

Evidence-based Forecasting for Climate Change Policies

Kesten C Green, University of South Australia, Australia
Willie Soon, Harvard-Smithsonian Center for Astrophysics, U.S.A.
J Scott Armstrong, University of Pennsylvania, U.S.A.

DRAFT

October 12, 2013
GAS-for-cli-pol-101.doc

In order to meet policy makers' need for climate forecasts, this paper extends the application of evidence-based forecasting of global mean temperatures. The extensions utilize more years of global mean temperature data and 34 years of better data. Forecasts from the no-trend model were compared with forecasts from six more-sophisticated methods that take account of more than the most recent data point to generate forecasts on the basis of recent levels, trends, and patterns. The more-sophisticated methods provided inconsequential improvements in forecast accuracy for horizons up to 10 years, beyond which horizon the no-change model provided the most accurate forecasts. The findings reinforce earlier conclusions that the IPCC scenario of dangerous increases in global mean temperatures fails to pass basic validation tests, and that the no-trend model provides the only scientific long-term climate forecasts. Without a scientific forecast of a global warming, there is no rational basis for government regulations, subsidies, funding, or penalties. Furthermore, a scientific forecast of global warming is not, by itself, a basis for action. That would require scientific forecasts that the warming would be harmful and forecasts that cost-effective policies would reduce the harm.

Key words: complexity; conservatism; cost benefit analysis; decision making, extrapolation; public policy; uncertainty; validation.

Latest draft available from <http://www.kestengreen.com/gas-improvements.pdf>

Basis of the IPCC's alarming scenarios

The United Nations' Intergovernmental Panel on Climate Change was established in 1988 in order to provide "the world with a clear scientific view on the current state of knowledge in climate change and its potential environmental and socio-economic impacts"¹. In its most recent report, the Intergovernmental Panel on Climate Change (IPCC) "projects" rapidly increasing global mean temperatures and of substantial harm arising from increasing temperatures². The IPCC attributes projections of increasing temperatures to carbon dioxide (CO₂) emissions from human activity. Lobby groups and politicians refer to the IPCC's reports to justify expensive policy recommendations such as abandoning coastal settlements, taxing carbon dioxide emissions, and subsidizing energy sources. Many governments have implemented such policies.

The IPCC relies upon scenarios derived from the outputs of computer models. While the climate modelers claim that their scenarios are *not* forecasts, the IPCC treats them as if they were. The words "forecast" and "predict" and their derivations, occur 127 times in the IPCC's AR4 report (Randall et al 2007).

The current global warming alarm movement parallels the global cooling alarm movement that spanned the 1970s³. President Carter created commissions that elicited judgmental forecasts from experts, which were then averaged.

The global warming movement also relies on experts' judgments, but this time the experts were asked to describe what causes temperatures to change. In their article examining the IPCC's procedures, Fildes and Kourentzes (2011) observe: "a major part of the model building is judgmental" (p. 970). They report that judgment is heavily relied upon for constructing a global temperature series from selected local readings, including and excluding variables, representing climate processes, estimating parameters, selecting initial conditions, and selecting and interpreting model outputs

As with the Carter Administration forecasts of an impending ice age, the global warming alarm is based on the judgments of selected experts. This reliance on the judgments of experts flies in the face of strong findings from forecasting research that expert judgments are not useful for forecasting in complex and uncertain situations (Armstrong 1980). This finding has been supported by subsequent research. In one study, 284 experts made 82,361 forecasts of political, social, and economic developments over a 20-year period. Their forecasts were little more accurate than the forecasts of novices, and less accurate than simple rules (Tetlock 2005).

The computer models upon which the IPCC rely, are simply representations of the judgments of some experts. The outputs of the models show the implications of those judgments. They are referred to as scenarios. As shown by Gregory and Duran (2001), scenarios provide inaccurate forecasts because they involve detailed descriptions of an imagined future that make them seem especially likely.

¹ From <http://www.ipcc.ch/organization/organization.shtml#UVPIqb-j5jg> on 8 October 2013.

² IPCC AR5 reports from Working Group I, II and III due out in September 2013, March 2014, and April 2014, respectively.

³ For an account, see <http://www.financialpost.com/story.html?id=bdc24964-7f82-4f7a-863c-f0ff43010278>

There have been claims that the nearly all climate scientists believe alarming global warming scenarios are valid forecasts. While this is a strange claim in view of the fact that 32,000 scientists signed the Oregon Petition to state that there is no basis for scenarios of dangerous manmade global warming, validation is not a matter for voting by scientists. Historically, voting has been regarded as a barrier to scientific advances.

To assess the validity of the IPCC process in a systematic way, Green and Armstrong (2007) conducted a structured audit of the climate models used by the IPCC, as described primarily in Randall *et al.* (2007). The audit examined the extent to which the scenarios of dangerous global warming were developed using evidence-based forecasting principles; scientific forecasting procedures, in other words. The principles, published in 2001, were due to the work of 39 scientists from different disciplines (psychology, economics, business, weather, criminology, etc.) and different countries, aided by 123 reviewers. Following that, the forecastingprinciples.com site was established to make the principles easily available to everyone. For over a decade, was the first item on a Google search for “forecasting.”⁴ The site also updates the principle when new evidence is received.

When global warming alarmists claimed that the principles did not apply to them, we asked for a list of relevant principles that do not apply and the evidence that they do not, and a list of principles that we were not aware of that do apply, and the evidence for them. Our requests to see the principles and evidence of their validity were ignored. The IPCC reports provided no references to support the validity of their procedures. We have been unable to find any such publications.

The audit revealed that the modeling procedures relied upon by the IPCC violated 72 of 89 relevant evidence-based forecasting principles. For example, IPCC lead authors resisted disclosing their data and methods (Michaels 2009). In other words, we found that the forecasting procedures were not valid.

