

## COMMENTS ON THE FOURTH NATIONAL CLIMATE ASSESSMENT

Patrick J. Michaels  
Director, Center for the Study of Science  
Cato Institute  
Washington DC 20001

*Note: The full review has been sent to [review@usgcrp.gov](mailto:review@usgcrp.gov) under filename Michaels\_complete\_review, which will be displayed here in its entirety.*

### 1. Introduction and Plain Language Summary

The draft fourth “National Assessment” (“NA4”) of climate change impacts is systematically flawed and requires a complete revision.

NA4 uses a flawed ensemble of models that dramatically overforecast warming of the lower troposphere, with even larger errors in the upper tropical troposphere. The model ensemble also could not accommodate the “pause” or “slowdown” in warming between the two large El Niños of 1997-8 and 2015-6. The distribution of warming rates within the CMIP5 ensemble is not a true indication of a statistical range of prospective warming, as it is a collection of systematic errors. Despite a glib statement about this Assessment fulfilling the terms of the federal Data Quality Act, that is fatuous. The use of systematically failing models does not fulfill the “maximizing the quality, objectivity, utility, and integrity of information” provision of the Act.

Institutional memory relating to the production of previous assessments is strong, and the process itself is long, as the first drafts of this version were written in the middle of the second Obama Administration. They were written largely by the same team that wrote the 2014 Assessment, which NOAA advertised, at its release, was “a key deliverable of President Obama’s Climate Action Plan.” The first (2000) Assessment used the two most extreme models of the 14 considered for temperature and precipitation. In my review I applied them to 10-year running means of lower-48 temperatures and the residual error was larger than the error of the raw data itself! The historical lineage of the fourth Assessment has all but guaranteed an alarming report, regardless of reality.

USGCRP should produce a reset Assessment, relying on a model or models that work in four dimensions for future guidance and ignoring the ones that don’t.

Why wasn’t this done to begin with? The model INM-CM4 is spot on, both at the surface and in the vertical, but using it would have largely meant the end of warming as a significant issue. Under a realistic emission scenario (which USGCRP also did not use), INM-CM4 strongly supports the “lukewarm” (endnote 1) synthesis of global warming. Given the culture of alarmism that has infected the global change community since before the first (2000) Assessment, using this model would have been a complete turnaround with serious implications.

The new Assessment should employ best scientific practice, and one that weather forecasters use every day. In the climate sphere, billions of dollars are at stake, and reliable forecasts are also critical.

[ PAGE \\* MERGEFORMAT ]

When making a forecast, it's a good idea to look out the window. Meteorologists decide what mix or what individual model is providing the most reliable guidance. Rarely do forecasters average up every available one, because some are better than others, depending upon the situation.

All of the fourth Assessment models other than INM-CM4 forecast the entire tropical troposphere too warm, especially in the upper reaches, and also have the surface too warm (endnote 2). The "pause," which is obvious in both the satellite and HadCRU4 data (endnote 3), wasn't accommodated, as noted by Fyfe et al. (2016). Because INM-CM4 doesn't run hot, it is able to further accommodate the lack of strong warming in the early part of the 21<sup>st</sup> century.

If one assumes, as the International Energy Agency (2017) does, that natural gas is going to continue to replace large amounts of coal energy, 21<sup>st</sup> century warming predicted by INM-CM4 is approximately 1.5°C, a value so low that the social costs of carbon become the social benefits of lukewarming.

In summary, the USGCRP must hit the reset button now. It should use a methodology that works—i.e. a model that works—rather than a family of failures that tout a future of unwarranted gloom and doom. It would also be wise to rely more heavily on a concentration pathway that recognizes the massive worldwide switch from coal to natural gas for both electrical generation and manufacturing. That's the right way, and the only way to produce a credible Assessment.

I would normally also supply an extensive commentary on the Key Findings, but because an entire new Assessment is warranted, the current ones are likely to change dramatically when the new drafts are released.

Administratively, resetting the Assessment will prove difficult. The leadership is long-standing and descended from the community that produced the previous Assessments. A more diverse team is needed to produce what is likely to be a dramatically different document.

