Global Warming: Forecasts by Scientists versus Scientific Forecasts*

J. Scott Armstrong†, The Wharton School, University of Pennsylvania

Kesten C. Green, Business and Economic Forecasting Unit, Monash University

Paper prepared for ISF 2007 in New York

June 19, 2007

Abstract

In 2007, a panel of experts established by the World Meteorological Organization and the United Nations Environment Programme issued its updated, Fourth Assessment Report, forecasts. The Intergovernmental Panel on Climate Change's Working Group One Report predicts dramatic and harmful increases in average world temperatures over the next 92 years. We asked, are these forecasts a good basis for developing public policy? Our answer is "no".

Much research on forecasting has shown that experts' predictions are not useful. Rather, policies should be based on forecasts from scientific forecasting methods. We assessed the extent to which long-term forecasts of global average temperatures have been derived using evidence-based forecasting methods. We asked scientists and others involved in forecasting climate change to tell us which scientific articles presented the most credible forecasts. Most of the responses we received (30 out of 51) listed the IPCC Report as the best source. Given that the Report was commissioned at an enormous cost in order to provide policy recommendations to governments, the response should be reassuring. It is not. The forecasts in the Report were not the outcome of scientific procedures. In effect, they present the opinions of scientists transformed by mathematics and obscured by complex writing. We found no references to the primary sources of information on forecasting despite the fact these are easily available in books, articles, and websites. We conducted an audit of Chapter 8 of the IPCC's WG1 Report. We found enough information to make judgments on 89 out of the total of 140 principles. We found that the forecasting procedures that were used violated 72 principles. Many of the violations were, by themselves, critical. We have been unable to identify any scientific forecasts to support global warming. Claims that the Earth will get warmer have no more credence than saying that it will get colder.

*Neither of the authors received funding for this paper.

† Information about J. Scott Armstrong can be found on Wikipedia.

"A trend is a trend,
But the question is, will it bend?
Will it alter is course
Through some unforeseen force
And come to a premature end?"
Alec Cairncross, 1969

Research on forecasting has been conducted since the 1930s. Of particular value are comparative empirical studies to determine which methods are most accurate in given situations. The findings, along with the evidence, were first summarized in Armstrong (1978, 1985). The forecasting principles project, begun in the mid-1990s, summarized knowledge as evidence-based principles (condition-action statements) to provide guidance on which methods to use in a given situation. The project led to the *Principles of Forecasting* handbook (Armstrong 2001), which involved 40 authors (all internationally known experts on forecasting methods) along with 123 reviewers (also leading experts on forecasting methods). The summarizing process alone required a four-year effort.

Efforts have been made to ensure that these principles are easy to find. They have been freely available on forecastingprinciples.com, a site that attracts over 200,000 visits per year. The site is intended to summarize all useful knowledge on forecasting methods and those who disagree are invited to contribute any missing knowledge. There is no other source that provides evidence-based forecasting principles. Over the past few years, if one entered the word forecasting on any of the major search engines, forecastingprinciples.com would appear first among the many millions of sites. The site is often updated and a recent update of evidence on some of the key principles was provided in Armstrong (2006).

Many of the principles go beyond common sense, and some are counter-intuitive. As a result, those who forecast in ignorance of the research literature are unlikely to produce useful predictions. For example, here are some of the well-established generalizations for situations involving long-range forecasts of complex issues where the causal factors are subject to uncertainty (as with climate):

- *Unaided judgmental forecasts by experts have no value.* This applies whether the opinions are expressed by words, spreadsheets, or mathematical models. It also applies regardless of how much scientific evidence is possessed by the experts. Among the reasons for this are:
    a) Complexity: People cannot assess complex relationships through unaided observations.
    b) Coincidence: People confuse correlation with causation.
    c) Feedback: People making judgmental predictions typically do not receive unambiguous feedback they can use to improve their forecasting.
    d) Bias: People have difficulty in obtaining or using evidence that contradicts their initial beliefs. This problem is especially serious for people who view themselves as experts.
- *Agreement among experts is weakly related to accuracy*. This is especially true when the experts communicate with one another and when they work together to solve problems. (As is the case with the IPCC process).
- *Complex models (those involving nonlinearities and interactions) harm accuracy because their errors multiply.* That is, they tend to magnify one another. Ascher (1978), refers to the Club of Rome's 1972 forecasts where, unaware of the research on forecasting, the developers proudly proclaimed, "in our model about 100,000

relationships are stored in the computer." (The first author was aghast not only at the poor methodology in that study, but also at how easy it was to mislead both politicians and the public.) Complex models are also less accurate because they tend to fit randomness, thereby also providing misleading conclusions about prediction intervals. Finally, there are more opportunities for errors to creep into complex models and the errors are difficult to find. Craig, Gadgil, and Koomey (2002) came to similar conclusions in their review of long-term energy forecasts for the US made between 1950 and 1980.

- *Given even modest uncertainty, prediction intervals are enormous*. For example, prediction intervals expand rapidly as time horizons increase so that one is faced with enormous intervals even when trying to forecast a straightforward thing such as automobile sales for General Motors over the next five years.
- *When there is uncertainty in forecasting, forecasts should be conservative*. Uncertainty arises when data contain measurement errors, when the series is unstable, when knowledge about the direction of relationships is uncertain, and when a forecast depends upon forecasts of related (causal) variables. For example, forecasts of no change have been found to be more accurate for annual sales forecasts than trend forecasts when there was substantial uncertainty in the trend lines (e.g., Schnaars & Bavuso 1986). This principle also implies that forecasters reverting to long-term trends when such trends have been firmly established, they do not waver, and there are no firm reasons to suggest that the trends will change. Finally, trends should be damped toward no change as the forecast horizon increases.

These conclusions were drawn from the forecasting principles in the edited handbook on forecasting (Armstrong 2001) and they are described at forecastingprinciples.com. A summary of the principles of forecasting, now numbering 140, is provided in the Forecasting Audit on forecastingprinciples.com. The principles are formatted as a checklist.

The Forecasting Problem

In determining the best policies to implement to deal with the climate of the future, a policy maker first has to select an appropriate statistic to use to represent the changing climate. By convention, the statistic is the averaged global temperature as measured with thermometers at ground stations throughout the world, though in practice this is a far from satisfactory metric (e.g., Essex et al., 2007). It is then necessary to obtain forecasts and prediction intervals for each of the following:

1. What will happen to the mean global temperature in the long-term (say 20 years or longer)?
2. *If* accurate forecasts of mean global temperature changes can be obtained *and* these changes are substantial, *then* it would be necessary to forecast the effects of the changes on the health of living things and on the health and wealth of humans.
3. *If* reliable forecasts of the effects of the temperature changes on the health of living things and on the health and wealth of humans can be obtained *and* the forecasts are for substantial harmful effects, *then* it would be necessary to forecast the costs and benefits of alternative policy proposals. The concerns about changes in global mean temperature are based on the assumption that the earth is currently at the optimal temperature and that variations over years (unlike variations within years) are

undesirable. For a proper assessment, costs and benefits must be comprehensive. (For example, policy responses to Rachel Carlson's *Silent Spring*, failed to anticipate that millions of people would die from malaria because efforts to reduce the use of DDT).