Is the current situation unique?

People in need of forecasts commonly assume that their situations are unique. Not surprisingly, then, the alarm over dangerous manmade global warming is based on a belief that things are different now. So different, we are told, that the Earth is at the tipping point of accelerating temperature increases, and that there is no time to lose (Hansen 2007; Vermeer and Rahmstorf 2009; Hanson *et al.* 2013).

To evaluate the claim that recent data display an unprecedented pattern and that the pattern is indicative of a turning point, we asked forecasters at the 2012 International Symposium on Forecasting in Boston to forecast the next 25 years of two similar-looking 50-year sequences of monthly global mean temperatures from the Hadley data series (Figure 1). We told the forecasters that both sequences occurred during the age of industrialism. In other words, both coincided with exponential increases in atmospheric carbon dioxide.

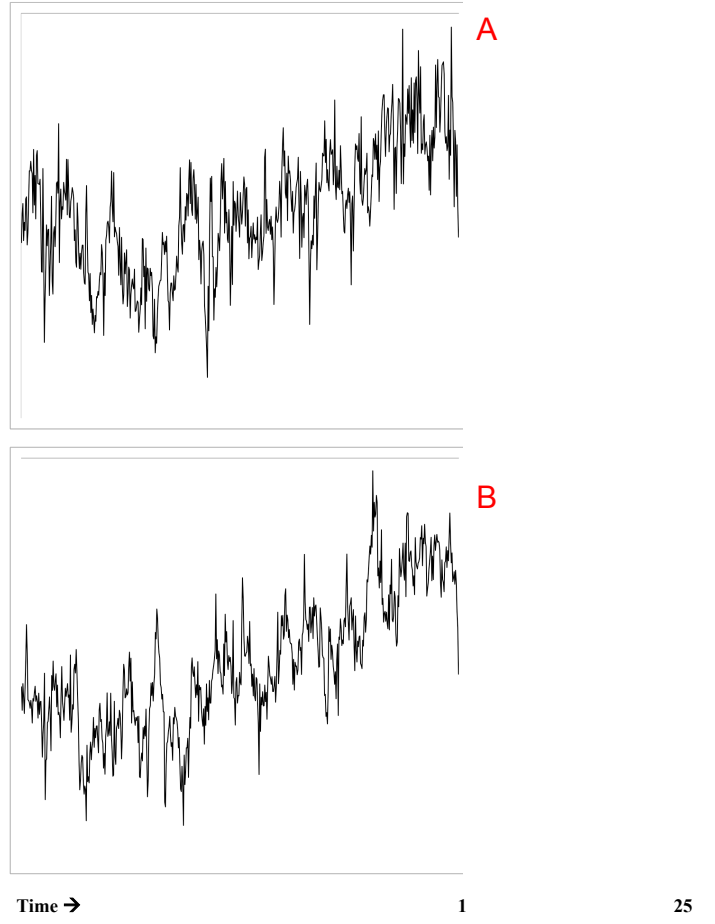
⁴ This changed when the site was repeatedly attacked by unknown hackers. It is currently risk free.

Figure 1: Forecasting task using Hadley (HadCRU3) global mean temperatures

**Test your forecasting skills:
Print this page and draw in your forecasts**

Monthly global mean temperatures over two half centuries*

Draw in your forecasts for the next 25 years for both charts.

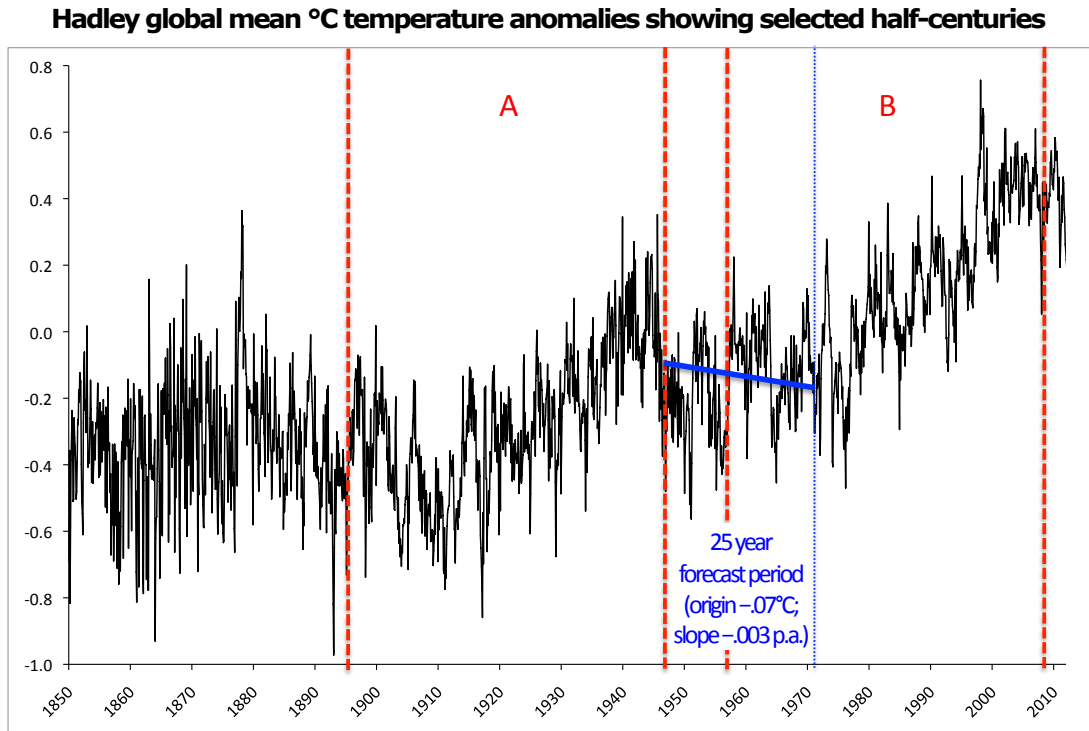


**Both during industrial era*

Nearly half (23 of the 51) respondents made forecasts that were consistent with carbon-dioxide causality, while another 27 forecast little or no trend over the forecast period. Figure 2 shows the data in context and a trend line for the 25-year period that the forecasters were asked to forecast and for which data are available. The similarity of the two sequences is inconsistent with the hypothesis that climate changes are different now than in earlier times.⁵

⁵ The idea behind this comparison came from a presentation by Meyer (2009).