(The entire review, including this introduction, containing figures and a table, is in a separate file that has been communicated to the USGCRP, and it will be an integral part of this submission.)

[ PAGE \\* MERGEFORMAT ]

## 2. Detailed Review

### *A Brief Historical Perspective*

This is the fourth National Assessment. It continues the tradition established by the first three.

The First National Assessment (2000) used models that were worse than a table of random numbers when applied to ten year running means for lower 48 temperature. The science team knew this and went ahead anyway (endnote 4). Given that these documents are very influential on national and international policy, that was tantamount, in my opinion, to scientific malpractice (endnote 5). It also chose the two most extreme models, for temperature and precipitation, of the suite that it examined (endnote 6). The second (2009) Assessment was so incomplete that it prompted an entire palimpsest (endnote 7). The third (2014) billed itself as “a key deliverable of President Obama’s Climate Action Plan,” which again received a detailed critical review about its content, illogic, and omissions (endnote 8).

### *Systematic problems with the Fourth Assessment models*

The Fourth National Assessment (hereafter, NA4) is model-based. Quoting from Chapter 2:

The future projections used in this assessment come from global climate models (GCMs) that reproduce key processes in the earth’s climate system using fundamental scientific principles.

It follows that if, as an ensemble, these models are systematically flawed in a significant fashion, it is improper to use them to project the impacts of the climate changes that they predict. That didn’t stop the first (2000) Assessment from using models worse than a table of random numbers, or the second and the third Assessments from using models with flaws similar to the ones in the this version (many are simply “improved” versions of second and third Assessment models). But perhaps this review will get a bit more attention than previous ones, as the political climate of Washington recently underwent an unforecast and abrupt change.

The growing disparity between predicted bulk tropospheric temperatures and observed values, especially at altitude in the tropics (see Figure 1), casts overall doubt on the utility of the large ensemble of general circulation models (GCMs) with regard to 21st century temperatures. The current model suite has an average equilibrium climate sensitivity (ECS) of 3.4°C (Andrews, 2012) (endnote 9) . The disparities may arise as a consequence of the recently acknowledged significant tuning of the GCMs in order for them to simply simulate the evolution of 20th century surface temperatures; see below. Regardless of the cause, these disparities cast doubt on the overall utility of the large ensemble of models with regard to 21<sup>st</sup> century temperatures.

[ PAGE \\* MERGEFORMAT ]