4. *If* reliable forecasts of the costs and benefits of alternative policy proposals can be obtained *and* at least one proposal is predicted to lead to net benefits, *then* it would be necessary to forecast whether the policy changes can be implemented successfully.

If reliable forecasts of policy implementation can be obtained *and* the forecasts clearly support net benefits for the policy, *and* the policy can be successfully implemented, *then* the policy proposal should be implemented. A failure to obtain scientifically validated forecasts at any stage would render subsequent stages irrelevant. Thus, we focus on the first of the four forecasting problems.

Is it necessary to use scientific forecasting methods – that is, to use methods that have been shown by empirical validation to be relevant to the types of problems involved with climate forecasting? Or is it sufficient to have leading scientists examine the evidence and make forecasts? We address this issue before moving on to our audits.

On the value of forecasts by experts

Many policy decisions are based on forecasts by experts. Research on persuasion has shown that people have substantial faith in the value of such forecasts. This faith increases when experts agree with one another.

Although they may seem convincing at the time, expert forecasts make for humorous reading in retrospect. Cerf and Navasky's (1998) 310 pages of examples, such as the following, are illustrative:

> "[The nickel-iron battery will put] the gasoline buggies…out of existence in no time."
> Thomas Alva Edison 1910
>
> "There is not the slightest indication that [nuclear] energy will ever be obtainable."
> Albert Einstein 1932
>
> "I think there is a world market for about five computers."
> Thomas J. Watson, Chairman of IBM, 1943
>
> "A few decades hence, energy may be free."
> John von Neumann, Fermi Award-winning scientist, 1956

Examples of faulty expert climate forecasts are easy to find, but are perhaps less humorous:

> "If present trends continue, the world will be about four degrees colder in 1990, but eleven degrees colder in the year 2000. This is about twice what it would take to put us into an ice age."
> Kenneth Watt, UC Davis ecologist, Earth Day,
> April 22, 1970 Swarthmore College speech

Are these examples merely a matter of selective perception? The first author's review of empirical research on this problem led to the "Seer-sucker theory," stating that, "No matter how much evidence exists that seers do not exist, seers will find suckers" (Armstrong 1980). The amount of expertise does not matter beyond a basic minimum level. There are exceptions to the Seer-sucker Theory: When forecasters get substantial amounts of well-summarized feedback about the accuracy of their forecasts and about the reasons why the forecasts were or were not

accurate, they can improve their forecasts. This situation applies for short-term (e.g., up to five days) weather forecasts, but it does not apply to long-term climate forecasts.

Research since 1980 has added support to the Seer-sucker Theory. In particular, Tetlock (2005) recruited 284 people whose professions included, "commenting or offering advice on political and economic trends." He asked them to forecast the probability that various situations would or would not occur, picking areas (geographic and substantive) within and outside their areas of expertise. By 2003, he had accumulated over 82,000 forecasts. The experts barely if at all outperformed non-experts and neither group did well against simple rules.

Comparative empirical studies have routinely concluded that judgmental forecasting by experts is the least accurate of the methods available to make forecasts. For example, Ascher (1978, p. 200), in his analysis of long-term forecasts of electricity consumption found that judgment was the least accurate method.

Those who are not familiar with the research on expert forecasting believe that surveys of experts provide useful forecasts. Things have changed little since the days of Malthus in the 1800s, when he predicted that the demand for food would outpace supply, thus leading to mass starvation. He expressed his opinions mathematically: The supply of food increases arithmetically while population grows at a geometric rate. The math showed that starvation was inevitable.

Experts' forecasts of climate changes have long been popular. Anderson and Gainor (2006) found the following headlines in their search of the *New York Times*:

Sept. 18, 1924:	"MacMillan Reports Signs of New Ice Age."
March 27, 1933:	"America in Longest Warm Spell Since 1776"
May 21, 1974:	"Scientists Ponder Why World's Climate is Changing: A Major Cooling Widely Considered to be Inevitable."
Dec. 27, 2005:	"Past Hot Times Hold Few Reasons to Relax About New Warming."

In each case, the forecasts were made with a high degree of confidence. Similar stories appeared in *Time Magazine*.

In the mid-1970s, there was a political debate raging about whether the global climate was changing. The United States' National Defense University addressed this issue in their book, *Climate Change to the Year* 2000 (NDU 1978). This study involved 9 man-years of effort by Department of Defense and other agencies, aided by experts who received honoraria, and a contract of nearly $400,000 (in 2007 dollars). The heart of the study was a survey of experts. It provided them with a chart of "annual mean temperature, $0\text{-}80^0$ N. latitude," that showed temperature rising from 1870 to early 1940 then dropping sharply up to 1970. The conclusion, based primarily on 19 replies weighted by the study directors, was that while a slight increase in temperature might occur, uncertainty was so high that "the next twenty years will be similar to that of the past" and the effects of any change would be negligible. Clearly, this was a forecast by scientists, not a scientific forecast. However, it proved to be quite influential. The report was discussed in The Global 2000 Report to the President (Carter) and at the World Climate Conference in Geneva in 1979.

The methodology used in the past few decades has shifted from surveys of experts' opinions to the use of computer models. However, based on the explanations that we have seen, such models are, in effect, mathematical ways for the experts to express their opinions. 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. For example, and Keepin and

Wynne (1984) wrote in the summary of their study of the IIASA's "widely acclaimed" projections for global energy that, "Despite the appearance of analytical rigour… [they] are highly unstable and based on informal guesswork".

International surveys of climate scientists from 27 countries, obtained by Brat and von Storch in 1996 and 2003, were summarized by Bast and Taylor (2007). Of over 1,060 respondents, 35% agreed with statement, "Climate models can accurately predict future climates," and 47% percent disagreed.

Auditing the forecasting procedures

We searched prior literature to determine whether anyone had examined the question of whether scientific forecasting had been used for long-term climate forecasts. We then assessed the extent to which those who have made climate forecasts used evidence-based forecasting procedures. To do this, we surveyed scientists and others who are involved in long-term climate forecasting. We examined whether the publications they nominated as sources of credible long-term forecasts of mean global temperature used evidence-based forecasting procedures. Finally, we conducted an audit of the document most commonly nominated by our survey respondents as the source of the most credible long-term forecasts of global average temperature.