Figure 2: Forecasts for the two similar graphs

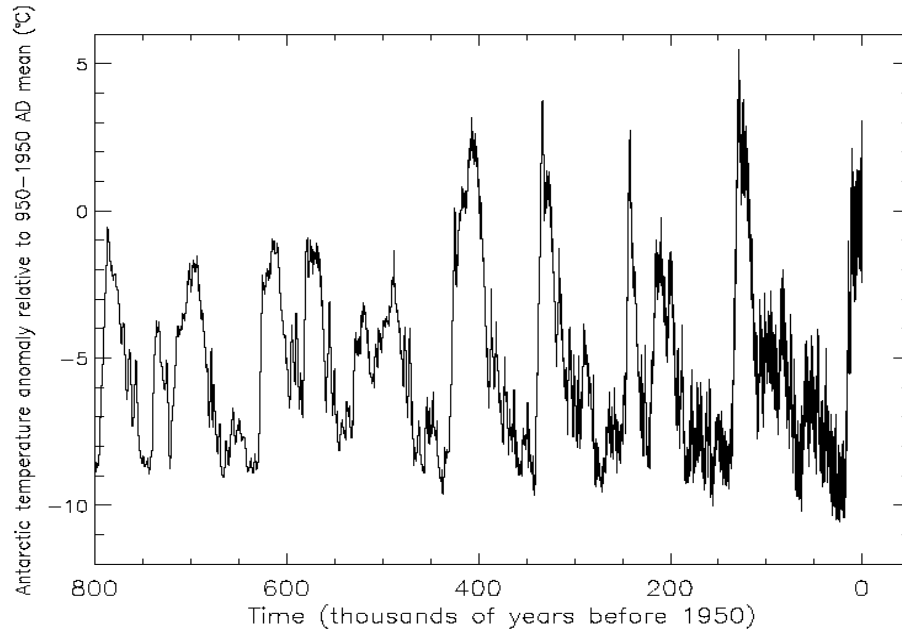


We are not aware of any strong evidence that the present is different from the past. We do recognize that opinions on this vary. For example, some people think there are changes that will lead to global warming while others believe there are changes that will lead to global cooling. Moreover, the levels of atmospheric carbon dioxide have been higher in the past (Kouwenberg *et al.* 2005).

Global warming alarmists have pointed to a correlation of 0.86 between atmospheric CO₂ and global mean temperature for the period 1850 to 2006 to support their contention that CO₂ has a strong effect on temperature. However, as is well known to those who study forecasting methods, correlations for non-experimental data provide a weak basis for causality. Indeed, even those untrained in forecasting often agree that correlation does not imply causality. For example, for the same period, the correlation between U.S. Postal rates and the Hadley temperature series was 0.84. And for the period 1970 to 2006, the correlation between National Oceanic and Atmospheric Administration (NOAA) expenditures and temperatures was 0.82.

We then looked at long temperature series (see, e.g., Figure 3 from Jouzel *et al.* 2007). In the absence of human influence (1) mean temperatures have changed substantially over long periods of time, (2) trends appear to reverse at irregular intervals, and (3) there is no underlying long-term trend. See Figure 3. There seems to be nothing unique about the recent history.

Figure 3: 800,000-year Record of Antarctic Temperature Change



Testing forecast validity

To date, there is only one peer-reviewed journal publication that claims to have made forecasts of long-term climate change. This paper, Green, Armstrong and Soon (2009), referred to as GAS below, developed and tested a model that, with minor exceptions, adhered to scientific forecasting principles. The exceptions were the use of data that were subject to potential biases, and that there was no combining of forecasts from alternative evidence-based methods. Based on the description of the situation, we hypothesized that the no-trend model was the most appropriate to use in this situation. When tested in the past, this simple model has been difficult to beat in complex and uncertain situations.

We compared the accuracy of the no-trend model with the IPCC's 1990 and 1992 forecasts of "future warming at rates of about 0.3°C/decade over the next century" (p. 17, Houghton, Callander, and Varney 1992). To do so, we produced no-trend forecasts from one-to-100 years ahead starting from 1851 and stepping forward year-by-year until 1975, the year before the current warming alarm was raised. We did the same for the IPCC scenario of temperatures increasing at a rate of 0.03°C per year in response to exponentially increasing human carbon dioxide emissions. This yielded 7,550 forecasts for each method,

The results? Overall, the no-trend forecast error was one-seventh the error of the IPCC scenario temperatures. The no-trend forecasts were as accurate or more accurate than the IPCC scenario temperatures for *all* forecast horizons. Crucially, the relative accuracy of the no-trend forecasts increased for longer horizons—for example, the no-trend forecast error was one-twelfth that of the IPCC scenario temperatures for 91-to-100-year ahead forecasts. The GAS study also examined the small sample of errors from *ex ante*

forecasts for 1992 through 2008. There was no difference in accuracy between the no-trend and IPCC forecasts for these short horizons.

Extensions of the GAS model

As noted in GAS, there were some shortcomings in our procedures. Given the small errors in the GAS forecast (e.g., mean absolute errors for 20- and 50-year horizons were 0.18°C and 0.24°C respectively), we concluded that there was little to be gained from further improvements.