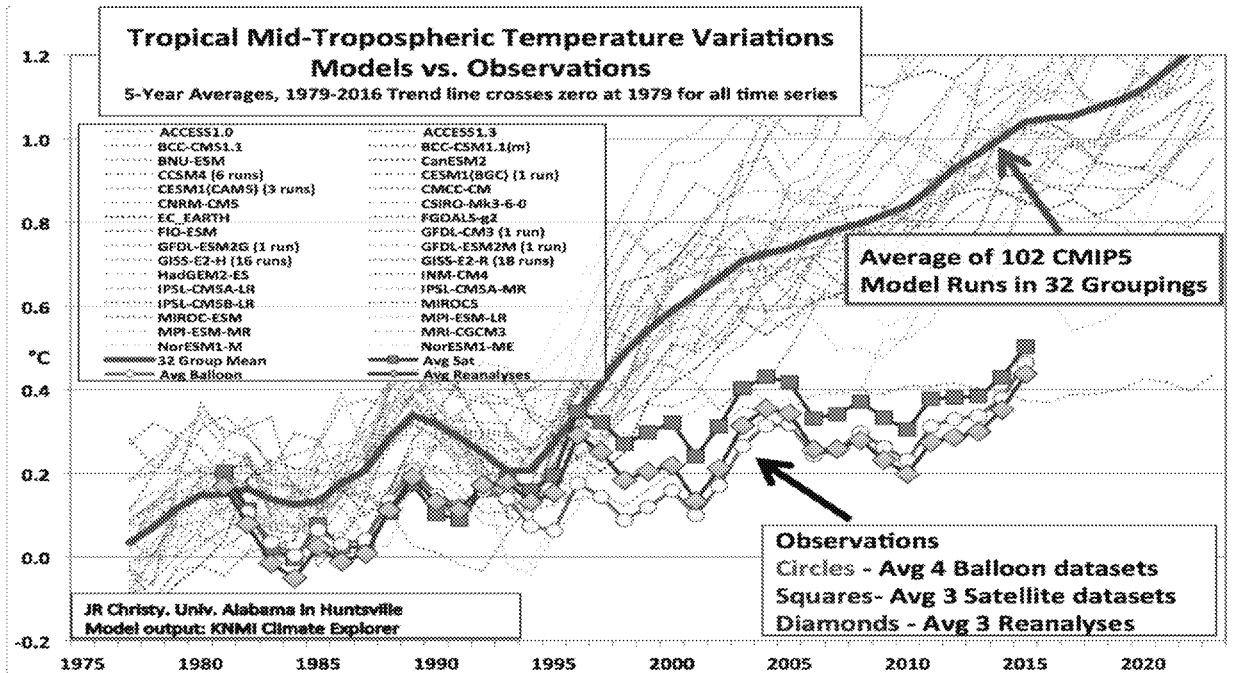


Figure 1. Modelled and observed tropical mid tropospheric (surface-100mb) temperature changes, 1979-2016. From testimony of John Christy to the House Committee on Science, Space and Technology, March 29, 2017. The one model that tracks the observations is INM-CM4. The data are also available in tabular form in the American Meteorological Society's "State of the Climate" report for 2016 (endnote 10).

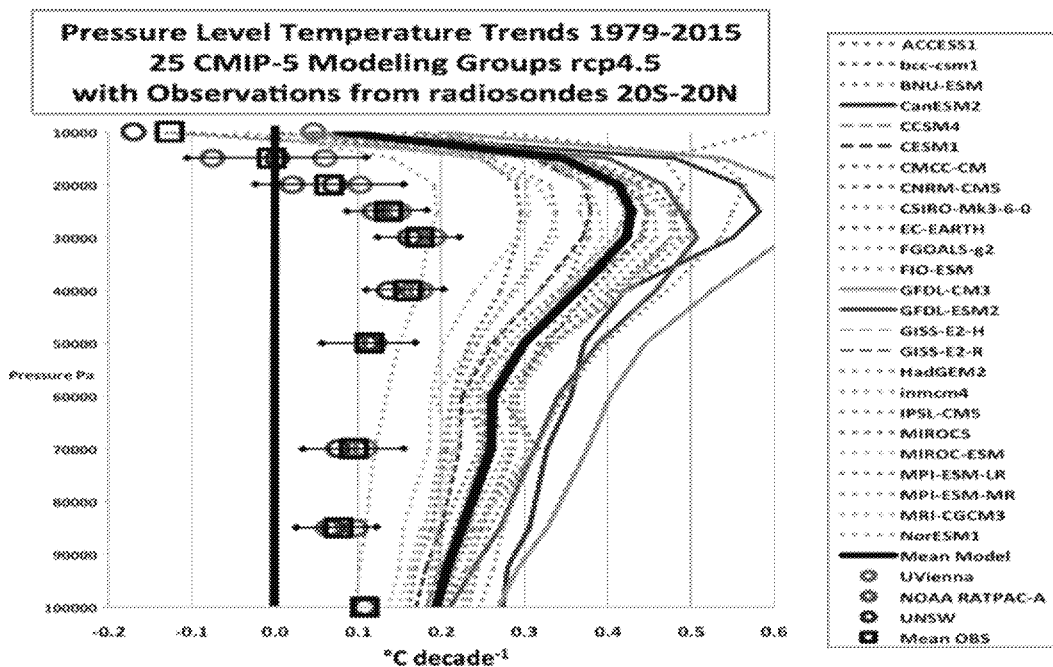


Figure 2. The vertical discrepancy between radiosonde-measured and model predicted tropical temperature trends, 20N-20S, is persistent and very large in the mid and upper troposphere. From Christy and McNider (2017); the exception is again the model INM-CM4.

