

Climate Data vs. Climate Models

Why do climate change assessments overlook the differences between the climate models they use and the empirical data?

BY PATRICK J. MICHAELS AND PAUL C. KNAPPENBERGER

The U.S. Global Change Research Act of 1990 created something called the U.S. Global Change Research Program (USGCRP), a 13-agency entity charged with conducting, disseminating, and assessing scientific research on climate change. Every four years, the program is supposed to produce a “national assessment” of climate change. One of the purposes of the assessment is to provide the Environmental Protection Agency with information that it can use to regulate carbon dioxide.

The first of these assessments was released in the fall of 2000. Despite the four-year mandate, none were produced during the George W. Bush administration. Finally, in 2009, a second assessment was published. While the 2009 report is still the document of record, a draft of the third report, scheduled for publication late this year, was circulated for public comment early this year. Unfortunately, none of the assessments provide complete and comprehensive summaries of the scientific literature, but instead highlight materials that tend to view climate change as a serious and emergent problem. A comprehensive documentation of the amount of neglected literature can be found in a recent Cato Institute publication, *Addendum: Global Climate Change Impacts in the United States*.

2000 Assessment

PATRICK J. MICHAELS is director of the Cato Institute’s Center for the Study of Science. PAUL C. “CHIP” KNAPPENBERGER is the center’s assistant director.

In the 2000 assessment, known officially as the “U.S. National Assessment of Climate Change,” the USGCRP examined nine different general circulation climate models (GCMs) to assess climate change impacts on the nation. They chose two GCMs to use for their projections of climate change. One, from the Canadian Climate Center, forecasted the largest temperature changes of all models considered, and the other, from the Hadley Center in the United Kingdom, forecasted the largest precipitation changes.

The salient feature of those models is that they achieved something very difficult in science: they generated “anti-information”—projections that were of less utility than no forecasts whatsoever. This can be demonstrated by comparing the output of the Canadian Climate Center’s model with actual temperatures.

The top half of Figure 1 displays the observed 10-year smoothed averaged maximum temperature departures from the climatological mean over the lower 48 states from data through 1998. The bottom half displays the difference between the model projections over the same time period and those same



