RECONCILING OBSERVATIONS OF GLOBAL TEMPERATURE CHANGE: 2013

by

David H. Douglass & John R. Christy (USA)

Reprinted from

ENERGY & ENVIRONMENT

VOLUME 24 No. 3 & 4 2013
ABSTRACT
In 2000 a panel of the U.S. National Research Council in a report with the same title suggested, among other things, that a “substantial disparity remains” between the observed warming rates of the surface and troposphere. Also, in 2000, the climate models showed more warming of the tropical atmosphere than was observed. Many papers have been written since then. We discuss the most recent papers on this subject and using the latest data show that the differences remain unresolved.

Key words: climate, models, disparity

1. INTRODUCTION.
A panel of the NRC [2000] was convened to address a discrepancy between temperature trends in the upper air (troposphere) and those of the surface. This was partly prompted by climate model simulations that depicted upper air temperatures warming at a rate significantly greater than that of their own surface temperatures. Thus there were two discrepancies that were raising questions, (1) the difference between the absolute rate of warming in the upper air in observations vs. models and (2) the difference in the relationship of temperature trends between the upper air and the surface in observations vs. models. These discrepancies were most prominent in the tropics (20°S – 20°N) where models consistently predicted more rapid warming in the troposphere (i.e. the “hot spot”) than the surface due to enhanced greenhouse gas forcing. Similar issues were also discussed in the Climate Change Science Program Synthesis and Assessment Product 1.1 “Temperature Trends in the Lower Atmosphere” [CCSP, 2006] in which it was reported “… discrepancies in the tropics remain to be resolved.”

Douglass et al. [2007] examined these tropical discrepancies by specifically comparing upper air temperature trends between the “model-average” available at that time from the Climate Model Intercomparison Project #3 or CMIP-3, and observations under the pre-condition, and fortunate coincidence, that the surface temperature trends of the “model-average” and the observations were essentially the same. The goal was
to see how models and observations compared in the upper air given that they both had the same surface trend. Our result indicated: (1) the model-average upper air trends were significantly greater than observations and therefore: (2) the amplification of the trends in the models’ atmospheres was significantly greater than seen in observations.

Santer et al. [2008], while claiming to dispute Douglass et al., actually performed a different analysis. They compared upper air trends between models and observations using models with surface trends differing from observations. Thus, rather than addressing Douglass et al.’s question which was, “If models and observations have the same surface trends, do their upper air trends agree?” Santer et al. asked, “Are there any upper air trends from models that might agree with observations, no matter what their surface trends might be?” This created the proverbial “apples to oranges” comparison (i.e. comparing upper air trends between models and observations when the two groups had different surface trends) and therefore did not address the specific finding of Douglass et al. [2007].

Other inconsistencies of Santer et al. included: (a) their elimination of surface temperatures over tropical land (roughly 25% of the tropics), (b) using obsolete versions of one SST dataset (ERSSTv2 and v3 rather than ERSSTv3b), and (c) using upper air observations known to be indirectly contaminated by aerosols from the eruption of Mt. Pinatubo. [See Christy, 2010] An account of the events leading up to the publication of the Santer et al. paper was described by Douglass and Christy [2009].

Christy et al. [2010] dealt with some of the issues contained in Santer et al. [2008] by employing and analyzing a different metric - the ratio of troposphere-to-surface trends for each data source to preserve the “apples to apples” comparisons. This effectively normalized all models and observations relative to their own surface trends and also eliminated differences created by differing interannual natural variations among the datasets. From observations, the ratio was calculated to be +0.8 ±0.3. From models, the average was 1.4. These values were significantly different, confirming the results of Douglass et al. (2007).

2. RECENT DEVELOPMENTS.
The most recent papers on this subject are by Santer et al.[2012] and Mitchell et al. [2013]. Santer et al. calculated 1979-2011 equivalent temperature trends for the satellite-based tropical lower troposphere (LT) from the average of 20 latest-generation climate models (CMIP-5) as +0.27 °C/decade. The average surface trend for these models was +0.194 °C/decade. The model-average ratio of troposphere-to-surface trends was therefore 1.39, as has been the case in previous model comparisons. While Mitchell et al. also showed that the basic trend magnitudes between models and observations were significantly different, they then chose a highly selective procedure to recalculate the amplification ratio to give the appearance that models and observations are in better agreement.

3. ANALYSIS OF THE ISSUES RAISED BY THE SANTER ET AL. AND MITCHELL ET AL. PAPERS
While Santer et al. [2012] reported model-mean tropospheric trends that were much above those of observations, they did not address the issue of the tropospheric
amplification of the surface trend values. We have compiled observations for the tropical LT and surface trends for 1979-2011 (i.e. the same period as Santer et al. [2012]) from various data sources are given in the table below (see Christy et al. [2010] for details of datasets). The three metrics of interest here are: (a) average tropospheric trend +0.103 ±0.03 °C/decade (where the error range captures all observations under the assumption that at least one is correct), (b) average surface trend +0.113 °C/decade, and (c) ratio of troposphere-to-surface average trends +0.91 ±0.2, again where the error range captures all possible combinations of the observations.