[ PAGE \\* MERGEFORMAT ]

Similarly, Figure 2 shows the vertical distribution of forecast and observed trends. Commenting on it, Christy and McNider (2017) note:

In every case, with the exception of the Russian model “inmcm4” below 250hPa, individual tropospheric model trends are larger than the observational average below 100 hPa with the discrepancies largest in the upper troposphere...

The point should be clear: unless INM-CM4 is also making systematic errors with major consequences (which are not apparent), the Assessment should be using it rather than the suite of models that is systematically and dramatically wrong.

This type of exercise is undertaken frequently in operational meteorology. Oftentimes the many global and regional forecast models give conflicting results for a given synoptic situation. Forecasters then examine which ones have been performing well, or which perform better given the situation, and then settle upon one or a blend of models to arrive at the final forecast. *They rarely average all of them up.* Emphasizing the ECMWF model in favor of the GFS for 2013 storm Sandy was a prudent choice in the longer timeframes. Averaging them would have been very costly.

Using the range of models that suffer from considerable bias in order to estimate the statistical distribution of a forecast is a folly of additive error, while using unbiased model(s) (in the global sense) minimizes the probability of such an error.

In the 2017 *Climate Science Special Report (CSSR)* for both surface temperatures and specific impacts, and the draft fourth National Assessment, the range of warming is generated almost exclusively by the models that don't work, and not the model that *works*. This is the central reason why the entire fourth Assessment process must be reset.

To reiterate: A collection of errors biased in one direction is hardly a true estimate of the range of a forecast. It is the opposite, a false estimate from models that are clearly warming the topical troposphere at over twice the observed rate. The warming rate forecast in the zone around 200mb is a stunning six *times* what has been observed in the last 36 years. About 38% of the earth's surface is underneath the zone studied.

#### *The Implications of Shale Gas were not Properly Considered*

To compound prospective future errors, the over-reliance on RCP 8.5 in the current Assessment is also questionable. To its credit, the NA4 does repeatedly mention the major displacement of coal with natural gas for electrical generation in the U.S., but fails to note the implication of large-scale international adoption of this switch, and the substitution of gas for coal in worldwide industry. The implication is that RCP 8.5 (mentioned in seven separate textual references (not counting the bibliographies)) is increasingly unlikely (endnote 11).

Quoting from the International Energy Agency (IEA)

The global natural gas market is undergoing a major transformation driven by new supplies coming from the United States to meet growing demand in developing countries and industry surpasses the power sector as the largest source of gas demand growth...

[ PAGE \\* MERGEFORMAT ]

The evolution of the role of natural gas in the global energy mix has far-reaching consequences on energy trade, air quality and carbon emissions...

Global gas demand is expected to grow by 1.6% a year...China will account for 40% of this growth.

NA4 should therefore rely more on RCP 6.0 rather than 8.5.

