

Scientific Forecasts of Climate Change? (for discussion)

Tom Fiddaman

[tom@metasd.com](mailto:tom@metasd.com)

Posted to the JDM discussion list at

<http://www.sjdm.org/mail-archive/jdm-society/2008-March/003453.html>

with comments open at

<http://blog.metasd.com/2008/03/07/evidence-on-climate-predictions/>

Scott Armstrong and Kesten Green contended recently on jdm that there are no valid forecasts of dangerous anthropogenic global warming ( <http://www.sjdm.org/mail-archive/jdm-society/2008-February/003436.html> ). I reviewed Green & Armstrong (2007) at [forecastingprinciples.org](http://forecastingprinciples.org) to see how they arrive at that conclusion. ( [http://www.forecastingprinciples.com/Public\\_Policy/WarmAudit31.pdf](http://www.forecastingprinciples.com/Public_Policy/WarmAudit31.pdf) )

To support the claim that there are no valid forecasts, the authors would have to show that:

1. Chapter 8 of the IPCC Fourth Assessment WG1 report ( <http://www.ipcc.ch/ipccreports/ar4-wg1.htm> ) shows that climate scientists do not follow "evidence-based forecasting" methods, as defined by Green & Armstrong (G&A).
2. The underlying models and processes summarized in Chapter 8 also do not follow proper methods.
3. It is impossible to create a valid forecast without following the authors' methods.

With respect to 1, I think it is fair to grant that Chapter 8 is weak on model validation details. However, it does not follow that models are invalid, especially because G&A make errors in their attributions of performance against the various principles. They apparently did not examine the primary climate literature in any detail, and clearly missed an enormous amount of relevant information, leading to serious misconceptions about models and their application; hence they can make no claim about 2. With respect to 3, the authors do not establish whether climate is a problem domain where models work or chaos reigns, and do not demonstrate viable alternative forecasts based on forecasting principles.

A demonstration of an actual defect in a model or a forecast would be much more convincing than any of the above. The authors include a variety of anecdotes indicating structural omissions and deficiencies in models. However, they provide no evidence that any of these are important or preclude forecasting. The authors did not directly examine models, data, or model output. Had they done so, they would have discovered that, over the last two decades, data supports early climate model predictions and rejects the authors' suggested no-change alternative hypothesis.

That is not to say that models are perfect. They are not, but this fact is already well-documented (just point Google scholar to "model intercomparison project" for examples). Many efforts have been made to assess the uncertainty around models through a variety of means, including testing of the effects of arbitrary omitted feedbacks on outcomes. An assessment of the modeling process is a useful check on such assessments, but only if it reflects the modeling process accurately.

Scientific Forecasts of Climate Change? (for discussion)

Tom Fiddaman

[tom@metasd.com](mailto:tom@metasd.com)

One of the A&G's first forecasting principles is "Make sure forecasts are independent of politics." The authors should take this one to heart. They have precluded thoughtful communication with climate scientists by creating a publicity circus around their paper at <http://theclimatebet.com/>. Their site plugs the Heartland Institute's climate change conference - a veneer of science over a media event (see [http://www.nytimes.com/2008/03/04/science/earth/04climate.html?\\_r=1&oref=slogin](http://www.nytimes.com/2008/03/04/science/earth/04climate.html?_r=1&oref=slogin) and <http://www.realclimate.org/index.php/archives/2008/01/what-if-you-held-a-conference-and-no-real-scientists-came/>). The culmination of all this wisdom is a bet, challenging Al Gore or any other taker to predict temperatures at 10 weather stations over 10 years (<http://theclimatebet.com/2007/06/16/a-global-warming-challenge/>). For all their talk of data quality, use of robust methods, etc., the authors are surely aware that this is a pitifully small sample, guaranteed to hide the signal in the noise. I can't help but think that, if their conclusions were sound and they wished to be taken seriously by scientists, they would have offered reasonable terms.

I believe it is valuable for researchers to critique decision processes outside their own discipline. However, in such cases the researcher is obligated to learn enough about the subject to make informed judgments. This paper does not meet that burden.

Further detailed observations on Armstrong & Green (2007):

"Forecasting" is a misleading term, at least in general usage. Forecasting implies the mindset, "If we know what will happen, we'll know what to do." In reality, we seldom know what will happen, partly due to irreducible uncertainties, but also in part because action we would take based on that knowledge would change the outcome.

What we really want is an operational understanding of the transfer function between decisions and outcomes we care about (e.g. between a carbon tax and welfare) together with an appreciation of the quality of that understanding. We would use that understanding, not to make point predictions, but to devise control strategies that guide what we can, and hedge against what we can't. In many cases it is possible to intelligently choose strategies in spite of great uncertainty, just as it is possible to sail to Hawaii without knowing which way the wind is going to blow.

Regardless of the objective, the forecasting principles described in the paper and at [audit.forecastingprinciples.com](http://audit.forecastingprinciples.com) strike me as useful. Most resonate with my own experience building mathematical models for policy and strategy, though one could argue as to whether and how they span the space of necessary and sufficient conditions for quality. With respect to models, for example, I do not see principles requiring tests of dimensional consistency and adherence to conservation laws, though those are essential tests, even in nonphysical models.

