01/29/public-and-scientists-views-on-science-and-society/pi_2015-01-29_science-and-society-03-15/.

Popper, Karl. 1959. *The Logic of Scientific Discovery*. New York: Basic Books, Inc.

Sargent, Robert G. 2005. "Verification and Validation of Simulation Models." In M. Kuhl, N. Steiger, F. Armstrong and J. Joines (eds.) *Proceedings of the 2005 Winter Simulation Conference*. Piscataway, NJ: IEEE, 130–43.

Skyrms, Brian. 1984. *Pragmatics and Empiricism*. New Haven: Yale University Press.

Sober, Elliott. 2008. *Evidence and Evolution*. Cambridge: Cambridge University Press.

---. 2012. "Coincidences and How to Reason About Them." In *EPSA Philosophy of Science: Amsterdam 2009*, edited by Henk W. de Regt, Stephan Hartmann, and Samir Okasha, 355–74. The European Philosophy of Science Association Proceedings 1. Springer Netherlands.

Stenhouse, Neil, Edward Maibach, Sara Cobb, Ray Ban, Andres Bleistein, Paul Croft, Eugene Bierly, Keith Seitter, Gary Rasmussen, and Anthony Leiserowitz. 2014. "Meteorologists' Views about Global Warming: A survey of American Meterological Society Professional Members. *Bulletin of the American Meteorological Society* 95 (7): 1029–40.

Vermeer, Martin, and Stefan Rahmstorf. 2009. "Global Sea Level Linked to Global Temperature." *Proceedings of the National Academy of Sciences* 106 (51): 21527–32.

Wallace, John M. and Peter V. Hobbs. 2006. *Atmospheric Science, Second Edition: An Introductory Survey*. Burlington, MA: Academic Press.