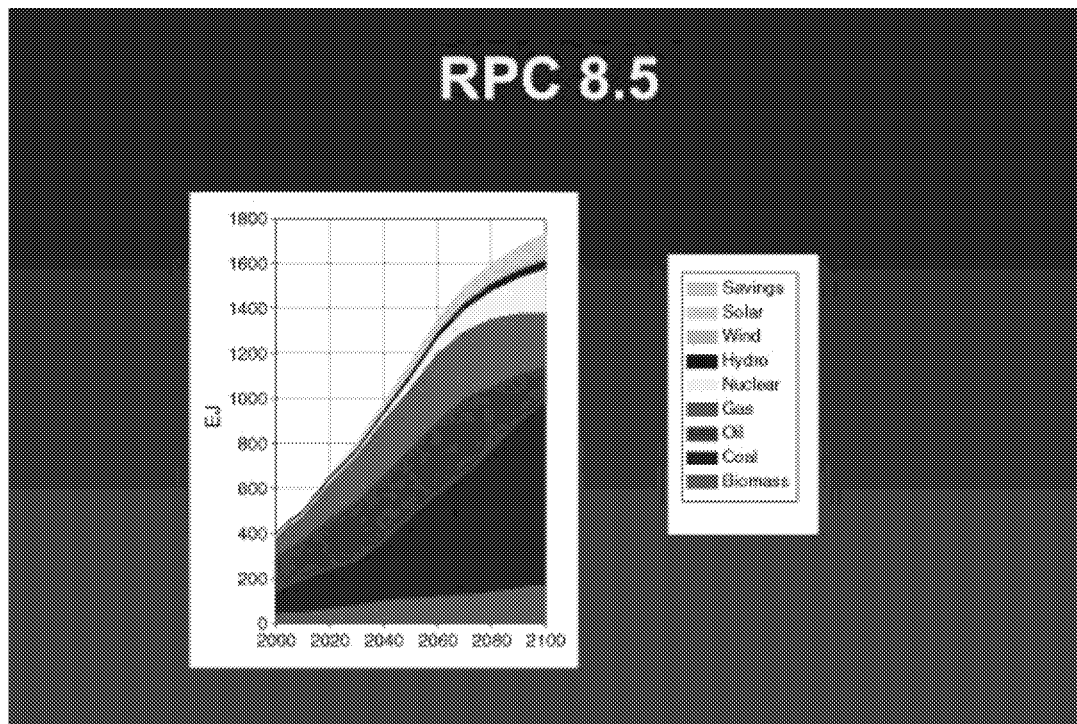


Figure 3. There is no evidence for rapidly increasing displacement of coal with natural gas for electrical generation in RCP 8.5, even though this is now forecast by the IEA worldwide.

The argument this is simply a U.S. phenomenon is premature. Unless the Chinese, who are the world's largest emitters, are different than people elsewhere, there will ultimately be restive demands to clean their unhealthy, coal-polluted air as their per capita income rises. The abundance of available gas at that time will almost certainly result in major fuel switching.

The reset NA4 needs to account for this, with an increased emphasis on RCP 6.0.

#### *The Social Cost (or Benefit) of Lukewarming*

INM-CM4 is decidedly lukewarm. I used *KNMI Explorer* to estimate 21<sup>st</sup> century warming—however, unlike for many of the other models, KNMI only has RCP 4.5 and 8.5 for INM-CM4. Using a warming slightly below the midpoint for those two gives a 21<sup>st</sup> century surface warming of approximately 1.5°. This is quite consistent with the empirical transient sensitivity recently calculated by Christy and McNider (2017).

[ PAGE \\* MERGEFORMAT ]

We therefore used their probability density function in a subsequent calculation by Kevin Dayaratna of the Heritage Foundation using the FUND model to determine an approximate social cost of carbon. We elected to follow the OMB (2004) guidelines that recommended using the robust historical average 7.0% discount rate, as well as the 3.0 it also recommends and the 5.0 used by the Obama Administration.

We show results of with equilibrium climate sensitivity/transient climate sensitivity ratios of 1.3 and 1.7.

### **Social Costs (Benefits) of a Ton of Carbon Dioxide and Probability of Benefit**

#### *1.3 Ratio ECS/TCS*

YEAR	3% D.R	5%	7%
2020	(0.55) (.55)	(1.36) (.64)	(1.31) (.72)
2050	1.19 (.46)	(0.39) (.52)	(0.77) (.57)

#### *1.7 Ratio ECS/TCS*

2020	4.04 (.23)	0.21 (.36)	(0.86) (.72)
2050	5.99 (.19)	1.25 (.31)	(0.23) (.57)

These results are very similar to what Dayaratna et al. (2017) published last year using the probability density functions for warming of Lewis and Curry (2015). This is expected because it is quite similar to what is derived from Christy and McNider (2017). I fully expect if we used a distribution from INM-CM4 run with RCP 6.0 that there would be similar results.

These are, of course, radically different from the cost estimates emanating from the previous Administration, but it is noteworthy that it specifically omitted the OMB-recommended robust historical discount rate of 7%.

We note that seven of the 12 estimates shown above are net benefits rather than costs (shown in parentheses). A reset Assessment using ICM-CM4 or a satellite/radiosonde derived probability function for 21<sup>st</sup> century warming is going to be radically different than estimates using the larger, warm-biased suite of climate models.

