Fact Sheet for "Consistency of Modelled and Observed Temperature Trends in the Tropical Troposphere", by B.D. Santer *et al.*[A]

Abstract

Using state-of-the-art observational datasets and results from a large archive of computer model simulations, a consortium of scientists from 12 different institutions has resolved a long-standing conundrum in climate science – the apparent discrepancy between simulated and observed temperature trends in the tropics. Research published by this group indicates that there is no fundamental discrepancy between modeled and observed tropical temperature trends when one accounts for: 1) the (currently large) uncertainties in observations; 2) the statistical uncertainties in estimating trends from observations. These results refute a recent claim that model and observed tropical temperature trends "disagree to a statistically significant extent". This claim was based on the application of a flawed statistical test and the use of older observational datasets.

Ben Santer, Peter Thorne, Leo Haimberger, Karl Taylor, Tom Wigley, John Lanzante, Susan Solomon, Melissa Free, Peter Gleckler, Phil Jones, Tom Karl, Steve Klein, Carl Mears, Doug Nychka, Gavin Schmidt, Steve Sherwood, and Frank Wentz

October 10, 2008

[A]This paper will be published online in the *International Journal of Climatology* during the week of Oct. 6-10, 2008.

QUESTION 1: What is the scientific context for the research published in the Santer *et al. International Journal of Climatology* paper?

Our paper compares modeled and observed atmospheric temperature changes in the tropical troposphere.[B] We were interested in this region because of an apparent inconsistency between computer model results and observations. Since the late 1960s, scientists have performed experiments in which computer models of the climate system are run with human-caused increases in atmospheric concentrations of greenhouse gases (GHGs).[C] These experiments consistently showed that increases in atmospheric concentrations of GHGs should lead to pronounced warming, both at the Earth's surface and in the troposphere. The models also predicted that in the tropics, the warming of the troposphere should be larger than the warming of the surface.[D]

Observed estimates of surface temperature changes are in good agreement with computer model results, confirming the predicted surface warming.[E] Until several years ago, however, most available estimates of tropospheric temperature changes obtained from satellites and weather balloons (radiosondes) implied that the tropical troposphere had actually cooled slightly over the last 20 to 30 years (in sharp contrast to the computer model predictions, which show tropospheric warming).

For nearly a decade, this apparent disconnect between models and reality has been used by some scientists and politicians to argue that:

- The surface thermometer record is wrong;

[B]The troposphere is the lowest layer of the atmosphere, where most weather phenomena take place. In the tropics, the troposphere extends from the surface to a height of about 10 miles (16 km) above the Earth's surface.

[C]Both climate models and the experiments performed with them have become more realistic over time. Since the mid 1990s, many climate model experiments have incorporated not only human-caused changes in GHGs, but also changes in other "forcing agents" that have effects on global or regional climate. Examples include human-caused changes in various aerosol particles (such as sulfate and soot aerosols), and natural changes in the Sun's energy output and the amount of volcanic dust in the atmosphere.

[D]This prediction of larger warming aloft than at the surface holds for all factors that tend to warm the surface of the Earth – it is not unique to human-caused changes in GHGs.

[E]This agreement between models and observations was also found for complex geographical patterns of surface temperature changes – not simply for trends in temperature changes averaged over very large areas (such as the tropics).

- The Earth has not experienced <u>any</u> surface or tropospheric warming since the beginning of satellite measurements of atmospheric temperature in 1979;

- Human-caused changes in greenhouse gases have no effect on climate;

- Computer models have no skill in simulating the observed temperature changes in the tropics, and therefore cannot be used to predict the climatic "shape of things to come" in response to further increases in greenhouse gases.

Our paper attempts to determine whether there is indeed a real and statistically significant discrepancy between modeled and observed temperature changes in the tropics, as was claimed in a paper published online in December 2007 in the *International Journal of Climatology*. As discussed in QUESTION 9, we find that this claim is incorrect.