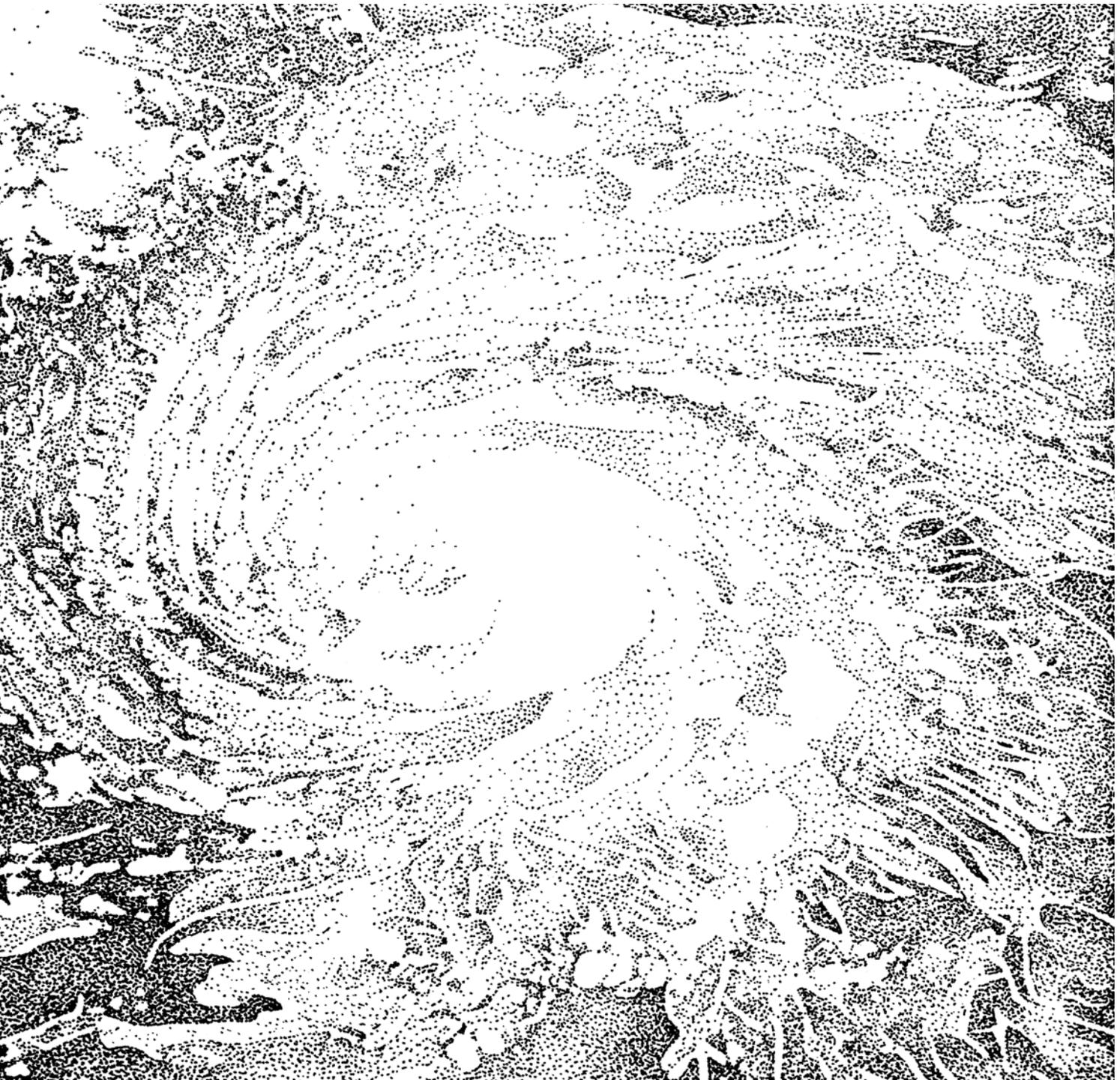


ILLUSTRATION BY MORGAN BALLARD

temperature observations. Statisticians refer to such differences as “residuals.” The figure includes the value of the variance—a measure of the variability of the data about its average value—for both the observed data and the climate model residuals. In this case, the residuals have over twice the variance of the raw data.

To analogize the relationship between the Canadian model and the empirical data, imagine the model output was in the form of 100 answers to a four-option multiple choice test. If the model simply randomly generated answers, within statistical limits it would provide the correct answer to about 25 percent of

the questions. But, analogously speaking, the Canadian model would do worse than random: it would only answer about 12 out of 100 correctly. That’s “anti-information”: using the model provides less information than guessing randomly.

We communicated this problem to Tom Karl, director of the National Climatic Data Center and the highest-ranking scientist in the USGCRP. He responded that the models were never meant to predict 10-year running means of surface average temperature. He then repeated our test using 25-year running means *and obtained the same result that we did*. But the assessment was issued unchanged.

2009 Assessment

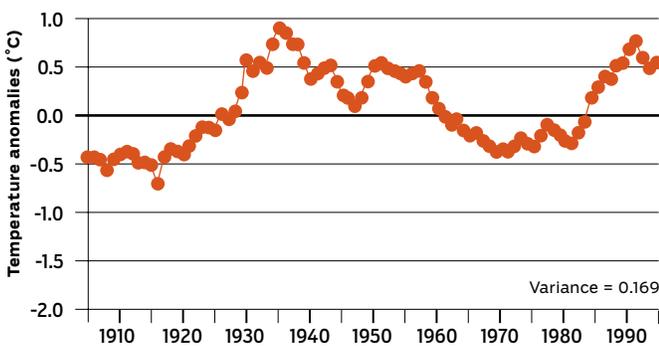
After a hiatus of nine years, the USGCRP produced its second national assessment, titled “Global Climate Change Impacts in the United States.” The environmental horrors detailed in the 2009 assessment primarily derive from a predicted increase in global surface temperature and changed precipitation patterns. But global surface temperatures have not increased recently at anywhere near the rate projected by the consensus of so-called “midrange emission scenario” climate models that formed much of the basis for the 2009 document. This has caused an intense debate about whether the pause in significant warming, now in its 16th year (using data from the Climate Research Unit (CRU) at the University of East Anglia—the data set that scientists cite the most), indicates that the models are failing because they are too “sensitive” to changes in carbon dioxide. That is, the models’ estimate of the increase in temperature resulting from a doubling of carbon dioxide concentration may simply be too high.