The problem with the principles is how they have been applied to arrive at the conclusion, that "we have been unable to find a scientific forecast to support the currently

## Scientific Forecasts of Climate Change? (for discussion)

Tom Fiddaman

[tom@metasd.com](mailto:tom@metasd.com)

widespread belief in ‘global warming.’” The primary basis for this statement is an analysis of Chapter 8 of the IPCC Fourth Assessment by scoring it against the 140 principles. There are serious problems both with the very idea of the scoring process, and with the details of the ratings themselves, as I will describe.

First, the IPCC report is a synthesis document, not a primary research article. One would not expect its methods section to describe every test of model validation actually performed in the process of creating the various source articles. Though the authors spent only one hour each rating the document, they were able to identify 72 “clear violations” of principles. One would think it necessary to further examine each violation to determine whether it arises as a fault of the primary research, or merely as a necessary omission for brevity. Yet the authors evidently failed to do so thoroughly – they cite very little primary scientific literature, and grumble about the 788 references provided in the report (page 1007).

Some of the “clear violations” and “apparent violations” are not even things that bear on the technical validity of climate science. “Test the client’s understanding of the methods” and “Obtain decision makers’ agreement on methods” are good advice for a consulting engagement, and perhaps imply a level of communication with decision makers that would be good for society, but these have little bearing on whether sub-grid-scale cloud dynamics are properly parameterized and lead to accurate predictions.

Of greater concern as I read the list is that many of the alleged violations appear to be a consequence of the brevity of AR4 Chapter 8 (compared to the climate literature) and the authors’ limited knowledge of that literature. For many of the items cited, I can think of counterexamples in the literature, and I am not a climate scientist. The four items identified by the authors as essential, “so important that any forecasting process that does not adhere to them cannot produce valid forecasts,” provide good examples.

Consider whether the events or series can be forecasted

“This principle refers to whether a forecasting method can be used that would do better than a naïve method. A common naïve method is to assume that things will not change.” (page 1011). The authors claim that uncertainty about global mean temperatures, the complexity of climate, and known model defects make the “no-change” model appropriate. There are three problems with their conclusion.

First, the no-change model can already be rejected on empirical grounds. A no-change forecast through the present, based on data prior to, say, 1988, would have performed poorly, as can be seen by looking at various global temperature series, e.g., <http://data.giss.nasa.gov/gistemp/graphs/Fig.A2.lrg.gif> Other variants of the no-change hypothesis also fail. For example, no-change in atmospheric water vapor content is inconsistent with data. <http://www.sciencemag.org/cgi/content/abstract/310/5749/841>

## Scientific Forecasts of Climate Change? (for discussion)

Tom Fiddaman

[tom@metasd.com](mailto:tom@metasd.com)

Second, the authors rely on Carter (2007), who says, “the slope and magnitude of temperature trends inferred from time-series data depend upon the choice of data end points. .... Accurate direct measurements of tropospheric global average temperature have only been available since 1979, and they show no evidence for greenhouse warming. Surface thermometer data, though flawed, also show temperature stasis since 1998” These statements are misleading. One outlier year does not create “stasis”. Tropospheric temperatures do show a clear trend, as do stratospheric temperatures, consistent with models (see [http://www.remss.com/msu/msu\\_data\\_description.html#figures](http://www.remss.com/msu/msu_data_description.html#figures) last panel, for example).

Third, one would think that an assessment of “forecastability” might look at actual past forecasts, but the authors apparently did not do so. A little effort would have led to the figures in the back of the 1990 Scientific Assessment or to Hansen, J. Global Climate Changes as Forecast by Goddard Institute for Space Studies 3-Dimensional Model, *Journal of Geophysical Research – Atmospheres* 93:9341 (1988). The latter is significant here because it has proven more reasonable than no-change, and because skeptic Patrick Michaels inappropriately cherry-picked trajectories from that paper in an attempt to discredit the work. Michaels is the editor *Shattered Consensus*, a skeptical book cited four times in the Armstrong & Green paper (vs. a total of 5 citations for primary climate literature).

### Keep forecasting methods simple

The corollary to this simple-minded rule should be, “but not too simple.” When the null model is rejected, one should seek alternatives. The next simplest model in the case of climate involves the radiative effects of greenhouse gases, which is well-established by physics. Put an extra blanket on the bed and it will be warmer under the covers. That too is insufficient to explain spatial and temporal weather patterns, for which it is necessary to include water vapor dynamics.

I find it inconsistent that the authors, in some places, criticize GCMs for their complexity, yet in others, suggest that they are invalid because they neglect a variety of 2<sup>nd</sup> and 3<sup>rd</sup> order effects. Complexity is not automatically a problem for models. Complexity is a problem when it involves speculative relationships combined with paucity of data such that aggregate results cannot be verified. Climate models use many experimentally testable relationships and enormous amounts of data (not just global means), as well as some more uncertain parameterized effects (sub-grid-scale clouds, for example). The authors fail to establish that the parameterized effects lead to overwhelming uncertainty.