Prior audits

Stewart and Glantz (1985) conducted an audit of the forecast by the NDU (1978) that was described above. They were critical of the report because it showed a lack of awareness of proper forecasting methodology. Their audit was hampered because the organizers of the study said that the raw data had been destroyed and a request to the Institute for the Future about the sensitivity of the forecasts to the weights went unanswered. Judging from a Google Scholar search, climate forecasters have paid little attention to this paper.

Carter, et al. (2006) examined the *Stern Review* (Stern 2007). They concluded that the *Report* authors made predictions without any reference to scientific forecasting.

Pilkey and Pilkey-Jarvis (2007) concluded that the long-term climate forecasts that they examined were based only on the opinions of the scientists. The opinions were expressed in complex mathematical terms. There was no validation of the methodologies. They referred to the following quote as a summary on their page 45: "Today's scientists have substituted mathematics for experiments, and they wander off through equation after equation and eventually build a structure which has no relation to reality. (Nikola Telsa, inventor and electrical engineer, 1934.)" Thus, while it is sensible to be explicit about beliefs and to formulate these in a model, the forecaster must go beyond this to demonstrate that the relationships are valid and well-supported, especially when the models are complex.

Carter (2007) examined evidence on the predictive validity of the general circulation models (GCMs) used by the IPCC scientists. He found that while the models included some basic principles of physics, scientists had to make "educated guesses" about the values of many parameters because knowledge about the physical processes of the earth's climate is incomplete. In practice, the GCMs failed to predict recent global average temperatures as accurately as simple curve-fitting approaches (Carter 2007, pp. 64 – 65) and also forecast greater warming at higher altitudes when the opposite has been the case (p. 64). Further, individual GCMs produce widely different forecasts from the same initial conditions and minor changes in parameters can result in forecasts of global cooling (Essex and McKitrick, 2002). Interestingly, modeling results that project global cooling are often rejected as "outliers" or "obviously wrong" (e.g., Stainforth et al., 2005).

Taylor (2007) compared seasonal forecasts by New Zealand's National Institute of Water and Atmospheric Research with outcomes for the period May 2002 to April 2007. He found NIWA's forecasts of average regional temperatures for the season ahead were, at 48% correct, no more accurate than chance. That this is a general result was confirmed by New Zealand climatologist Dr Jim Renwick, who observed that NIWA's low success rate was comparable to that of other forecasting groups worldwide. He added that "Climate prediction is hard, half of the variability in the climate system is not predictable, so we don't expect to do terrifically well." Dr Renwick is an author on Working Group I of the IPCC 4th Assessment Report, and also serves on the World Meteorological Organisation Commission for Climatology Expert Team on Seasonal Forecasting; His expert view is that current GCM climate models are unable to predict future climate any better than chance (New Zealand Climate Science Coalition 2007).

In another example, the US National Hurricane Center's report on hurricane forecast accuracy noted, "No routinely-available early dynamical model had skill at 5 days" (Franklin 2007). This comment probably refers to forecasts for the paths of known, individual storms, but seasonal storm ensemble forecasts are clearly no more accurate. For example, the NHC's forecast for the 2006 season was widely off the mark. On June 7, Vice Admiral Conrad C. Lautenbacher, Jr. of the National Oceanic and Atmospheric Administration gave the following testimony before the Committee on Appropriations Subcommittee on Commerce, Justice and Science of the United States Senate (Lautenbacher 2006, p. 3):

> NOAA's prediction for the 2006 Atlantic hurricane season is for 13-16 tropical storms, with eight to 10 becoming hurricanes, of which four to six could become major hurricanes. … We are predicting an 80 percent likelihood of an above average number of storms in the Atlantic Basin this season. This is the highest percentage we have ever issued.

By the beginning of December, Gresko (2006) was able to write "The mild 2006 Atlantic hurricane season draws to a close Thursday without a single hurricane striking the United States".

Use of the scientific literature on forecasting methods

In April 2007, we used the Advanced Search function of Google Scholar to get a general sense of the extent to which environmental forecasters use scientific literature on forecasting methods. When we searched for "global warming" and "forecasting principles," we found no relevant sites. Nor did we find any relevant sites for "forecastingprinciples.com" and "global warming." Nor were there any relevant sites mentioned for the relevant-sounding paper, "Forecasting for Environmental Decision-Making" (Armstrong 1999) published in a book with a relevant title: *Tools to Aid Environmental Decision Making*.

We examined the references in Chapter 8 (on evaluation) of the 2007 Intergovernmental Panel on Climate Change's Working Group One report. There were 788 references. Of these, none had any apparent relationship to forecasting methodology. The task was not that difficult as most papers had titles such as "Using stable water isotopes to evaluate basin-scale simulations of surface water budgets," and "Oceanic isopycnal mixing by coordinate rotation."

It is hard to understand how scientific forecasting could be conducted without any reference to the literature on how to make such forecasts.

Surveys of scientists involved in climate forecasts

We conducted a survey of scientists involved in long-term climate forecasting and of policy makers. Our primary concern was to identify the most important forecasts and how those forecasts were made. In particular, we wished to know if the most widely accepted forecasts of global average temperature were based on the opinions of experts or on scientific forecasting methods. Given the conclusion from our Google search that many scientists are unaware of evidence-based findings related to forecasting methods, our hypothesis was that the forecasts were based on the opinions of scientists.

We sent a questionnaire to experts who had expressed diverse opinions on global warming. We generated lists of experts by identifying key people and asking them to identify others. (The lists are provided in Appendix A.) Most (70%) of the 240 experts on our lists were IPCC reviewers and authors.

The questionnaire asked the experts to provide references for what they regarded as the most credible source of long-range forecasts of mean global temperatures.

We strove for simplicity to minimize resistance to our request. Even busy people should have time to send a few references, especially if they believe that it is important to evaluate the quality of the forecasts that will influence major decisions.

We received useful responses from 51 people, 42 of whom provided references to what they regarded as credible sources of long-range forecasts of mean global temperatures. Interestingly, 8 respondents provided references in support of their claims that no credible forecasts exist. Of the 42 expert respondents who were associated with global warming views, 30 referred us to the IPCC's report.

Forecasting audit for global warming

Based on the replies to this survey, it was clear that the IPCC's Working Group 1 Report contained the forecasts that are viewed as most credible by the climate community. These forecasts are contained in Chapter 10 of the Report and the models that are used to forecast climate are assessed in Chapter 8, "Climate Models and Their Evaluation" (Randall et al. 2007). Chapter 8 provided the most useful information on the forecasting process used by the IPCC to derive forecasts of mean global temperatures, so we audited that chapter.

In apparent contradiction to claims by some climate experts that the IPCC provides "projections" and not "forecasts, the word "forecast" and its derivatives occurred 37 times, and "predict" and its derivatives occur 90 times in the body of Chapter 8. Recall also that most of our respondents (29 of whom were IPCC authors or reviewers) nominated the IPCC report as the most credible source of forecasts (not projections) of global average temperature.