In the only other study to examine the application of scientific forecasting methods to climate, Fildes and Kourentzes (2011) (denoted below as F&K) derived forecasts of global mean temperatures for horizons up to 20 years ahead. This extension was important for two reasons. First, do they obtain forecasts that support the IPCC forecasts? And, second, might some of their methods be useful for the longer-term forecasts?

Given that they followed many forecasting principles, it is no surprise that their forecasts were much more accurate than the IPCC scenario. The errors from the least accurate of F&K's six forecasting methods were less than half the IPCC forecast errors for forecast horizons of 1 to 4 years, and also for the 10-year ahead forecasts. As with our study, their improvements were achieved using inexpensive methods.

F&K's forecasts up to ten-years ahead were slightly more accurate than those presented in GAS. Over the same horizons, the GAS no-change forecast errors were about one-third less than those from the IPCC scenario. F&K followed evidence-based principles as described in Armstrong (2001). These include using simple methods for a complex and uncertain situation, using reasonable alternative methods, considering that different methods might be relevant for short-term than for long-term forecasting, combining forecasts from evidence-based methods, conducting validation tests on data unknown to the models.⁶ In doing so, they used valid and relevant error measures, provided full disclosure, and responded to questions about their research.

F&K tested the no-change forecasting model against six other alternatives to the IPCC's General Circulation Models, one GCM (Smith et al.'s 2007 Decadal Climate Prediction System referred to as DePreSys in F&K), and a combined forecast using the seven non-GCM methods. (F&K refer to the naïve model as the random walk model or RW, and we refer to it as the "no-trend model" in this paper.) To test the forecasting models, F&K used a version of the same temperature series that was used in GAS (HadCRUT3)⁷.

The no-trend model GAS used ignored important temporary influences on climate such as volcanic eruptions and El Nino weather patterns. Thus, we expected, that forecasts from the F&K models would be relatively more accurate than forecasts from the GAS naïve model across the short horizons examined by F&K.

⁶ In the case of F&K, the modelers knew the data used for validation tests. We discuss this issue below.

⁷ HadCRUT3 has at the time of writing been superceded by HadCRUT4. We use HadCRUT4 in this paper, except where the earlier series are needed for consistency. The series all commence in 1850 and consist of adjusted composites of manual temperature readings from selected weather stations and ships-

F&K did not propose which method they expected to provide the most accurate forecasts, so we compared forecast errors from the naïve model with the average of the errors from the six F&K alternatives that provided sufficient forecasts.

To make the comparison, we summed all 176 absolute errors for each method's forecasts across the 1–4, 10, and 20-year time horizons that F&K used (66 observations from their Table 1, and 60 and 50 observations from Table 2; F&K p. 984 and p. 985). Surprising to us, the sum of the absolute errors from the GAS no-trend model forecasts were 5% *smaller* than the average of the six F&K methods' sums of absolute errors.

F&K subsequently provided us with all the forecasts that they described in their paper: Namely, 1,190 forecasts for horizons from one to twenty years from each of their six methods. For the full set of forecasts, GAS no-trend model errors were 2% smaller than the average of the six F&K methods' errors.⁸ Forecasts from F&K's models for their longer, 11 to 20 year, horizons were nearly 8% less accurate than the equivalent forecasts from the no-trend model.

Extensions for improving long-term climate forecasting

We considered possible improvements to long-range climate forecasting including the use of causal forces and combining forecasts.

Inclusion of causal factors: Scientists dispute the nature and importance of causes of global mean temperature changes. The models used by the IPCC assume that manmade changes in CO₂ and other trace gases in the atmosphere are major drivers of temperature and hence of climate changes (Randall *et al.* 2007). Other researchers dispute that assumption, concluding there is no direct evidence of any substantive human effect on global climate (e.g., Beenstock *et al.* 2012). Leading climate scientists propose other sources for climate change such as variations in solar radiation reaching the Earth (e.g., Soon *et al.* 2011), unexplained variations in cloudiness (Lindzen and Choi 2011), and variations in ocean and atmospheric currents (Wunsch 2010). In contrast, there is little dispute that major volcanic eruptions can lead to cooling over several years (see e.g., Breitenmoser *et al.* 2012), or that the Sun is an important influence on climate change (see e.g., Soon and Legates 2013).

Combining forecasts is one of the most powerful concepts in forecasting. Combining improves accuracy under almost all conditions and it reduces the risk of large errors. Under ideal conditions it can reduce cut forecast errors in half (Graefe, et al. 2014). Conditions for combining are, however, poor in climate forecasting due to the scarcity of evidence-based forecasts.

We tested the accuracy of an average of the combination of forecasts from F&K's six methods. Bearing in mind that F&K's methods were applied to a small sample that was known to trend generally upward, we expected forecasts that incorporated a small trend would perform better than the no-trend forecasts for that period.

⁸ Tables summarising our analysis of the F&K forecasts are available at <http://kestengreen.com/gas2012.pdf>.

And so it did. The cumulative relative absolute error (CumRAE from Armstrong and Collopy 1992) for the resulting 1,190 “conservative trend” forecasts was .922 (the no-trend model score at 1.0). The figure represents an error reduction of almost 8% relative to the no-trend forecast, and a reduction of 10% relative to the forecasts from the average of F&K’s six models.

F&K’s autoregression (AR) models⁹ are also conservative trend models. On examining the forecasts F&K provided, we found the models forecast trends of between 0.001°C and 0.007°C, with a mean of 0.0037°C or 0.37°C per century. That is roughly half the 0.70°C per century rate at which the Hadley data increased during the 1914 to 2006 period over which the models were variously estimated, and one-eighth of the IPCC’s projected warming rate. Moreover, the AR models trends were inconsistent with the claim of *accelerating* warming: models estimated with more recent data did not forecast faster warming than those estimated with older data.