#### *We May Never Know the Cause of the Overestimated Bulk Warming*

It may be nearly impossible to determine the cause(s) of overforecast bulk warming, but its effects are manifold. By forecasting a much warmer upper troposphere than is being observed, the models must be systematically underestimating tropical precipitation (endnote 12). It would also seem that descending air into the subtropical high pressure systems would be warmer than what is being observed. These two simple examples would have consequences for vegetation; a drier tropical regime would affect the vast tropical rainforests, and warmer descending air is likely to increase desertification in the persistent Hadley cells. Both of these processes will then create their own secondary feedbacks to surface temperature and sensible weather.

[ PAGE \\* MERGEFORMAT ]

If these problems can't be corrected, the reset NA4 may as well exit the business of predicting climate impacts, especially on vegetation, agriculture, and sea level rise. Those impacts are all primarily driven by a rise in temperature, and if too much bulk warming is being demonstrably predicted, NA4 becomes not unlike NA1 (2000), when the science team went ahead anyway after being told (and finding out themselves) that the models were actually supplying negative knowledge, inducing larger residual errors *after* applying them to the raw data. "Damn the data, full speed ahead" can no longer be tolerated.

The problem is that we may never know what has gone wrong with the models as an ensemble. In a paper detailing the process of model tuning, Mauritsen (2012) noted it is apparently impossible to completely know what was done to these models over their historical development. In Mauritsen's words, "model development happens over generations, and it is difficult to describe comprehensively (endnote 13)."

Significant portions of climate models are therefore black boxes with varying degrees of subjectivity. Recently, Hourdin et al., (2017) issued a rather strident call for more transparency about model tuning.

Left to their own devices, it has long been known (endnote 14) that climate models run with increasing atmospheric carbon dioxide only produce too much warming. As a result, internal parameters that ultimately predict future climate are tuned to reproduce the global temperature history of the 20<sup>th</sup> century (endnote 15). Model parameters are tuned to what Hourdin et al. called an "anticipated acceptable range."<sup>1</sup>

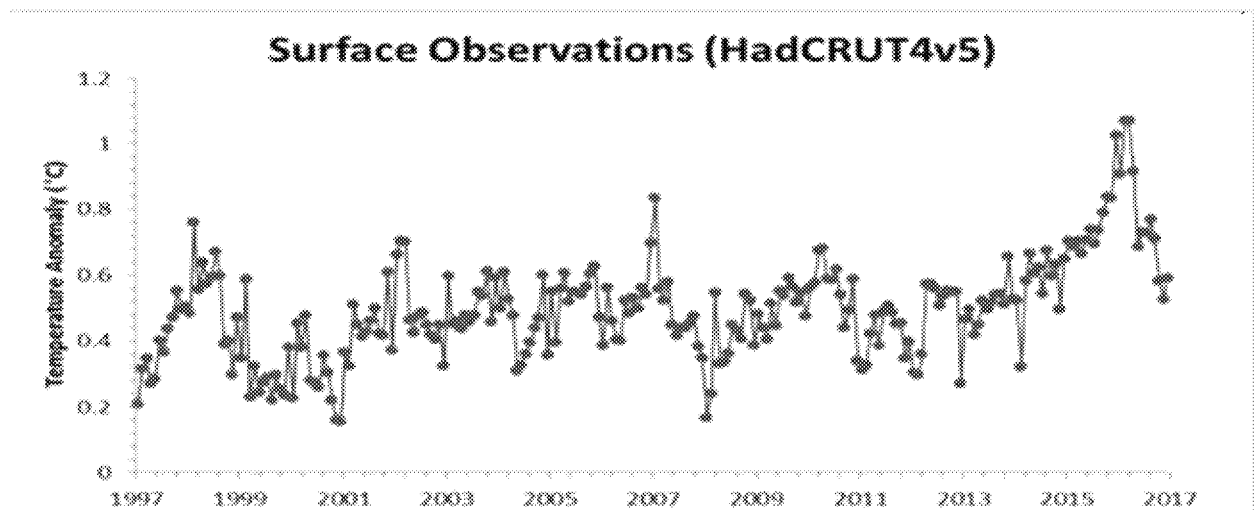
NA4 and the accompanying Climate Science Special Report repeatedly state that models show anthropogenic emissions are responsible for almost all 20<sup>th</sup> century warming.