Chapter 8 was, in our judgment, poorly written. The writing showed little concern for the target readership, provided extensive detail on items that are of little interest in judging the merits of the forecasting process, provided references without describing what readers might find, and imposed an incredible burden on readers by providing 788 references. The readability of the chapter was low. For example, Section 8.2.1.3 "Parametrization," a critical section for understanding the forecasting process, scored 23 for Flesch-Kinkaid Reading Ease ("plain English" is 60 and the *Harvard Law Review* scores 32), and a "grade level" Gunning-Fog Index of 21 (5 is "very readable" and 20 is "very difficult"). In addition, the Chapter reads in places like a sales brochure. In the three-page executive summary, the terms, "new," and "improved," and related derivatives appeared 17 times. Most significantly, the chapter omitted key details on the assumptions and the forecasting process that were used by its authors. Given these problems, we found it puzzling that so many of our respondents nominated the IPCC report as the most credible paper on long-term climate forecasts.

We each made a formal, independent audit of IPCC Chapter 8 in May 2007. To do so, we used the Forecasting Audit software on the forecastingprinciples.com site, which is based on material originally published in Armstrong (2001). To our knowledge, it is the only evidence-based tool for the evaluation of forecasting procedures.

While Chapter 8 required many hours to read, it took us each about one hour to rate the forecasting approach described in the Chapter using the Audit software. We have each been involved with developing the Forecasting Audit program, so other users would require much more time. Ratings are on a 5-point scale from -2 to +2. A rating of +2 indicates the forecasting procedures were consistent with a principle, and a rating of -2 indicates failure to comply with a principle. The Audit software also has options to indicate that there is insufficient information to rate the procedures or that the principle is not relevant to a particular forecasting problem.

Our overall average ratings were similar at -1.37 and -1.35. We compared the individual ratings for individual principles and discussed inconsistencies. In some case we averaged the ratings, truncating toward zero, in other cases we decided that there was insufficient information or that the information was too ambiguous to rate with confidence. Our final ratings are fully disclosed in the Special Interest Group section of the forecastingprinciples.com site that is devoted to Public Policy (publicpolicyforecasting.com).

Of the 140 principles in the Forecasting Audit, we judged that 127 were relevant for auditing the forecasting problem addressed in Chapter 8. The Chapter provided insufficient information to rate the forecasting procedures that were used against 38 of these principles. The number of items that we were unable to rate was quite high. For example, we did not rate the Chapter against Principle 10.2, "Use all important variables." At least in part, our difficulty in auditing the Chapter was due to the fact that it is poorly written, as the readability statistics show. It was sometimes difficult to know whether the information we sought was present or not.

Of the 89 forecasting principles that we were able to rate, the Chapter violated 72. Adherence to some of the key principles is necessary for forecasts to be valid. We address three such principles, all based on strong empirical evidence: violation of any one of them would render the *IPCC* climate forecasts invalid. In the event, all three of the key principles are violated by the forecasting procedures described in IPCC Chapter 8.

1.4 Consider whether a formal forecasting procedure can beat a naïve method

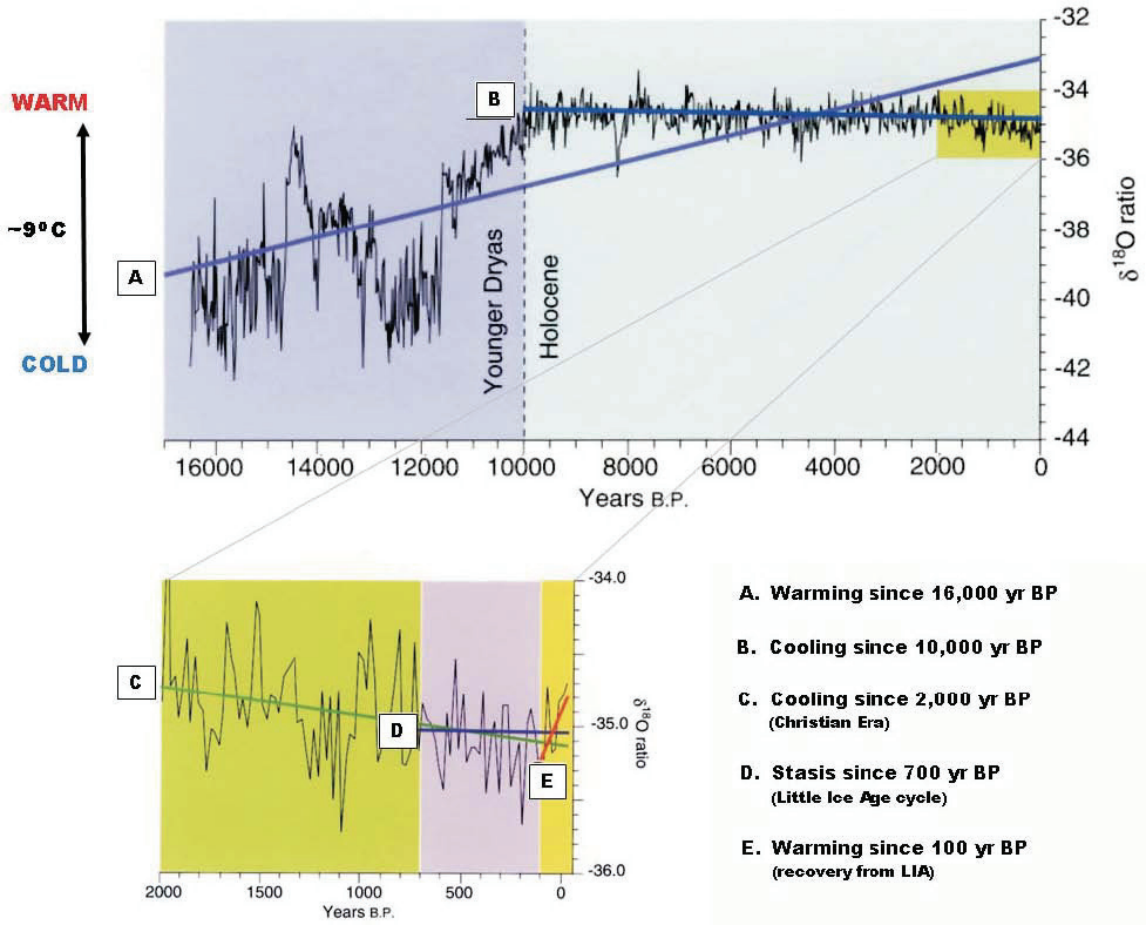
This principle refers to whether a forecasting method can be developed that would do better than a naïve method. To a large extent, forecastability depends on the level of uncertainty. To the extent that uncertainty is high, the forecasting method should emphasize the naïve method. (This is illustrated by regression model coefficients: when uncertainty increases, the coefficients tend towards zero.) Importantly, departures from the naïve model tend to increase forecast error when uncertainty is high.