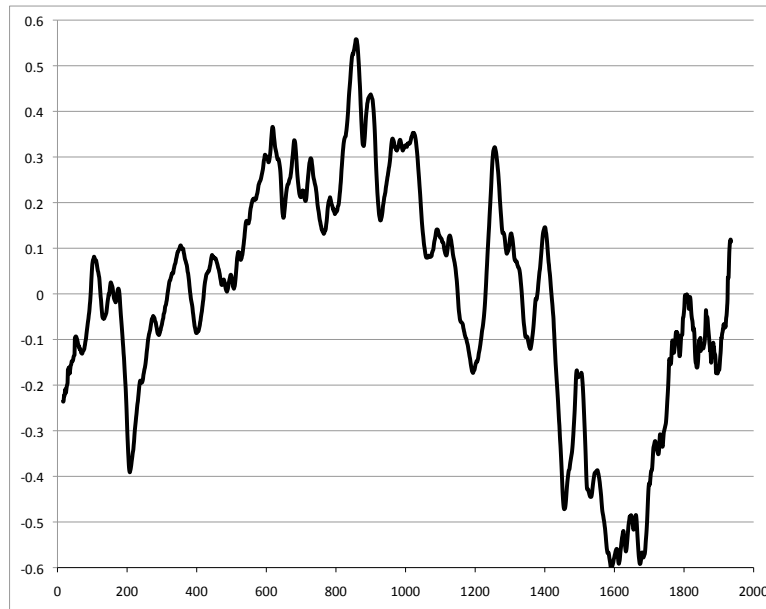
Validate methods with more and longer-term forecasts

When the number of years in the forecast horizon exceeds that in the historical period, uncertainty increases. To properly test methods for long-term forecasting, then, long-term validation samples are needed. There are currently only 21 years of truly out-of-sample forecasts from the IPCC models for which actual values are available. In GAS, we got around this validation problem for the IPCC scenarios by applying the IPCC’s forecast that exponentially increasing atmospheric CO₂ levels would cause global mean temperatures to increase by 0.03°C per year to forecast the Hadley temperature series from 1851. (The period was one of exponentially increasing industrialization and energy use and is thus analogous to the IPCC’s 0.03°C per year business-as-usual scenario.) We suggest that other researchers also use the *full* Hadley data set to test their forecasts versus other approaches. This approach provided 10,750 forecasts for 1-year to 100-year horizons. In contrast, F&K’s validation tests were based on 1,190 forecasts for shorter (1–20 year) horizons from each model. While it is well known that the Hadley data series trends broadly upward, the full series does include downward trending periods, and so it provides a stronger test.

Using the full Hadley data set does not entirely avoid the problem of lack of out-of-sample data because forecasters are aware of the upward trends in both CO₂ concentrations and (very broadly) temperatures over the period of the series. This is contrary to the principle that one should “test all models for performance with data not used in estimation” (p. 340, Allen and Fildes 2001). To better avoid the problem, we suggest using temperature series that are much longer than the Hadley series, such as the roughly 2000-year Loehle and McCulloch (2008) series (Figure 4), for validating long term forecasting methods.

⁹ Autoregression models forecast the next period of a time series using earlier values of the series.

Figure 4: 2000-year record of global temperature change¹⁰



Adjust validation data for known biases, and use alternative validation data

Studies of the Hadley data series have found biases. These include the heat island effect, the reduction of the number of weather stations in isolated areas, poor maintenance of weather stations, and the increasing use of electronic thermometers, which can record very short duration highs that can arise from heat eddies (Watts *et al.* 2012). In some cases, recorded temperatures had been adjusted without explanation.

The biases all appear to be in the direction of producing a warming trend. One analysis led to the conclusion that the actual increase in temperatures from 1979 to 2002 was half as large as the increase in the Hadley data (McKittrick and Michaels 2007; McKittrick and Nierenberg 2010; Watts *et al.* 2012).

Researchers in New Zealand constructed an unadjusted temperature series using data from the same weather stations that are used in the pre-2010 official (adjusted) temperature series. (The official series is used in the construction of the Hadley temperature series.) Prior to their 2010 review, the National Institute of Water and Atmospheric Research (NIWA) provided neither a record of the adjustments that were made in order to construct the official series, nor the reasons for them. The official series has an increasing trend of 1.0°C per century, whereas the unadjusted has a trend of 0.3°C per century.

The lower troposphere temperature data from the University of Alabama at Huntsville (UAH) is the most valid of the global mean temperature series currently available. Satellite-mounted instruments are used to

¹⁰ We followed the conservative presentation of findings in Loehle and McCulloch (2008) by showing in Figure 3 the smoothed paleo-temperature series available from 16 AD to 1935 AD only. There is an inherent unresolvable question related to how one can connect these paleo-temperature series to a series based on thermometer readings (see Soon *et al.* 2003), not the least of which is how to correct for the non-climatic contamination of the thermometer records. We are nevertheless satisfied of the robustness of Loehle and McCulloch's (2008) finding that "The peak value of the MWP [Medieval Warm Period] is 0.526 Deg C above the mean over the period (again as a 29 year mean, not annual, value). ... While instrumental data are not strictly comparable, the rise in 29 year-smoothed global data from NASA GISS ... from 1935 to 1992 (with data from 1978 to 2006) is 0.34 Deg C. Even adding this rise to the 1935 reconstructed value, the MWP peak remains 0.07 Deg C above the end of the 20th Century values..." (pp. 97-98).

detect microwave emissions. This method avoids contamination from near-surface structures and other non-climatic influences (Christy *et al.* 2010; McNider *et al.* 2012).