Nor are all climate models complex. Models exist at all levels of aggregation, down to univariate 1<sup>st</sup> order. Simple energy balance models are consistent with the findings of much larger models. See for example <http://www.ncdc.noaa.gov/paleo/pubs/crowley.html>

Scientific Forecasts of Climate Change? (for discussion)

Tom Fiddaman

[tom@metasd.com](mailto:tom@metasd.com)

Do not use fit to develop the model

The authors write that, “It was not clear to us to what extent the models described in Chapter 8 (or in Chapter 9 by Hegerl et al. 2007) are either based on, or have been tested against, sound empirical data.” Yet they are somehow able to determine with confidence that modelers have committed a “clear violation” by using fit to develop their models.

It should be obvious that, as with simplicity above, not all use of fit is bad – otherwise science would have to forego measurement entirely. Certainly fit can be misused when relationships are arbitrarily added to models to improve it (as when mindlessly adding explanatory variables to a regression). The authors provide no evidence that climate modelers have done this.

Indeed, evidence suggests otherwise. For a long time, models were inconsistent with tropospheric temperatures from the UAH analysis of microwave sounding unit data; skeptics trumpeted this apparent deficiency loudly. However, modelers did not change their models to resolve the discrepancy, because there was no physical basis for doing so. Recently, it has turned out that the errors were in UAH’s processing of the data, not in the models.

Here and elsewhere, the paper repeats the common notion that larger models are easier to fit to data arbitrarily. This is true for simple linear models, but not in general for large scale nonlinear dynamic models. The latter generally respond to relatively few tuning parameters (counterintuitively at that) and require conformity to much more data, making it extremely difficult to tune for a desired outcome. Large scale climate models are simply too slow and too noisy to fit by hill climbing; most of their parameters are established a priori.

Use out-of-sample (ex ante) error measures

Again, the fact that Chapter 8 of the IPCC report does not discuss ex ante forecast results does not constitute an indictment of climate science. Had the authors looked further, they could have discovered that (a) models are not slavishly fit to data (as above), and (b) models are routinely compared against one another and against extraordinary circumstances, such as volcanic eruptions and alternate earth orbital and continental configurations (in an attempt to replicate paleoclimates). These are mentioned in Chapter 8 (page 600-601) but apparently went unnoticed because they are not explicitly described as ex ante.

Given that the four headline “clear violations” appear to be misconceptions based on an incomplete view of the climate literature, it is hard to see how one can expect better of the others, and thereby conclude that there are no valid climate predictions. The authors’ line of questioning about the process by which predictions are made undoubtedly has value, but ought to be pursued with a deeper understanding of the problem domain.

## Scientific Forecasts of Climate Change? (for discussion)

Tom Fiddaman

[tom@metasd.com](mailto:tom@metasd.com)

Throughout the paper, the authors take a dim view of models, seemingly equating them with any kind of structured or unstructured thinking, “Climate models are, in effect, mathematical ways for the experts to express their opinions.” Models are actually much more than that; they permit assimilation of data and accurate predictions of the behavioral consequences of structural assumptions. Their performance can be tracked over time, so that they may be improved or discarded as appropriate. The authors go on, “To our knowledge, there is no empirical evidence to suggest that presenting opinions in mathematical terms rather than in words will contribute to forecast accuracy.” (page 1002). This statement may be true in social system forecasting, but it is obviously not true in general. Only a fool would bet on an astrologer, or even an unaided astronomer, against a computer simulation for predicting the timing of a solar eclipse in a distant future year.

I would argue that predictability is not only a function of process and maturity of understanding; it also depends on the characteristics of the problem domain. Some situations are highly predictable, given sufficient understanding and assistance from formal models. Others are hopelessly unpredictable, even with considerable knowledge, perhaps because they are chaotic (in the formal sense) or because unknown nonstationary phenomena dominate. In between are many intermediate cases, including chaotic and stochastic systems for which we can at least place bounds on behavior. Our choice of models vs. null forecasts should be based on evidence about both the domain and our level of understanding. While this paper raises interesting questions arising from other problem domains, it does little to establish that they are relevant to the climate problem.

There is a whiff of condescension throughout the paper – as if natural scientists were a bunch of superstitious bumbler until Scientific Forecasting came along to show them the light. Perhaps this is the reason no researchers have taken the authors up on their call for alternative ratings (page 1014). The authors report that climate modelers have failed to cite the forecasting literature, and wonder “how scientific forecasting could be conducted without reference to the research literature on how to make forecasts.” (page 1015). In part, I suspect that most scientists do not see their role as forecasting; they see it as the construction of tools (models) for contingent prediction. Thus it is not surprising that they do not cite the forecasting literature. I’d also hazard a guess that, for better or worse, many natural scientists regard social and economic forecasting as irrelevant to their endeavors. Science managed to make useful predictions long before there was a “forecasting principles” literature. It strikes me as the height of arrogance to imagine that natural science could not have independently arrived at some useful understanding about how to apply models to make judgments in complex situations.