In our judgment, the uncertainty in forecasting global mean temperature is extremely high. For example, what, if any, is the current trend? Carter (2007, p. 67) wrote:

> …the slope and magnitude of temperature trends inferred from time-series data depend upon the choice of data end points. Drawing trend lines through highly variable, cyclic temperature data or proxy data is therefore a dubious exercise. 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.

Uncertainty over the trend in mean temperature is illustrated by Figure 1 (taken from Carter (2007) and originating in Davis and Bohling (2001)).

Figure 1



Climatic cycling over the last 16 000 years as indicated by averaged 20-year oxygen isotope ratios from the GISP2 Greenland ice core. Trend lines A-E all extend up to the end of the 20th century, fitted through the data for the last 16 000, 10 000, 2000, 700 and 100 years, respectively. The trends are indicative of both warming and cooling, depending upon the chosen starting point.

Global climate is complex. Scientific evidence on many key relationships is weak or absent; e.g., does CO_2 cause high temperatures or do high temperatures cause CO_2 (e.g. Jaworowski 2007)? And what effect do variations in solar activity have? (See for example Figure 2, reproduced from Soon 2005, and Svensmark's 2007 paper on the sun's influence on cloud seeding by cosmic rays). Measurements of key variables such as local temperatures and a representative global temperature are contentious in the case of modern measurements, because of possible artifacts such as the urban heat island effect, and often speculative in the case of ancient ones, such as those climate proxies derived from tree ring and ice-core data (Carter 2007). Finally, it is difficult to forecast the causal variables.

The already high level of uncertainty rises rapidly as the forecast horizon increases.

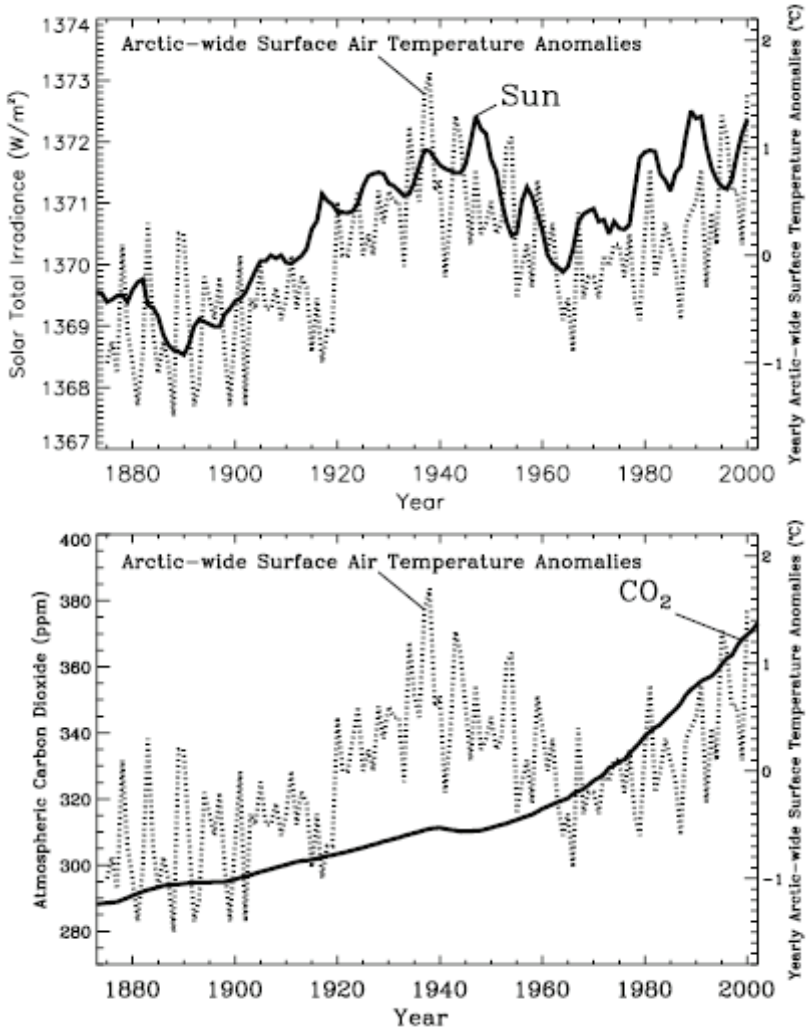
While the authors of Chapter 8 claim that the forecasts of global mean temperature are well-founded, their language is imprecise and relies heavily on such words as "generally," "reasonable well," "widely," and "relatively" [to what?]. The report makes many explicit references to uncertainty. For example, the phrases ". . . it is not yet possible to determine which estimates of the climate change cloud feedbacks are the most reliable" and "Despite advances since the TAR, substantial uncertainty remains in the magnitude of cryospheric feedbacks within AOGCMs" appear on p. 593. In discussing the modeling of temperature, the authors wrote, "The extent to

which these systematic model errors affect a model's response to external perturbations is unknown, but may be significant" (p. 608), and, "The diurnal temperature range… is generally too small in the models, in many regions by as much as 50%" (p. 609), and "It is not yet known why models generally underestimate the diurnal temperature range." The following words and phrases appear at least once in the Chapter: unknown, uncertain, unclear, not clear, disagreement, uncertain, not fully understood, appears, not well observed, variability, variety, difference, unresolved, not resolved, and poorly understood.

Given the high uncertainty, the naïve method for this situation would be the "no-change" model. Remarkably, nowhere does the IPCC Report address the issue of forecastability. It should have been addressed prior to spending enormous sums on complex forecasting models.

In effect, given the current state of uncertainty regarding climate, prior evidence on forecasting methods suggests that attempts to improve upon the naïve model might increase forecast error. To reverse this conclusion, one would have to produce validated evidence in favor of certain methods. Such evidence is not provided in Chapter 8 of the IPCC report.

Figure 2



Annual-mean Arctic-wide air temperature anomaly time series (dotted lines) correlated with the estimated total solar irradiance (top panel; solid lines) and with the atmospheric carbon dioxide, CO_2, mixing ratio (bottom panel; solid lines) from 1875 to 2000.

We are not suggesting that climate change cannot be forecast, only that this has yet to be demonstrated. We expect that such methods as the naïve model with drift, rule-based forecasting, well-specified causal models, and combined forecasts might prove useful. All of those methods are discussed in Armstrong (2001).To our knowledge, none of these methods has been examined to date.