This is claimed despite the fact that of the two twentieth-century warmings; the first one, approximately from 1910 to 1945, could hardly have been a result of carbon dioxide emissions. The 1910-1945 warming is statistically similar in slope to the 1976-1997 warming.<sup>2</sup>

---

<sup>1</sup> Hourdin, F., et al., 2017. The art and science of climate model tuning. *Bulletin of the American Meteorological Society*. <https://doi.org/10.1175/BAMS-D-15-0013335.1>

<sup>2</sup>



Post-1998 remains controversial. There is a clear "pause" from late 1997 through 2014 (or 2002-2014,

[ PAGE \\* MERGEFORMAT ]

Ice core data from Law Dome show the surface concentration was only around 298ppm when the first warming began, which gives a CO<sub>2</sub> forcing of +0.35 w/m<sup>2</sup> based upon the standard formula (dRF= 5.35ln(298/279)). Stevens (2015),<sup>3</sup> citing Carslaw et al. (2013) gives a sulfate forcing of -0.3 watts/m<sup>2</sup>, resulting in a near-zero net combined forcing. Tuning the models to somehow account for this warming implies an enormous sensitivity. If that were actually true, current temperatures would be so high that there would be little policy debate.

Tuning the models to mimic the historical record and then claiming that anthropogenic emissions explain the early warming is circular reasoning at its finest; *reset NA4 needs to be explicit about this.*

Consequently, we are left with the following unhappy circumstance: it is the modeler, and not the model that decides what the “anticipated acceptable range” of parameters is in order to fit the double peak of warming in the 20<sup>th</sup> century. Claiming that this is evidence for the reliability of the models’ future prediction is fatuous.

In fact, the opposite is true. Each time a model is tuned in search of a particular result, an increment of potential future instability is added. It’s not surprising that, in forecast mode, the models make such egregious errors over the entire tropical troposphere.

#### *Data Quality Act*

Any Assessment must comply with the Data Quality Act, including a reset NA4. It is doubtful that relying on systematically failing models with parameters tuned to an “anticipated acceptable range” fulfills the Act’s requirement to “maximize the quality, objectivity, utility, and integrity of information.”

### **3. Conclusion**

This review has demonstrated that NA4 suffers from a fundamental methodological flaw in assuming that models making large bulk errors are representative of a range of future warming. Ubiquitous tuning of the models to the 20<sup>th</sup> century history hardly increases their reliability. NA4 also pays inadequate attention to the implications of an ongoing seismic shift in world energy towards natural gas. Warming predicted by the one model that does not suffer the bulk errors, coupled with a slightly lower concentration pathway because of forecast switching from coal to natural gas, becomes a net benefit rather than a social cost.

Going back to 2000, there have been persistent problems throughout the entire assessment process, underscoring the need for major administrative change.

For these and other reasons, draft NA4 should be shelved and reset, so that time and resources can be devoted to a new Assessment that corrects and addresses the first three Assessments and the draft NA4.

---

after the first ENSO cycle in this plot is complete) evident both the latest version of HadCRUT4 and the UAH lower tropospheric satellite-sensed data.

<sup>3</sup> Stevens, B., 2015. Rethinking the lower bound of aerosol radiative forcing. *J. Clim.* **28**, 4794-4819.