Figure 2 compares the trends (through the year 2012) in the CRU’s observed global temperature history over periods ranging

FIGURE 1
Actual Temperatures and the Canadian Climate Change Model

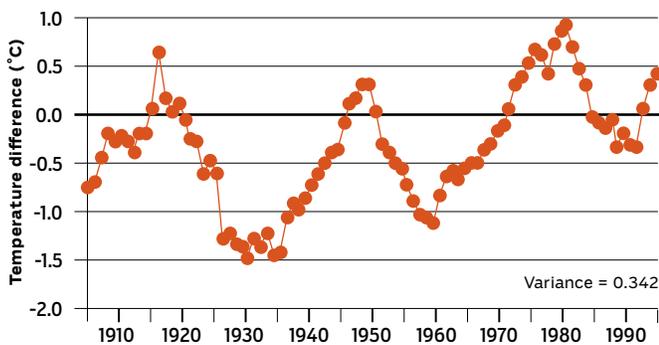
Observed Temperatures

Observed 10-year smoothed averaged maximum temperature departures from the recent climatological mean over the lower 48 states.



Model Projections Minus Observed Temperatures

Predicted minus observed averaged maximum temperature departures for the Canadian Climate Model that was the basis for the 2000 assessment.



from five to 15 years to the complete collection of climate model runs in the most recent climate assessment from the United Nations’ Intergovernmental Panel on Climate Change (IPCC). This suite of models is run under the midrange emissions scenario, which has tracked fairly accurately with real emissions rates, especially given the displacement of coal by cheaper abundant natural gas from shale formations worldwide.

The average forecast warming is approximately 0.2 degrees Celsius per decade, or 0.30 degrees Celsius over the entire 15-year period. The solid gray lines near the edges of the gray “trumpet” depict the 95 percent confidence interval surrounding the predicted trend of 0.2 degrees Celsius per decade, as different runs of different models produce different forecasts. As you can see, the actual trend has been around 0 rather than 0.2 degrees Celsius per decade, and for about the last five periods has been right at the edge of the 95 percent confidence interval. That is, we are 95 percent confident that the models are not consistent with the actual temperature record.

While there appears to be no actual trend in the average actual temperature over the last 16 years, there is some apparent “greenhouse” signal in the *distribution* of surface temperature change since the mid-1970s. As predicted, the Northern Hemisphere has warmed more than the Southern, and the high-latitude land regions of the Northern Hemisphere have warmed the most. That is consistent with climate physics, which predicts that dry air must warm more than moist air if carbon dioxide is increased. (Antarctica is an exception to this because of its massive thermal inertia.)

However, there is a major discrepancy between observed and forecast temperatures about seven to 10 miles in altitude over the tropics. (Remember that the tropics cover nearly half of the planetary surface.) While this very large block of the atmosphere is predicted to show substantially enhanced warming when compared to lower altitudes, it actually has not warmed preferentially. This is an extremely serious problem for the climate models because precipitation is largely determined by the temperature difference between the surface and higher altitudes.

With regard to precipitation, the 2009 assessment flouts the usual conventions of statistical significance. In the notes on maps of projected precipitation changes, the report states that “this analysis uses 15 model simulations.... [H]atching [on the map] indicates at least two out of three models agree on the sign of the projected change in precipitation,” and that this is where “[c]onfidence in the projected changes is highest.”

The USGCRP used 15 climate models, and the cumulative probability that 10 or more of the models randomly agree on the sign of projected precipitation change is 0.15. Scientific convention usually requires that the probability of a result being the product of chance be less than 5 percent (0.05) before the result is accepted, so this criterion does not conform to the usual conventions about “confidence.”

Another problem is that a “higher emissions scenario” (not the “midrange” one) was used to generate the report’s climate change maps. Because of the expected worldwide dispersal of

hydraulic fracturing and horizontal drilling for shale gas and the resulting shift toward natural gas and away from other fuels (particularly coal), the higher (carbon) emissions scenario is not appropriate for use in projections. Given that “highest confidence” areas of increased precipitation over the United States

The models’ estimate of the increase in temperature resulting from a doubling of carbon dioxide concentration may simply be too high.

are relatively small even in the “higher emissions scenario,” they are likely to be exceedingly small or perhaps nonexistent under the midrange one.

2013 Draft Assessment

On January 14, 2013, the USGCRP released a draft of its next assessment. The draft weighed in at over 1,200 pages, compared to the 190 pages of the 2009 version.

Unfortunately, the draft also has precipitation and temperature problems that aren’t dissimilar to those in the 2009 report.