9.3 Do not use fit to develop the model.

It is not clear to what extent the models produced by the IPCC are either based on, or have been tested against, sound empirical data. However, some statements were made about the ability of the models described in Chapter 8 to fit historical data, after tweaking of their parameters. Extensive research has shown that the ability of models to fit historical data has little relationship to forecast accuracy (See "Evaluating Methods" in Armstrong 2001.) It is well known that fit can be improved by making the model more complex. The consequence of increasing complexity to improve fit, however, is to decrease the accuracy of forecasts. The 12 authors of Chapter 8 appeared to be unaware of this principle.

13.26 Use out-of-sample data to test the forecasts.

Chapter 8 did not provide evidence on the accuracy of *ex ante* long-range forecasts from the models used to generate the IPCC's forecasts of climate change. It would have been feasible to assess the accuracy of alternative forecasting methods for short- and medium-range forecasts by using "successive updating." This involves with-holding data on a number of years, then providing forecasts for one-year ahead, then two-years ahead, and so on up to 20 years. The actual years could be disguised during these validation procedures. Furthermore, the years could be reversed (without telling the forecasters) to assess back-casting accuracy. If the methods were unable to improve accuracy on these forecasts, beyond the naïve method, it is difficult to imagine why accuracy could improve beyond 20 years.

A list of all 73 violations of forecasting principles by the authors of IPCC Chapter 8 is provided on the Public Policy Special Interest Group Page at forecastingprinciples.com. The many violations provide further evidence that the authors were unaware of evidence-based principles for forecasting. If they were aware of them, it would have been incumbent on them to present evidence to justify their departures from best forecasting practice. They did not do so. We conclude that because the forecasting processes described in the Chapter overlook scientific evidence on forecasting, the IPCC forecasts of climate change are not scientific.

Others are invited to provide audits of Chapter 8 or other studies related to climate forecasting. Audits should also be done for other studies in an attempt to find climate studies that do *not* violate evidence-based methods for forecasting.

The Role of the Forecasting Audit

The Forecasting Audit should be used early and often by climate change forecasters and their clients, as it should be by those involved in other public policy issues, such as national health plan proposals and proposals for gun control. Doing so would help to ensure that they are using appropriate forecasting procedures. Outside evaluators should also be encouraged to conduct audits. The audits should be made available to both the sponsors of the study and the public by posting on an open web site.

Gaining acceptance of scientific findings

History is filled with the poor treatment of those who attempt to introduce science into arenas where emotions are high and vested interests are threatened. Galileo springs to mind. Scientists in the West at least no longer face death when presenting their findings. Nevertheless, the scientific review system currently acts to prevent the publication of research findings that conflict with commonly held beliefs (for a review of research on this matter, see Armstrong 1997).

We recommend the use of objective evidence-based (scientific) procedures to assess the validity of global warming forecasts. Our belief is that science will win out in the long run. The problem is that we may waste enormous resources over a short-run that might last for the next few decades.

Prior to conducting an audit, one might ask policy makers to say what information would be sufficient to change their opinions. People who are able to specify such evidence are often able to change their opinions. When we have used this question among academic researchers and students, we find that many of them are willing to specify such information. Disturbingly, however, many others are unable to even imagine that any information could possibly change their minds.

Conclusions

The *Stern Review* concluded that, "The scientific evidence is now overwhelming; climate change presents very serious global risks, and it demands an urgent global response" (Stern 2007, p. xv). We have not been able to find any scientific evidence to support such a claim. We can only hope that before committing resources, decision makers will insist on scientific forecasts rather than accept the opinions of some scientists.

To provide forecasts that are useful for policy-making, one would need to prepare forecasts not only of global temperature, but also of the net effects of any temperature change; then on the effects of policy changes aimed at reducing temperature changes or the negative effects of it, the costs of such changes, and the likelihood of successful implementation. A failure at any stage would nullify any value to the forecasts.

We have shown that failure occurs at the first stage of analysis. Specifically, we have been unable to find a single scientific forecast to support the currently widespread belief in dangerous, human-caused "global warming". Prior research on forecasting suggests that a naïve (no change) forecast would be superior to current predictions which are, in effect, experts' judgments only.

Based on our Google searches, those forecasting long-term climate change have no apparent knowledge of evidence-based forecasting methods, so we expect that the same conclusions would apply to the other three necessary parts of the forecasting problem.

By relying on evidence-based forecasting methods, we conclude that policies founded on predictions of man-made global warming from models based on the opinions of scientists will be harmful.

Given the conditions involved in long-term global forecasts and the high uncertainty involved, prior research on forecasting suggests that even if the forecasting methods were properly applied, it may not be possible to improve upon the naïve, "no-change," forecast. We do not even have evidence that it is possible to make useful medium term (e.g., one to five year) forecasts.

Our paper is concerned with rational assessments of public policy, not with public opinions. People will continue to believe that serious manmade global warming exists as they will continue to believe other things that have no scientific support (e.g., the biblical creation story, astrology, minimum wages as a way to help poor people, and so on). It is appropriate for concerned individuals to donate money or time for such perceived issues as long as they do not harm others

in these efforts. It is not appropriate to impose on others policies that have not been shown to have scientific merit.

One might say that it is important to consider steps to prevent global warming, but we have the same level of confidence in saying that we should take steps to prevent global cooling. The more important question is "what is the best way to invest our resources for the benefit of mankind?" This would lead to such trade-offs as asking whether it is better to spend a dollar on reducing AIDS or air pollution or malaria or breast cancer, where we know what policies will work, or to spend it on controlling future climate control, where we have no scientific evidence. Given the large uncertainties of climate change science, government polices on climate control are unwarranted and will reduce the well-being of people who are not the beneficiaries of the wealth redistribution that will occur as a result of such policies. For example, wealthy owners of beach-front properties, climate researchers, consultants, those involved in carbon trading markets, and manufacturers of some energy sources might expect to gain from global warming policies,

Those who advocate various positions on the climate owe it to the people who would be affected by the policies they recommend to base their advocacy on scientific forecasts that address all four of the key areas that are necessary for a rational analysis of the problem.

Next Steps

We have posted a copy of this paper for peer review on publicpolicyforecasting.com. We will use the peer review to inform our revisions. With the authors' permission, we will post the peer review, although it will be moderated to exclude "advertisements" and ad hominem arguments. All posted peer review will need to include the authors' names, position, email address, and any matter that might be construed as biasing the reviewer.

We will also post applications of the evidence-based Forecasting Audit (available on forecastingprinciples.com) to papers that address any of the four forecasting problems related to the formulation of public policy relevant to climate changes. As with peer review, we will require all relevant information on the people who conduct the audits prior to posting the audits.

Acknowledgements

Bob Carter made many helpful suggestions on the content and writing of the paper. Paul Goodwin and Tom Yokum provided useful suggestions on an early draft.