[ PAGE \\* MERGEFORMAT ]

## ENDNOTES

1. “Lukewarmers” agree on a human influence on global climate, but at or below the bottom of the range specified by the U. N.’s Intergovernmental Panel on Climate Change.
2. See the review of this Assessment submitted by Richard McNider and John Christy.
3. See Footnote 22 near the end of this document.
4. I wrote to the chief scientist, Tom Karl, and he emailed me back, that “we ran the test you did but changed the averaging period” from 10-year running means to 1, 5, 10, 20 and 25 years. He kindly included a graph that showed *at all time intervals tested* that the residual variance after applying the models was larger than the raw variance. A modified version (for clarity) can be found as Figure 25, page 109 of my 2016 book *Lukewarming*. I first documented it in 2003 in a chapter, “Science or Political Science? An Assessment of the U.S. National Assessment of the Potential Consequences of Climate Variability and Change.” In Gough, M., Ed., *Politicizing Science: The Alchemy of Policymaking*. Hoover, Palo Alto.
5. This action was exactly analogous to a physician prescribing a medication he or she knows will make the patient worse.
6. Documented on page 209 of my 2004 book *Meltdown*.
7. <https://object.cato.org/pubs/Global-Climate-Change-Impacts.pdf>
8. <https://www.cato.org/publications/the-missing-science-from-the-draft-national-assessment>
9. The associated Fifth Assessment Report of the IPCC gives this figure as 3.2°C, but the calculated average is 3.37.
10. <https://www.ametsoc.org/ams/index.cfm/publications/bulletin-of-the-american-meteorological-society-bams/state-of-the-climate/>
11. This will be a result of the increasing real per-capita incomes and GDP. When certain levels of affluence are reached, environmental protection becomes affordable and is publicly demanded. This happened in the mid-20<sup>th</sup> century in the US, beginning with the miasmatic air of Pittsburgh. The horrific air quality in urban China will likely be the first target of any nascent green movement there. The amount of retrofitting of their newer coal plants remains unknown, but as IEA indicates (above), China will be responsible for the largest percent of gas usage growth in the next five years.
12. Unless, of course, this output is systematically “tuned”. But, as Mauritsen implied, and Hourdin lamented, we may never know what was done.
13. Grad students and postdocs marching through the models didn’t always keep good notes on what they did.
14. Tom Wigley wrote of this in an in-house journal *Climate Monitor* in 1987, and it was explicitly acknowledged in the second (1996) Assessment Report of the IPCC.
15. From Voosen, 2016: “Indeed, whether climate scientists like to admit it or not, nearly every model has been calibrated precisely to 20<sup>th</sup> century climate records—it would have ended up in the trash. “It is

[ PAGE \\* MERGEFORMAT ]

fair to say that all models have tuned it”, says Isaac Held, a scientist at the Geophysical Fluid Dynamics Laboratory, another prominent modelling center in Princeton, New Jersey.” (Voosen, P., 2016. Climate scientists open up their black boxes to scrutiny. *Science* **354**, 401-402.)

#### REFERENCES CITED

Andrews, T., et al., 2012. Forcing, feedbacks, and climate sensitivity in CMIP5 coupled ocean-atmosphere models. *Geophys. Res. Lett.* **39**, DOI: 10.1029/2012GL051607.

Christy, J. R., and R. T. McNider, 2017. Satellite bulk tropospheric temperatures as a metric for climate sensitivity. *Asia-Pac. J. Atmos. Sci.* **53**, 511-518.

Dayaratna, K., McKittrick, R., and D. Kreutzer, 2017. Empirically constrained climate sensitivity and the social cost of carbon. *Climate Change Economics*, April 2017. DOI: <http://dx.doi.org/10.1142/S2010007817500063>

Fyfe, J.C., et al., 2016. Making sense of the early 2000s warming slowdown. *Nature Climate Change* **6**, 224-228.

Hourdin, F., et al., 2017. The art and science of climate model tuning. *Bulletin of the American Meteorological Society*. <https://doi.org/10.1175/BAMS-D-15-0013335.1>

IEA, 2017. IEA Sees Global Gas Demand Rising to 2022 as U S Drives Market Transformation. <https://www.iea.org/newsroom/news/2017/july/iea-sees-global-gas-demand-rising-to-2022-as-us-drives-market-transformation.html>

Lewis, N., and J.A. Curry, 2015. The implications for climate sensitivity of AR5 forcing and heat uptake estimates. *Cli. Dyn.*, **46** 1387-1396.

Mauritsen, T., et al, 2012. Tuning the Climate of a Global Model. *J. Adv. Modelling Earth Systems* **4**, DOI: 10.1029/2012MS000151