<u>QUESTION 2: What arguments were made to support this claim?</u>

David Douglass, John Christy, Benjamin Pearson, and S. Fred Singer[1] devised a statistical test to determine whether modeled and observed atmospheric temperature trends in the tropical troposphere were significantly different. They applied this test in several different ways. First, they considered temperature trends in two different layers of the troposphere (the lower troposphere and the mid- to upper troposphere). In each of these layers, their test suggested that the modeled warming trends were larger than and significantly different from the warming trends estimated from satellite data. Second, they compared trends in the temperature <u>differences</u> between the surface and the lower troposphere – a measure of the "differential warming" of the surface and lower atmosphere. Once again, their test pointed towards the existence of statistically significant differences in modeled and observed trends.

The bottom-line conclusion of Douglass *et al*. was that "*models and observations disagree to a statistically significant extent*". As discussed in QUESTIONS 6-8, we show that this statistical test is flawed, and that the conclusions reached by Douglass *et al*. are incorrect.

<u>QUESTION 3: But hadn't the scientific community already resolved this issue?</u>

The community had already achieved a partial resolution of this issue in a 2006 Report issued by the U.S. Climate Change Science Program (CCSP)[2]. The CCSP Report concluded that, when one examined temperature changes at the global scale, newer satellite and weather balloon datasets showed "*no significant discrepancy*" between surface and tropospheric warming trends, and were therefore consistent with computer model results. But the same CCSP Report noted that it was not possible (in 2006) to reconcile modeled and observed temperature changes in the tropics, where "*most observational datasets show more warming at the surface than in the troposphere, while most model runs have larger warming aloft than at the surface*".

The CCSP Report relied almost exclusively on published literature. At the time of its publication in 2006, there were no peer-reviewed studies on the formal statistical significance of differences between modeled and observed tropical temperature trends. The Douglass *et al*. paper attempted to assess the statistical significance of the model-versus-observed tropical trend differences noted in the CCSP Report.

QUESTION 4: What was the thrust of your new research?

Our primary goal was to determine whether the findings of Douglass *et al*. were sound. As noted above, Douglass *et al*. reported that "*models and observations disagree to a statistically significant extent*". They interpreted their results as evidence that computer models are seriously flawed, and that the projections of future climate change made with such models are untrustworthy. If Douglass *et al*. were right, this would imply that there was some fundamental flaw – not only in all state-of-the-art climate models, but also in our basic theoretical understanding of how the climate system should respond to increases in GHGs. We wanted to know whether such a fundamental flaw really existed.

QUESTION 5: What specific issues did you focus on?

We focused on two issues. First, Douglass *et al*. claimed that they had applied a "*robust statistical test*" to identify statistically significant differences between modeled and observed temperature trends. We sought to understand whether their test was indeed "*robust*" and appropriate. Second, Douglass *et al*. claimed to be using the "*best available updated observations*" for their study. We did not believe that this claim was accurate.

We decided to check their analysis by applying a variety of different statistical tests to modeled and observed temperature trends, and by employing temperature data from more recent observational datasets – datasets that were either unavailable to Douglass *et al*. at the time of their study, or which were available, but had not been used by them.

QUESTION 6: What did you learn about the appropriateness of the Douglass *et al*. test?

We found that there was a serious flaw in the "*robust statistical test*" that Douglass *et al*. had used to compare models and observations. Their test ignored the effects of natural climate "noise" on observed temperature trends, and the resulting statistical uncertainty in estimating the "signal component" of these trends (see QUESTION 7 for a definition of the "signal component").

QUESTION 7: Why was this a problem?

We know that in the real world, changes in temperatures are due to a combination of human effects and natural factors. The "natural factors" can be things like volcanic eruptions or changes in the Sun's energy output. Another type of "natural factor" is referred to as "internal variability", which is unrelated to changes in the Sun or volcanic dust, and involves phenomena like El Niños, La Niñas, and other natural climate oscillations. In the tropics in particular, El Niños and La Niñas have a substantial effect on surface and atmospheric temperature. They introduce climate "noise", which complicates the separation of human and natural effects on temperature.