References

Anderson, R. Warren and Dan Gainor (2006). "Fire and Ice: Journalists have warned of climate change for 100 years, but can't decide weather we face an ice age or warming." Business and Media Institute, May 17. Available at http://www.businessandmedia.org/specialreports/2006/fireandice/FireandIce.pdf

Armstrong, J. Scott (1980), "The Seer-sucker theory: The value of experts in forecasting," *Technology Review* 83 (June-July), 16-24.

Armstrong, J. Scott (1978; 1985), *Long-Range Forecasting: From Crystal Ball to Computer."* New York: Wiley-Interscience.

Armstrong, J. Scott (1997), "Peer review for journals: evidence on quality control, fairness, and innovation", *Science and Engineering Ethics*, 3, 63-84

Armstrong, J. Scott (1999). "Forecasting for environmental decision-making," in V.H. Dale and M.E. English, eds., *Tools to Aid Environmental Decision Making*. New York: Springer-Verlag, pp. 192-225.

Armstrong, J. Scott (2001), *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Kluwer Academic Publishers.

Armstrong, J. Scott (2006), "Findings from evidence-based forecasting: Methods for reducing forecast error," *International Journal of Forecasting*, 22, 583-598.

Ascher W. (1978), *Forecasting: An Appraisal for Policy Makers and Planners*. Baltimore: Johns Hopkins University Press.

Bast, Joseph and James M. Taylor (2007), *Scientific consensus on global warming*. The Heartland Institute: Chicago Illinois, Available http://downloads.heartland.org/20861.pdf. [The responses to all questions in the 1996 and 2003 surveys by Brat and von Storch are included as an appendix.]

Carter, Robert M. (2007), "The myth of dangerous human-caused climate change," The Aus/MM New Leaders Conference, Brisbane May 3, 2007. Full text available from http://members.iinet.net.au/~glrmc/new_page_1.htm

Carter, Robert M., C. R. de Freitas, I. M. Goklany & D. Holland (2006), "The Stern review: A dual critique: Part 1," *World Economics*, 167-198.

Cerf, Christopher & Victor Navasky (1984), *The Experts Speak*. New York: Pantheon. Baltimore, MD: Johns Hopkins Univ. Press.

Craig, Paul P., Ashok Gadgil, and Jonathan G, Koomey (2002). "What Can History Teach Us? A Retrospective Examination of Long-Term Energy Forecasts for the United States." *Annual Review of Energy and the Environment, 27*, 83-118.

Davis, John C., and Geoffrey Bohling (2001). "The Search for Patterns in Ice-Core Temperature Curves" in Gerhard, Lee C., William E. Harrison, and Bernold M. Hanson, eds.,2001, *Geological Perspectives of Global Climate Change: American Association of Petroleum Geologists Studies in Geology #47*, Tulsa, OK, p. 213-230.

Essex, C., McKitrick, R. & Andresen, B. (2007). "Does a global temperature exist?" *Journal of Non-Equilibrium Thermodynamics 32*, 1-27.

Essex, C. & McKitrick, R. (2002). *Taken by Storm. The Troubled Science, Policy & Politics of Global Warming*, Key Porter Books, Toronto,

Franklin, James L. (2007). "2006 National Hurricane Center Forecast Verification Report." National Hurricane Center. Available at http://www.nhc.noaa.gov/verification/pdfs/Verification_2006.pdf

Gresko, Jessica (2006). "2006 Hurricane Forecast a Flop." Associated Press, posted 01 December 10:28 am ET. Available at http://www.livescience.com/environment/061201_ap_hurricane_flop.html

Ihaka, James (2007). "Half chance of getting weather forecast right." *New Zealand Herald. June 6*. Available at http://203.99.65.121/section/1/story.cfm?c_id=1&objectid=10443866

Jaworowski, Z. (2007) "CO_2: The greatest scientific scandal of our time". *Executive Intelligence Review*, 34(11), 38-53. Available at http://www.warwickhughes.com/icecore/zjmar07.pdf

Keepin, B. and Wynne, B. (1984) "Technical analysis of IIASA energy scenarios". *Nature*, 312, 691-695.

Lautenbacher, Conrad C. Jr. (2006). "Written testimony of Vice Admiral Conrad C. Lautenbacher, Jr. U.S. Navy (ret.) Under Secretary Of Commerce For Oceans And Atmosphere And Administrator For National Oceanic and Atmospheric Administration U. S. Department Of Commerce Oversight Hearing on '2006 Hurricane Season' before the Committee On Appropriations Subcommittee On Commerce, Justice and Science of the United States Senate June 7, 2006." National Oceanic and Atmospheric Administration. Available at http://www.legislative.noaa.gov/Testimony/lautenbacher060706.pdf.

NDU (1978), *Climate Change to the Year 2000*. National Defense University. Washington, D.C.

New Zealand Climate Science Coalition (2007). "World climate predictors right only half the time." Media release 7 June. Available at http://www.scoop.co.nz/stories/SC0706/S00026.htm

Pilkey, Orrin H. and Linda Pilkey-Jarvis (2007), *Useless Arithmetic Why Environmental Scientists Can't predict the Future.* New York: Columbia University Press.

Randall, D.A., R.A. Wood, S. Bony, R. Colman, T. Fichefet, J. Fyfe, V. Kattsov, A. Pitman, J. Shukla, J. Srinivasan, R.J. Stouffer, A. Sumi and K.E. Taylor (2007). "Climate Models and Their Evaluation," in Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller eds., *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK and New York, NY, USA: Cambridge University Press.

Soon, W. W.-H. (2005). "Variable solar irradiance as a plausible agent for multidecadal variations in the Arctic-wide surface air temperature record of the past 130 years", *Geophysical Research Letters, 32,* Available from http://www.agu.org/pubs/crossref/2005/2005GL023429.shtml

Schnaars, Steven P. & R. J. Bavuso (1986), "Extrapolation models on very short-term forecasts," *Journal of Business Research*, 14, 27-36.

Stainforth, D A, Aina, T, Christensen, C, Collins, M, Faull, N, Frame, D. J, Kettleborough, J A, Knight, S, Martin, A, Murphy, J M, Piani, C, Sexton, D, Smith, L A, Spicer, R A, Thorpe, A J and Allen, M R, (2005). Uncertainty in predictions of the climate response to rising levels of greenhouse gases, *Nature*, *433*, 403-406.

Stern, N. (2007). *The Economics of Climate Change: The Stern Review*. New York: Cambridge University Press. Available from http://www.hmtreasury.gov.uk/independent_reviews/stern_review_economics_climate_change/sternreview_index.cfm

Stewart, Thomas R & Michael H. Glantz (1985), "Expert judgment and climate forecasting: A methodological critique of '*Climate Change to the Year 2000'*," *Climate Change*, 7, 159-183.