Unfortunately, the series started only in late-1978, and so it is not sufficient to adequately test long-term forecast accuracy. Nevertheless, the data presents an opportunity to compare the relative accuracy of forecasts from F&K's six models and from the GAS naïve model against the IPCC model.

We did not have access to the F&K models to derive forecasts based on the UAH data. To get around the problem, for each model's forecasts we compared the change from the previous year with the actual UAH changes in order to derive forecast errors. We expected that F&K's models would have an advantage in forecasting out to 10 and perhaps 20 years, but that the no-trend model would be a strong contestant for longer horizons when uncertainty is greater.

The first UAH annual change that can be calculated is for 1980 and the last F&K forecast was for 2007, there are 532 F&K forecasts per model for horizons from 2 to 10 years over this period. We found that the average absolute errors from the F&K model forecasts, expressed as annual temperature changes, were 1.1% smaller than the equivalent GAS no-trend model forecasts. The 190 errors from the F&K model forecasts for F&K's longer horizons (11 to 20 years), however, were 0.5% larger than the comparable GAS no-trend model forecast errors.

Damp the forecasted trends to compensate for the complexity and uncertainty of the situation

Complexity and uncertainty are high for climate forecasting due to problems in obtaining valid and reliable estimates of important climate variables. This problem increases as the forecast horizon lengthens. Damping often improves accuracy.

To test the effect of damping, we damped the forecasts from each of the F&K models progressively, starting with the 11-year horizon by using the formula $F_m \text{ damped} = F_m + [(F_N - F_m) / (21 - h)]$; where F_N is the naïve GAS no-trend model forecast and h is the horizon from 11 to 20. This, the trend disappear after year 20. The damped F&K forecast errors were nearly 1% smaller than the naïve model forecast errors; an improvement in accuracy of nearly 3% relative to the original F&K forecasts. We made no attempt to optimize the damping procedure.

Further research

Given the remarkable accuracy of the no-trend model, it is difficult to see how improvements in accuracy can provide any practical benefit. Nevertheless, should curiosity or new information spur researchers to further work on this problem, they should rely solely on the use of validated evidence-based forecasting procedures. To our knowledge, no area is immune from these principles.

Developed from studies conducted over more than half a century, the forecasting principles have stood for more than a decade with only one major revision and one minor one. We have heard claims that the principles are not relevant to climate forecasting, but our appeals for information as to which of the principles do not apply to climate and as to where climate-forecasting principles have been published remain unanswered.

Policy implications

In 1896 the Swedish Nobel Prize winner in chemistry, Svante Arrhenius, speculated that increases in atmospheric CO₂ from industrial activity would cause warming (Baliunas and Soon 1999). Arrhenius's idea attracted little attention at the time, possibly because he assumed that global warming was desirable.

In the 1960s, experts predicted human emissions of soot and aerosols would more than overpower the effects of increases in atmospheric CO₂ thereby threatening the world with dangerous manmade global cooling (Rasool and Schneider 1971). Calls of alarm that the world was in danger of entering a new ice age faded when the declining trend in global mean temperatures stalled.

As with the cooling alarm, temperatures have not changed in accord with the dangerous warming hypothesis. Despite the lack of forecasting skill, governments have implemented some of the expensive policies that have been advocated by the alarmists. Why?

Experiments by Sunstein and Zeckhauser (2011) found that when people fear that a very bad thing might happen, such as a five-meter sea level rise, they ignore probabilities. They also seek to blame others, such as large corporations or foreigners and, with little regard for the cost, demand that government solve the problem. In contrast, individuals are reluctant to spend their own money to make their fear go away. These normal human biases lead to voting and policy decisions that reduce welfare.

Such behavior is not unusual. In our study of situations that are analogous to the current alarm over scenarios of dangerous manmade global warming, we identified 26 earlier movements based on scenarios of manmade disaster, including the global cooling alarm in the 1960s. None of them were based on scientific forecasts. Yet governments imposed costly policies in response to 23 of them. In no case did the forecast of major harm come true (Green and Armstrong 2011).

Despite the lack of scientific support for climate policies, vast resources have already been devoted to policies that are intended to stop predicted dangerous manmade global warming (e.g., Hansen and Sato 2012). This situation can be explained by the motivations of the parties involved and the incentives they face. Governments and interest groups are attracted to opportunities for increasing tax revenues and for allocating that revenue to favored projects and causes, supporters, and important voting blocks. Businesses that can arrange their affairs to take advantage of government policies stand to make profits. Many individual citizens, on the other hand, may not see that the new taxes and a reduction in overall efficiency of the economy affect them, although it is often obvious to those who lose jobs.

F&K's forecasts have reinforced the findings in Green and Armstrong (2007) and GAS that the hypothesis of dangerous warming over the 21st Century and beyond should be rejected. The additional analyses in this current paper provide additional evidence for that conclusion.

Climate policy rests on a three-legged stool. First, it is necessary that valid and reliable scientific forecasts show a strong, consistent and causally based trend in temperatures. Given that this has not happened for over 800,000 years, this seems remote. Second, it is necessary that valid and reliable

scientific forecasts show that the net effects of the trend in temperatures will be very harmful. Nor has it been shown that the warming over the past century and a half has been harmful. Third, it is necessary that valid and reliable scientific forecasts show that each proposed policy (e.g., polar bears protection measures; Armstrong, Green and Soon 2008) would provide a substantive net benefit relative to taking no action. A failure of any leg of the stool invalidates policy action. As things stand now, there are no scientific forecasts to support any of the three legs.

Some people believe it is reasonable to make policy on the basis that a vividly imagined harmful event might occur, and call this belief the “precautionary principle”¹¹. This is an anti-scientific belief that decisions should be made by judgment rather than by scientific study. The also defies logic; for example, in the case of climate change the precautionary principle can also be used to advocate policies in response to possible dangerous global cooling¹².