Douglass *et al*. effectively assumed that the observed surface and tropospheric temperature trends were perfectly-known, and that these trends were purely due to human-caused changes in greenhouse gases[F]. The inappropriateness of this assumption is immediately obvious by looking at any observed temperature time series, such as the surface and tropospheric temperature time series shown below.

[F]In their paper, Douglass *et al*. claim to be testing "*the proposition that greenhouse model simulations and observations can be reconciled*". The model simulations of 20th century climate change that they used to test this proposition, however, include a variety of different human and natural forcing factors, such as changes in sulfate and soot aerosols, volcanic dust, the Sun's energy output, and land surface properties. These so-called "20CEN" experiments are not just driven by human-caused increases in GHGs. Douglass *et al*.'s proposition that they are only testing the response of climate models to GHG increases is simply incorrect.

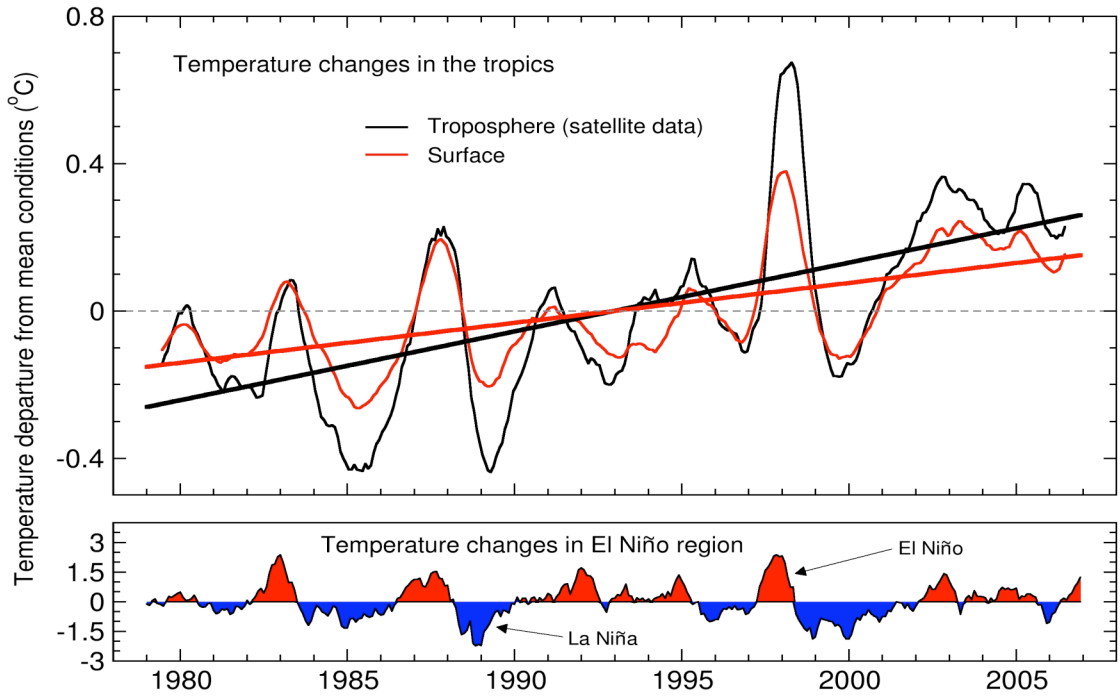

Figure Caption: Estimates of observed temperature changes in the tropics (30°N-30°S). Changes are expressed as departures from average conditions over 1979 to 2006. The top panel shows results for the surface[G] and lower troposphere.[H] The thin red and black lines in the top panel are 12-month running averages of the temperature changes for individual months. The thick straight lines are trends that have been fitted to the time series of surface and tropospheric temperature changes. The warming trend is larger in the tropospheric temperature data than in the surface temperature record, in accord with computer model results. The bottom panel shows a commonly-used index of El Niño and La Niña activity, consisting of sea-surface temperature changes averaged over the so-called Niño 3.4 region of the tropical Pacific. The bottom panel shows that much of the year-to-year variability in surface and lower tropospheric temperatures is related to changes in El Niños and La Niñas.