Operationally meaningless precipitation forecasts | Precipitation changes will not exert discernible effects until they arise from the noise of historical data. If the time for this to occur is exceedingly long, the forecasts are useless in terms of utility. Should we really plan for slight changes forecast to occur 300 years in the future? What is the opportunity cost of acting based upon such a time horizon?

The 2013 draft precipitation forecasts look very similar to those in the 2009 version, except that the cross-hatched (“highest confidence”) areas now are where 90 percent or more of the model projections are of the same sign, instead of 67 percent. This new criterion does meet the 0.05 significance level based upon binomial probability.

Conspicuously absent is any analysis about when a signal of changed precipitation will emerge from the background noise. That is, given the predictions of the climate models, how long will it take to distinguish a “wetter” or “drier” climate from normal fluctuations in precipitation?

In order to analyze this, we examined the draft report’s state and seasonal combinations in which predictions were “confident.” We then calculated the number of years it would take, assuming linear climate change, for the average precipitation to reach 1 standard deviation (which contains approximately two-thirds of the observations) above or below the historical mean.

The projections in the draft are the average precipitation change for the period 2070–2099 (midpoint of 2085). There were

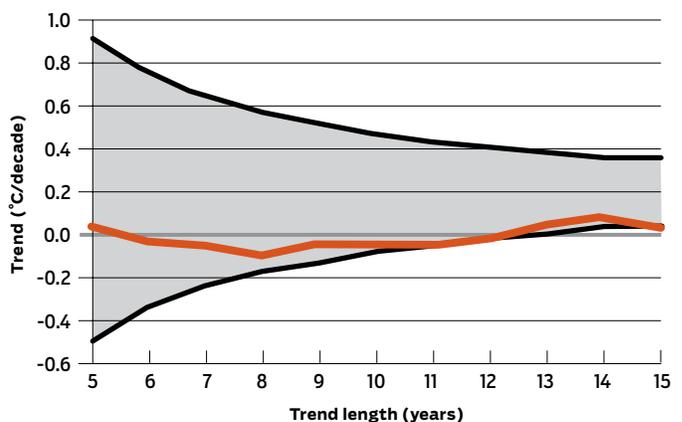
84 separate season-state combinations of “highest confidence.” In nine of those, the predicted change has already emerged from the noise (some 70 years ahead of time—which may or may not be good news for the models). Of those nine, eight are precipitation increases, with most occurring in the spring.

There were 75 remaining cases for which the observed change to date is currently less than the model projected when it estimated change for the 2070–2099 averaging period. During the summer—when rainfall is important to agriculture—the average time for the projected changes to emerge from the natural noise is 520 years. In the winter, it is 330 years. Averaged across all seasons, it will take approximately 297 years before a state’s projected seasonal precipitation changes emerge from background variability—and this is only for those state-season combinations where 90 percent of the models agree on sign. For the rest (vast majority) of the country, the climate models can’t even agree on whether it will become wetter or drier.

‘Sensitive’ citation problem | A convenient and important metric of climate change is the “sensitivity” of temperature to changes in greenhouse gases—generally, the amount of warming that results from an effective doubling of atmospheric carbon dioxide. High sensitivity, or a probability distribution of sensitivity that has “fat tails” at the high values, indicates a more urgent issue or a nontrivial probability of urgency. Lower sensitivity can reduce the effects of global warming to the degree that it is counterproductive to reduce emissions dramatically. Robert Mendelsohn, a Yale economist, has cal-

FIGURE 2
Actual Temperatures and the 2009 Assessment

Observed trend in global average surface temperature for trend lengths of 5–15 years ending in 2012 (red); average projected trend in global average surface temperatures from a collection of climate models (gray line) and the 95% confidence interval of the model trend projections (gray trumpet-shaped region).



culated that a net global warming of as much as 2.5 degrees Celsius confers net economic benefit. The change from net benefit to net cost occurs somewhere between 2.5 and 4.0 degrees Celsius.

The most serious problem with the 2013 draft is the failure to address the growing number of recently published findings that suggest that the climate sensitivity is much lower than previously estimated. This is evident in Figure 3, in which one can see the plethora of recent publications with lowered mean sensitivity and probability distributions where high-end warming is restricted compared to the models used by the IPCC in its most recent (2007) climate assessment.

The mean equilibrium sensitivity of the climate models in the 2013 IPCC draft assessment is 3.4 degrees Celsius. The mean equilibrium climate sensitivity from the collection of recent papers (illustrated in Figure 3) is close to 2.0 degrees Celsius—or about 40 percent lower. New

estimates of climate sensitivity and the implications of a 40 percent lower equilibrium climate sensitivity should have been incorporated in the new national draft report, but they were not.