Properly forecasting the net effects on welfare of climate changes and of policy proposals relative to decentralized adaption would be challenging. For example, people in many places would benefit from a warmer climate and higher levels of CO₂ if they were left to adapt as necessary and as they saw fit to any local changes that might occur.¹³ It is hard to imagine that people, given their diverse local climates, circumstances, and preferences could possibly agree on an optimum global mean temperature and on how much they would be willing to pay for it. It is even harder to imagine agreement when there is uncertainty over whether such an optimum could ever be achieved in practice.

Conclusions

In GAS, we concluded that forecasts from the no-trend model are so accurate for long-term forecasts (up to 100 years), that there would be no economic benefit from improving the accuracy of the forecasts.

We found little evidence to suggest that things are different now in climate forecasting. While some expert will disagree, the evidence shows that experts’ opinions as well as their confidence are unrelated to the accuracy of their forecasts for complex and uncertain situation.

Attempts to improve upon the no-change model suggested small gains in accuracy might be possible for short-term forecasts of up to ten years. Unfortunately, this has no value for policy analysis. Also, unfortunately, the evidence-based methods that provided some help for the short-term forecasts did not more accurate than the no-change methods for forest horizon beyond 10 years.

Responsible policy makers need scientific long-range forecasts of global warming. Then they would need scientific forecasts of the effects of alternative feasible policies and of decentralized adaption. Only then would it be possible to conduct proper cost-benefit analyses in order to identify policy alternatives that could

¹¹ See Green and Armstrong’s 2008 essay “Uncertainty, the Precautionary Principle, and Climate Change” at <http://kestencgreen.com/uncertainty-precautionary.pdf>

¹² See, for example, ‘Russian Lake Vostok Scientists Say “New Ice Age is Unavoidable”’ at <http://notrickszone.com/2012/04/27/russian-lake-vostok-researchers-say-ice-age-is-unavoidable/>

¹³ See, for example, Robinson, Robinson, and Soon (2007) on the beneficial effects of elevated levels of atmospheric CO₂.

reasonably be expected to substantially improve human welfare. In other words, policy changes rely on a three-legged stool of forecasts, and currently none of the legs provide any support.

We recommend rescinding government funding and regulations related to the dangerous manmade global warming alarm until such time as they can be supported by cost/benefit analyses based on unbiased, independent, audited scientific forecasts. In other words, governments should follow the advice to “Don’t just do something, stand there!”

Acknowledgments

We thank Robert Fildes and Nikolaos Kourentzes for supplying data and for useful suggestions on the paper. Donald J. Peters provided a review. We presented the paper in a panel session at the International Symposium on Forecasting in Boston, MA, in 2012, and are grateful for the comments and suggestions we received from attendees.

References

- Allen, P.G. & Fildes, R. (2001). Econometric forecasting. In Armstrong, J.S. (ed.) *Principles of Forecasting*, Kluwer: Boston. 303–362.
- Armstrong, J. S. & Collopy, F. (1992). Error measures for generalizing about forecasting methods: empirical comparisons. *International Journal of Forecasting*, 8, 69–80
- Armstrong, J. S., Green, K. C., & Soon, W. (2011). Research on forecasting for the manmade global warming alarm. *Energy and Environment*, 22, 1091-1104.
- Armstrong, J. S., Green, K. C., & Soon, W. (2008). Polar bear population forecasts: A public-policy forecasting audit. *Interfaces*, 38, 382-404.
- Baliunas, S. & Soon, W. (1999). Pioneers in the greenhouse effect, *World Climate Report*, 4 (no. 19, June 14 issue), 1-5. (available http://www.worldclimaterreport.com/archive/previous_issues/vol4/v4n19/cutting.htm).
- Beenstock, M., Reingewertz, Y., and Paldor, XN (2012), Polynomial cointegration tests of anthropogenic impact on global warming, *Earth System Dynamics*, 3, 173-188.
- Breitenmoser, P., Beer, J., Bronnimann, S., Frank, D., Steinhilber, F., & Wanner, H. (2012). Solar and volcanic fingerprints in tree-ring chronologies over the past 2000 years. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 313-314, 127–139.
- Christy, J. R., Herman, B., Pielke, R., Sr., Klotzbach, P., McNider, R. T., Hnilo, J. J., Spencer, R. W., Chase, T., & Douglass, D. (2010). What do observational datasets say about modeled tropospheric temperature trends since 1979? *Remote Sensing*, 2, 2148-2169.
- Fildes, R. & Kourentzes, N. (2011). Validation and forecasting accuracy in models of climate change. *International Journal of Forecasting*, 27, 968–995.
- Graefe, A., Armstrong, J.S. Cuzan, A. & Jones, R.J. (2014). [Combining forecasts: An application to elections](#),” *International Journal of Forecasting*, 30, 43-54.
- Green, K. C. & Armstrong, J. S. (2007). Global warming: forecasts by scientists versus scientific forecasts. *Energy and Environment*, 18, No. 7+8, 995–1019.
- Green, K. C., Armstrong, J. S. & Soon, W. (2009). Validity of Climate Change Forecasting for Public Policy Decision Making. *International Journal of Forecasting*, 25, 826–832.
- Green, K. C., & Armstrong, J. S. (2011), The global warming alarm: Forecasts from the structured analogies method. Working Paper.
- Hansen, J.E. (2007) Scientific reticence and sea level rise, *Environmental Research Letters*, 2, doi:10.1088/1748-9326/2/2/024002.
- Hansen, J.E., and Mki. Sato, 2012: Paleoclimate implications for human-made climate change. In Climate Change: Inferences from Paleoclimate and Regional Aspects. A. Berger, F. Mesinger, and D. Šijački, Eds. Springer, pp. 21-48, doi:10.1007/978-3-7091-0973-1_2.
- Hansen, J.E., Mki. Sato, G. Russell, P. Kharecha, 2013: Climate sensitivity, sea level and atmospheric carbon dioxide. *Philosophical Transactions of The Royal Society A*, 371, 20120294.
- Houghton, J. T., Callander, B. A., and Varney, S. K. (eds.) (1992). *Climate Change 1992: The supplementary report to the IPCC scientific assessment*. Cambridge, U.K.: Cambridge University Press for Intergovernmental Panel on Climate Change.