This Figure illustrates that both tropical surface and tropospheric temperatures have gradually warmed since 1979. Superimposed on this overall warming is climate "noise", which in this case arises primarily from El Niños and La Niñas. When temperatures are averaged over the tropics (and indeed, over the globe), El Niños tend to warm the surface and lower atmosphere, and La Niñas tend to

[G]Surface data are from version 3 of the Extended Reconstructed Sea Surface Temperature Dataset (ERSST) produced by the National Oceanic and Atmospheric Administration (NOAA).
[H]Lower tropospheric temperatures are from version 3.0 of the TLT retrieval produced by Remote Sensing Systems in Santa Rosa, California.

cool these regions.[l] As is visually obvious, El Niños and La Niñas introduce considerable year-to-year variability in surface and tropospheric temperature.

Because of the climate noise introduced by El Niños and La Niñas, there is uncertainty in estimating any underlying temperature trend, such as that arising from slow, human-caused increases in GHGs. In the real world and in many model simulations of 20th century climate change, this underlying trend in temperature is not caused by GHG increases alone – it results from the combined changes in GHGs and other external forcing factors, and is partly masked by climate noise.

The underlying "signal trend" is what we really want to compare in climate models and observations. Any meaningful statistical test of the differences between modeled and observed temperature trends must therefore account for the statistical uncertainty in estimating this "signal trend" from noisy observational data. The Douglass *et al*. test did not account for this uncertainty.

QUESTION 8: What were the consequences of the flaw in the Douglass *et al*. test?

The primary consequence was that Douglass *et al*. reached incorrect conclusions about the true statistical significance of differences between modeled and observed temperature trends in the tropics. When we applied modified versions of their test – versions that properly accounted for uncertainties in estimating the "signal component" of observed temperature trends – we obtained results that were strikingly different from theirs. Like Douglass *et al*., we applied our tests to modeled and observed temperature trends:

- In individual layers of the troposphere;
- In the trend difference between surface and tropospheric warming rates.

Unlike Douglass *et al*., however, we found that most of our tests involving temperature trends in individual layers of the troposphere did not show statistically significant differences between models and observations. This result was relatively insensitive to which model or satellite dataset we chose for the trend comparison.

[l]For example, 1998 was unusually warm because of the effects of a very large El Niño.

The situation was a little more complex for tests involving the trend difference between surface and tropospheric warming rates. In this case, the statistical significance of the differences between models and observations was sensitive to our choice of observational datasets. When we used a satellite-based tropospheric temperature dataset developed at Remote Sensing Systems (RSS) in Santa Rosa, California, we found that the warming in the tropical troposphere was always larger than the warming at the surface.[J] This behavior is consistent with the behavior of the climate models and with our understanding of the physical processes that govern tropospheric temperature profiles. It is contrary to the findings of Douglass *et al*.

However, when we used a satellite-based tropospheric temperature dataset developed at the University of Alabama at Huntsville (UAH)[K], the tropospheric warming was less than the surface warming. But even when we employed UAH data, our statistical test showed that the observed difference between surface and tropospheric warming trends was not always significantly different from the trend difference in model simulations. Whether or not trend differences were statistically significant was dependent on the choice of model and the choice of observed surface dataset used in the test.[L]

QUESTION 9: So what is the bottom line of your study?