Conclusion

The three national assessments suffer from common problems largely related to inadequate climate models and selective sci-

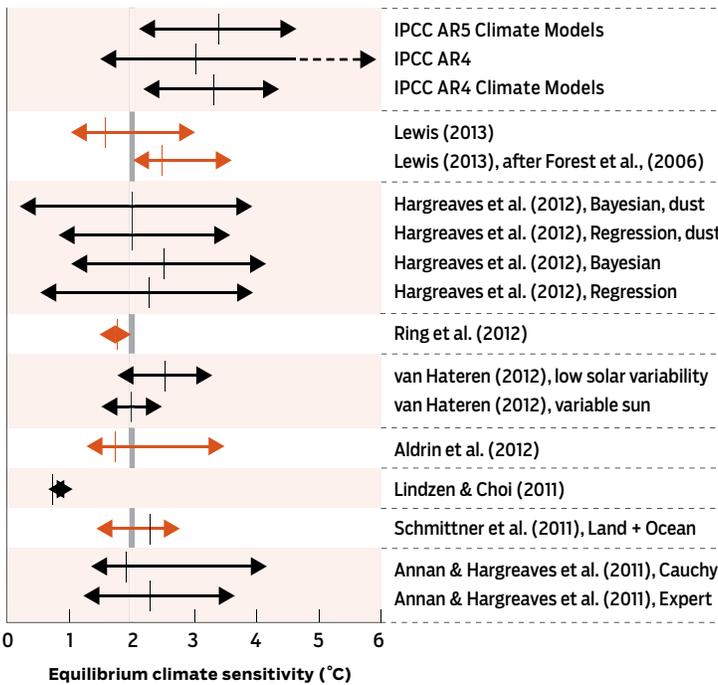
The most serious problem with the 2013 draft is the failure to address the growing number of recently published findings that suggest that the climate sensitivity is much lower than previously estimated.

entific citations. The 2000 assessment used models that were the most extreme ones available and had the remarkable quality of generating “anti-information.” The 2009 assessment used a very low significance criterion for the essential precipitation variable. The draft version of the newest assessment employs operationally meaningless precipitation forecasts. In all three assessments, the climate models likely have too high a climate sensitivity, resulting in overestimates of future change.

In the assessment process, scientists whose professional advancement is tied to the issue of climate change are the authors. Critics participate only during the public comment period, and there is no guarantee that the criticism will alter the contents. For example, the calculation that the residual variance in the first assessment forecasts was greater than the variance of the observed climate data should have been fatal—but, despite that calculation (which was seconded by the most senior climate scientist on the panel that produced it), the report subsequently appeared as if nothing had happened.

The second and third assessments suffer from discrepancies between model predictions and actual temperatures. This discrepancy may be an indication that the “sensitivity” of temperature to carbon dioxide (which can only be specified, as it cannot be calculated from first scientific principals) has been estimated to be too high. The absence of these lower estimates in the most recent 2013 draft assessment is a serious omission and, unless it is corrected, will render the report obsolete on the day it is published. **R**

FIGURE 3
A Less Sensitive Planet?
Climate sensitivity estimates from new research published since 2010, compared with the range given in the IPCC Fourth Assessment Report (AR4) and Fifth Assessment Report (AR5).



NOTE: The arrows indicate the 5% and 95% confidence bounds for each estimate along with the best estimate (median of each probability density function, or the mean of multiple estimates vertical line). Ring et al. (2012) present four estimates of the climate sensitivity and the box encompasses those estimates. The right-hand side of the IPCC AR4 range is dotted to indicate that the IPCC does not actually state the value for the upper 95% confidence bound of its estimate and the left-hand arrow only extends to the 10% lower bound as the 5% lower bound is not given. The vertical bar is the mean of the 14 best estimates from the new findings. The mean climate sensitivity (3.4°C) of the climate models used in the latest IPCC report (2013) is 13 percent greater than the IPCC’s “best estimate” of 3.0°C, and 70% greater than the mean of recent estimates (2.0°C).

READINGS

- *Addendum: Global Climate Change Impacts in the United States*, produced by the Center for the Study of Science. Cato Institute, 2012.
- *The Missing Science from the Draft National Assessment on Climate Change*, by Patrick J. Michaels et al. Cato Institute, 2013.