Svensmark, H. (2007). Cosmoclimatology: a new theory emerges. *Astronomy and Geophysics*, 48, 1.18-1.24. Available from http://www.blackwell-synergy.com/doi/abs/10.1111/j.1468-4004.2007.48118.x

Taylor, Malcolm (2007). An evaluation of NIWA's climate predictions for May 2002 to April 2007. Climate Science Coalition. Available at http://www.climatescience.org.nz/assets/2007691051580.ClimateUpdateEvaluationText.pdf Data available at http://www.climatescience.org.nz/assets/2007691059100.ClimateUpdateEvaluationCalc.xls.pdf

Tetlock, Philip E. (2005), *Expert Political Judgment: How Good Is It? How Can We Know?* (Princeton Univ. Press).

Appendix A: People to whom we sent our questionnaire (* indicates a relevant response)

IPCC Working Group 1
Myles Allen, Richard Alley, Ian Allison, Peter Ambenje, Vincenzo Artale, Paulo Artaxo, Alphonsus Baede, Roger Barry, Terje Berntsen, Richard A. Betts, Nathaniel L. Bindoff, Roxana Bojariu, Sandrine Bony, Kansri Boonpragob, Pascale Braconnot, Guy Brasseur, Keith Briffa, Aristita Busuioc, Jorge Carrasco, Anny Cazenave, Anthony Chen*, Amnat Chidthaisong, Jens Hesselbjerg Christensen, Philippe Ciais*, William Collins, Robert Colman*, Peter Cox, Ulrich Cubasch, Pedro Leite Da Silva Dias, Kenneth L. Denman, Robert Dickinson, Yihui Ding, Jean-Claude Duplessy, David Easterling, David W. Fahey, Thierry Fichefet*, Gregory Flato, Piers M. de F. Forster*, Pierre Friedlingstein, Congbin Fu, Yoshiyuki Fuji, John Fyfe, Xuejie Gao, Amadou Thierno Gaye*, Nathan Gillett*, Filippo Giorgi, Jonathan Gregory*, David Griggs, Sergey Gulev, Kimio Hanawa, Didier Hauglustaine, James Haywood, Gabriele Hegerl*, Martin Heimann*, Christoph Heinze, Isaac Held*, Bruce Hewitson, Elisabeth Holland, Brian Hoskins, Daniel Jacob, Bubu Pateh Jallow, Eystein Jansen*, Philip Jones, Richard Jones, Fortunat Joos, Jean Jouzel, Tom Karl, David Karoly*, Georg Kaser, Vladimir Kattsov, Akio Kitoh, Albert Klein Tank, Reto Knutti, Toshio Koike, Rupa Kumar Kolli, Won-Tae Kwon, Laurent Labeyrie, René Laprise, Corrine Le Quéré, Hervé Le Treut, Judith Lean, Peter Lemke, Sydney Levitus, Ulrike Lohmann, David C. Lowe, Yong Luo, Victor Magaña Rueda, Elisa Manzini, Jose Antonio Marengo, Maria Martelo, Valérie Masson-Delmotte, Taroh Matsuno, Cecilie Mauritzen, Bryant Mcavaney, Linda Mearns, Gerald Meehl, Claudio Guillermo Menendez, John Mitchell, Abdalah Mokssit, Mario Molina, Philip Mote*, James Murphy, Gunnar Myhre, Teruyuki Nakajima, John Nganga, Neville Nicholls, Akira Noda, Yukihiro Nojiri, Laban Ogallo, Daniel Olago, Bette Otto-Bliesner, Jonathan Overpeck*, Govind Ballabh Pant, David Parker, Wm. Richard Peltier, Joyce Penner*, Thomas Peterson*, Andrew Pitman, Serge Planton, Michael Prather*, Ronald Prinn, Graciela Raga, Fatemeh Rahimzadeh, Stefan Rahmstorf, Jouni Räisänen, Srikanthan (S.) Ramachandran, Veerabhadran Ramanathan, Venkatachalam Ramaswamy, Rengaswamy Ramesh, David Randall*, Sarah Raper, Dominique Raynaud, Jiawen Ren, James A. Renwick, David Rind, Annette Rinke, Matilde M. Rusticucci, Abdoulaye Sarr, Michael Schulz*, Jagadish Shukla, C. K. Shum, Robert H. Socolow*, Brian Soden, Olga Solomina*, Richard Somerville*, Jayaraman Srinivasan, Thomas Stocker, Peter A. Stott*, Ron Stouffer, Akimasa Sumi, Lynne D. Talley, Karl E. Taylor*, Kevin Trenberth*, Alakkat S. Unnikrishnan, Rob Van Dorland, Ricardo Villalba, Ian G. Watterson*, Andrew Weaver*, Penny Whetton, Jurgen Willebrand, Steven C. Wofsy, Richard A. Wood, David Wratt, Panmao Zhai, Tingjun Zhang, De'er Zhang, Xiaoye Zhang, Zong-Ci Zhao, Francis Zwiers*

Union of Concerned Scientists
Brenda Ekwurzel, Peter Frumhoff, Amy Lynd Luers

Channel 4 "The Great Global Warming Swindle" documentary (2007)
Bert Bolin, Piers Corbyn*, Eigil Friis-Christensen, James Shitwaki, Frederick Singer, Carl Wunsch*

Wikipedia's list of global warming "skeptics"
Khabibullo Ismailovich Abdusamatov*, Syun-Ichi Akasofu*, Sallie Baliunas, Tim Ball, Robert Balling*, Fred Barnes, Joe Barton, Joe Bastardi, David Bellamy, Tom Bethell, Robert Bidinotto, Roy Blunt, Sonja Boehmer, Andrew Bolt, John Brignell*, Nigel Calder, Ian Castles*, George Chilingarian, John Christy*, Ian Clark, Philip Cooney, Robert Davis, David Deming*, David Douglass, Lester Hogan, Craig Idso, Keith Idso, Sherwood Idso, Zbigniew Jaworowski, Wibjorn Karlen, William Kininmonth, Nigel Lawson, Douglas Leahey, David Legates, Richard Lindzen*, Ross Mckitrick*, Patrick Michaels, Lubos Motl*, Kary Mullis, Tad Murty, Tim Patterson, Benny

Peiser*, Ian Plimer, Arthur Robinson, Frederick Seitz, Nir Shaviv, Fred Smith, Willie Soon, Thomas Sowell, Roy Spencer, Philip Stott, Hendrik Tennekes, Jan Veizer, Peter Walsh, Edward Wegman

Other sources
Daniel Abbasi, Augie Auer, Jonathan Boston, Daniel Botkin*, Reid Bryson, Robert Carter*, Ralph Chapman, Al Gore, Kirtland C. Griffin*, David Henderson, Christopher Landsea*, Bjorn Lomborg, Tim Osborn, Roger Pielke*, Henrik Saxe, Thomas Schelling*, Matthew Sobel, Nicholas Stern*, Brian Valentine*, Antonio Zichichi.