- IPCC (2007). *Climate Change 2007: Synthesis Report, Contribution of Working Group I, II and III to the Fourth Assessment Report to the Intergovernmental Panel on Climate Change*. Edited by Core Writing Team, Pachauri, R.K. and Reisinger, A., IPCC Geneva, Switzerland. [Available from: http://www.ipcc.ch/publications_and_data/publications_ipcc_fourth_assessment_report_synthesis_report.htm]
- Jouzel, J., and 31 co-authors (2007). Orbital and millennial Antarctic climate variability over the past 800,000 years. *Science*, 317, 793–796.
- Lindzen, R.S., and Choi, Y.-S. (2011). On the observational determination of climate sensitivity and its implications. *Asia-Pacific Journal of Atmospheric Sciences*, vol. 47, 377–390.
- Loehle, C., & McCulloch, J. H. (2008). Correction to: A 2000-year global temperature reconstruction based on non-tree ring proxies. *Energy & Environment*, 19, 93–100.
- McKittrick, R.R., and Michaels, P.J. (2007) Quantifying the influence of anthropogenic surface processes and inhomogeneities on gridded global climate data. *Journal Geophysical Research*, vol. 112, doi:10.1029/2007JD008465.
- McKittrick, Ross R. and Nicolas Nierenberg (2010). Socioeconomic Patterns in Climate Data. *Journal of Economic and Social Measurement*, 35(3,4) pp. 149–175.
- McNider, R.T., Steeneveld, G.J., Holstag, A.A.M., Pielke Sr., R.A., Mackaro, S., Pour-Biazar, A., Walters, J., Nair, U., and Christy, J.R. (2012) Response and sensitivity of the nocturnal boundary layer over land to added longwave radiative forcing. *Journal Geophysical Research*, vol. 117, doi:10.1029/2012JD017578.
- Meyer, W. (2009). Catastrophe denied: A critique of catastrophic man-made global warming theory. Annotated slides from a presentation in Phoenix, AZ, November 10. [Available from <http://www.climate-skeptic.com/Climate%20Presentation%20Annotated%201-1-2010.pdf>]
- Michaels, P. J. (2009). The dog ate global warming. *National Review Online*, September 23, <http://www.nationalreview.com/articles/228291/dog-ate-global-warming/patrick-j-michaels>
- New Scientist (2007). Climate Catastrophe, 28 July issue, 30–34.
- New Zealand Climate Science Coalition (2011). Statistical Audit of the NIWA 7-Station Review. The New Zealand Climate Science Coalition [Available from http://nzclimatescience.net/images/PDFs/main_rept_on_niwa_7ss_review%20v5.pdf]
- Randall, D.A., Wood, R.A., Bony, S., Colman, R., Fichetef, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R. J., Sumi, A. and Taylor, K.E. (2007). Climate Models and Their Evaluation, in Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M. and Miller, H.L. 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.
- Rasool, S. I., & Schneider, S. H. (1971). Atmospheric carbon dioxide and aerosols: Effects of large increases on global climate. *Science*, 173 (3992), 138–141.
- Robinson, A. B., Robinson, N. E., & Soon, W. (2007). Environmental effects of increased atmospheric carbon dioxide. *Journal of American Physicians and Surgeons*. 12, 79–90.
- Soon, W., Dutta, K., Legates, D.R., Velasco, V. and Zhang, W. (2011) Variation in surface air temperature of China during the 20th Century. *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 73, 2331–2344.
- Soon, W. and Legates, D.R. (2013) Solar Irradiance Modulation of Equator-to-Pole (Arctic) Temperature Gradients: Empirical Evidence for Climate Variation on Multi-decadal Timescales. *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 93, 45–56.
- Sunstein, Cass R. & R. Zeckhauser (2011), Overreaction to fearsome risks, *Environmental and Resource Economics*, 48, 435–449.
- Tetlock, P.E. (2005). *Expert political judgement*. Princeton.
- Turner, J., Bracegirdle, T., Philips, T., Marshall, G.J., and Hosking, J.S. (2012), An initial assessment of Antarctic sea ice extent in the CMIP5 models, *Journal of Climate*, in press, doi:10.1175/JCLI-D-12-00068.1.
- USGCRP (2009) *Global Climate Change Impacts in the United States*, Thomas R. Karl, Jerry M. Melillo, and Thomas C. Peterson (eds), Cambridge University Press.
- Washington Post Editorial Board (2013). A sweltering planet's agenda. January 13.
- Vermeer, M. and Rahmstorf, S. (2009), Global sea level linked to global temperature, *Proceedings of the National Academy of Sciences*, 106, 21527–21532.
- Watts, A., Jones, E., McIntyre, S., and Christy, J.R. (2012) An area and distance weighted analysis of the impacts of station exposure on the U.S. Historical Climatology Network temperatures and temperature trends. PREPRINT DRAFT DISCUSSION PAPER [available for open peer review here: http://wattsupwiththat.files.wordpress.com/2012/07/watts-et-al_2012_discussion_paper_webrelease.pdf]
- Wunsch, C. (2010) Towards understanding the Paleocene. *Quaternary Science Reviews*, vol. 29, 1960–1967.