The bottom line is that we obtained results strikingly different from those of Douglass *et al*. The "*robust statistical test*" that they used to compare models and observations had at least one serious flaw – its failure to account for any uncertainty in the "signal component" of observed temperature trends (see QUESTION 7). This flaw led them to reach incorrect conclusions. We showed this by applying their test to randomly generated data with the same statistical properties as the observed temperature data, but without any underlying "signal trend". In this "synthetic data" case, we knew that significant differences in temperature trends could occur by chance only, and thus would happen infrequently. When we applied the Douglass *et al*. test, however, we found that even randomly generated data showed statistically significant trend differences much more frequently than we would expect on the basis of chance alone. A test that fails to behave properly when used with random data – when one knows in

[J]Irrespective of which one of four different observational datasets was used to characterize changes in tropical surface temperatures.
[K]Developed by John Christy (one of the co-authors of the Douglass *et al*. paper), Roy Spencer, and colleagues.
[L]See Table V in our paper.

advance what results to expect – cannot be expected to perform reliably when applied to real observational and model data.

Q10: Final question: Have you reconciled modeled and observed temperature trends in the tropics?

We've gone a long way towards such a reconciliation. There are at least two reasons for this.[M] The first reason is that we have now applied appropriate statistical tests for comparing modeled and observed temperature trends in the tropics. Unlike the Douglass *et al*. test, our test properly accounts for uncertainty in estimating the "signal component" of observed temperature trends. Results from these more appropriate tests do not support the claim that there are fundamental, pervasive, and statistically significant differences between modeled and observed tropical temperature trends. This claim is not tenable for temperature trends in individual layers of the troposphere. Nor is it tenable for the differences in the warming rates of the surface and troposphere.

Second, we now have many more estimates of recent temperature changes. These have been produced by a number of different research groups, often using completely independent methods.

Research groups involved in the development of newer sea surface temperature datasets have reported improvements in the treatment of information from buoys and satellites. This has led to slightly reduced estimates of the warming of the tropical ocean surface (relative to the warming in the earlier surface temperature datasets used by Douglass *et al*. and in the CCSP Report). Additionally, newly-developed satellite and radiosonde datasets now show larger warming of the tropical troposphere than was apparent in the datasets used by Douglass *et al*. The enhanced tropospheric warming is due to improvements in our ability to identify and adjust for biases introduced by changes over time in the instruments used to measure temperature.[N]

[M]A third reason is that several studies published within the last 12 months provide independent evidence for substantial warming of the tropical troposphere. These studies have documented pronounced increases in surface specific humidity and atmospheric water vapor that are in accord with tropospheric warming.

[N]Several of the newer radiosonde and satellite datasets that exhibit pronounced tropospheric warming are based on novel approaches to the construction of homogeneous datasets. These approaches often involve bringing in data from new sources (such as hitherto unused satellite data, or data on the physical relationship between temperature and wind) in order to better constrain uncertainties in estimated tropospheric temperature changes.

Access to such a rich variety of independently produced datasets has provided us with a valuable perspective on the inherent uncertainty in observed estimates of recent climate change. Based on our current best estimates of these observational uncertainties, there is no fundamental discrepancy between modeled and observed tropical temperature trends. In fact, many of the recently-developed observational datasets now show tropical temperature changes that are larger aloft than at the surface – behavior that is entirely consistent with climate model results.

One of the lessons from this work is that even with improved datasets, there are still important uncertainties in observational estimates of recent tropospheric temperature trends. These uncertainties may never be fully resolved, and are partly a consequence of historical observing strategies, which were geared towards weather forecasting rather than climate monitoring. We should apply what we learned in this study toward improving existing climate monitoring systems, so that future model evaluation studies are less sensitive to observational ambiguity.

---

[1] Douglass DH, Christy JR, Pearson BD, Singer SF. 2007. A comparison of tropical temperature trends with model predictions. *International Journal of Climatology* **27**: doi:10.1002/joc.1651.

[2] Karl TR, Hassol SJ, Miller CD, Murray WL (*eds*). 2006. *Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences*. A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research. National Oceanic and Atmospheric Administration, National Climatic Data Center, Asheville, NC, 164 pp.