<table>
<thead>
<tr>
<th>LOWER TROPOSPHERE</th>
<th>1979-2011 Trend °C/decade</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>RICH v1.5</td>
<td>+0.125</td>
<td>Radiosonde balloons</td>
</tr>
<tr>
<td>Raobcore v1.5</td>
<td>+0.085</td>
<td>Radiosonde balloons</td>
</tr>
<tr>
<td>HadAT2</td>
<td>+0.109</td>
<td>Radiosonde balloons</td>
</tr>
<tr>
<td>RATPAC_B</td>
<td>+0.097</td>
<td>Radiosonde balloons</td>
</tr>
<tr>
<td>UAHuntsville</td>
<td>+0.075</td>
<td>Satellite</td>
</tr>
<tr>
<td>Remote Sensing Systems</td>
<td>+0.124</td>
<td>Satellite</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SURFACE</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NOAA/NCDC</td>
<td>+0.111</td>
<td>Surface stations/buoys/ships</td>
</tr>
<tr>
<td>HadCRUT4</td>
<td>+0.115</td>
<td>Surface stations/buoys/ships</td>
</tr>
</tbody>
</table>

In addition, we note that the 20 CMIP-5 models included the effects of volcanoes and ozone, so they would be relatively consistent with the forcings experienced by observations, i.e. “apples to apples”. Thirty-three years is also long enough for the interannual fluctuations due to ENSO to have little impact on the overall trend. The result shows the model-average trend depicts warming of the tropical troposphere that is more than twice the rate of observations since 1979 (0.27 vs. 0.103 °C/decade) - a significant difference. Klotzbach et al. [2009, 2010] found statistically different trends on the global scale. (We note that Santer et al. [2012] also report that global LT trends of models are on average 1.9 times those of observations.)

Further, the ratio of troposphere-to-surface warming rates with 33 years of observations produces values of +0.9 ±0.2, still well below the model-average ratio of 1.4 (see McKitrick et al. [2010, 2011] for a different yet detailed analysis). This suggests a fundamental difference in the processes which govern the vertical flux of heat between observations and models.

Mitchell et al. [2013] attempted to demonstrate that if certain carefully designed conditions are applied, models are not as bad as they seem relative to observations in the tropics. Regarding our first issue, where we show that the absolute magnitudes of trend differences between models and observations are significantly different, they agree. However Mitchell et al. then perform a somewhat deceptive analysis of the
amplification ratio, relying on only one observational dataset RICHv1.5 (the warmest) then generating an unrealistic error margin for the observed amplification ratio and finally claiming that some model results are captured by RICH’s very warm and now wide-range of possibilities.

That RICH is an outlier dataset is demonstrate by calculating the 1979-2008, 300-150 hPa temperature trends for RICH, RAOBCORE, HADAT2 and RATPAC respectively which are +0.13, +0.03, -0.02 and +0.02 °C decade⁻¹. Being at least 0.10 °C decade⁻¹ warmer than the other datasets serves as evidence that RICH is not a reasonable representative for “observations.” Regarding the use of RICH for “observed” amplification ratios, we note that based on the Mitchell et al. error model, RICH does not even encompass the observed ratio values of the other three data sets at several pressure levels (e.g. ratios at 500 hPa for RAOBCORE, HADAT2 and RATPAC are respectively 0.25, 0.36 and 0.47, while RICH’s range is 0.8 to 1.8.) Further evidence of RICH’s unrepresentativeness is seen when we take the results of Mitchell et al. and apply them to our satellite-layer amplification ratio. In this calculation Mitchell et al. would have produced “observed” outcomes which are greater than 1.5. This magnitude strongly violates the calculated observed values of 0.9 ±0.2, demonstrating the weakness of using RICH as the only set of upper-air observations. Finally, it is somewhat disturbing that Mitchell et al. (a) did not cite considerable evidence, much of which contradicts their results, from previous papers (e.g. Christy et al. [2010]; McKitrick et al. [2010, 2011]; Bengtsson and Hodges, [2011]), and (b) did not explain the lack of investigation of the other upper air datasets (HadAT2 dataset - Titchner et al. [2009]; RATPAC dataset - Free et al. [2005]; RAOBCORE dataset - Haimberger et al. [2012] cited by Mitchell et al. but only for the RICH dataset). As a result, the evidence presented by Mitchell et al. does not adequately address the comparison of amplification ratios between models and observations.

Given the above, the original conclusions of Douglass et al. [2007] are confirmed, i.e. that (1) models in general have a basic warm bias in the critical region of the tropics where models predicted a prominent tropospheric “hot spot” to have developed by now, and (2) for a given warming rate of the tropical surface temperature, models produce an amplification of the warming in the upper air that is greater than observed.

4. CONCLUSIONS

After thirteen years and two new generations of climate models (CMIP-3, -5) the results still suggest that model-averages are unable to represent (1) the observed magnitudes of tropical temperature trends throughout the tropospheric levels, and (2) the relationship between the temperature trends of the surface and troposphere. The disparity remains.

REFERENCES


Christopher D. Miller, and William L. Murray, editors. A report by the Climate Change Science Program and the subcommittee on Global Change Research, Washington, DC.


