

# Reexamining the warming in the tropical upper troposphere: Models versus radiosonde observations

Dian J. Seidel,<sup>1</sup> Melissa Free,<sup>1</sup> and James S. Wang<sup>1,2</sup>

Received 13 September 2012; revised 14 October 2012; accepted 16 October 2012; published 22 November 2012.

[1] A recent study of 1979–2010 tropical tropospheric temperature trends in climate model simulations and satellite microwave sounding unit (MSU) observations concluded that, although both showed greater warming in the upper than lower troposphere, the vertical amplification of warming was exaggerated in most models. We repeat that analysis of temperature trends, vertical difference trends, and trend ratios using five radiosonde datasets. Some, but not all, comparisons support the notion that vertical amplification in models exceeds that observed. However, larger ranges of radiosonde trends compared with those for MSU, and the sensitivity of results to the upper-tropospheric level analyzed, make it difficult to conclude unambiguously that models are inconsistent with radiosonde observations. The larger ranges are due to the availability of more radiosonde datasets with different approaches for adjusting measurement biases. Together these two studies highlight challenges of using imperfect observations of tropical tropospheric temperature over a few decades to assess climate model performance.

**Citation:** Seidel, D. J., M. Free, and J. S. Wang (2012), Reexamining the warming in the tropical upper troposphere: Models versus radiosonde observations, *Geophys. Res. Lett.*, 39, L22701, doi:10.1029/2012GL053850.

## 1. Introduction

[2] Examining 1979–2010 tropical tropospheric temperature trends in 19 climate models and in satellite microwave sounding unit (MSU) observations, *Fu et al.* [2011, hereinafter FMJ] found that both observations and models indicate both lower-middle and upper-middle tropospheric warming, with greater warming aloft. However, vertical amplification was exaggerated in most of the models compared with MSU observations. Because the vertical structure of temperature trends is related to tropospheric lapse rate and humidity trends, their finding suggests models may not accurately simulate associated climate feedback processes. Here we use radiosonde data to test the reproducibility of the main conclusion of FMJ – that models show larger vertical amplification of tropical tropospheric warming trends than do observations.

[3] Radiosonde data offer several distinct advantages for this investigation. First, better vertical resolution allows clearer differentiation of upper and lower tropospheric trends. Radio-

sonde reports have 11 mandatory pressure levels between 1000 and 100 hPa, while MSU has two channels that sense emissions from deep layers, which FMJ manipulated to obtain temperatures representative of two deep tropospheric layers [*Fu et al.*, 2004; *Fu and Johanson*, 2005]. Second, there are more long-term radiosonde datasets than MSU datasets, offering a better opportunity to assess uncertainty associated with removing time-varying biases.

[4] Many studies and several comprehensive reports have examined modeled and observed temperature trend differences between the surface and the troposphere, with recent work suggesting that the largest uncertainties are associated with tropical tropospheric trends [see *Climate Change Science Program*, 2006; *Thorne et al.*, 2011, and references therein]. Before FMJ, only a few studies had focused on differences between the upper and lower tropical troposphere (rather than the surface). *Gaffen et al.* [2000] examined 700 and 500 hPa temperature and lapse rate trends and found cooling at both levels in unadjusted radiosonde data during 1979–1997, a shorter time span than FMJ. *Bengtsson and Hodges* [2011] found greater warming in the lower than middle troposphere in MSU data over tropical ocean regions during 1979–2008, but their “middle” tropospheric layer is actually higher than FMJ’s “upper-middle” layer, because FMJ removed the stratospheric influences (including cooling that counteracts tropospheric warming).

## 2. Radiosonde Data, Model Simulations and Methods

[5] This study is based on modeled and observed tropical (20°N to 20°S) monthly tropospheric temperature anomalies for 1979–2010. We focus on 300 and 700 hPa, near the centers of the MSU layers FMJ and *Po-Chedley and Fu* [2012] refer to as the upper-middle troposphere ( $T_{24}$ ) and lower-middle troposphere ( $T_{2LT}$ ) (Figure 1 in FMJ), and 200 hPa, where FMJ report maximum warming in models. The  $T_{24}$  weighting results in temperatures that are more purely representative of the upper troposphere than of the deeper layer sensed by MSU Channel 2 [*Fu et al.*, 2004; *Fu and Johanson*, 2005].

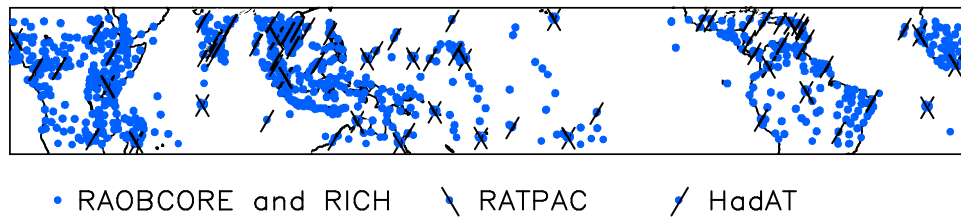
[6] Using tropical-average upper and lower tropospheric temperature anomaly time series, and upper-minus-lower-tropospheric difference time series, we computed trends as in FMJ, using least squares linear regression, with 95% confidence intervals that account for time series autocorrelation. Trends from different datasets and models are compared via two-sided t tests [*Lanzante*, 2005; *Santer et al.*, 2008; FMJ]. The ratio of the two trends (not the trend in time series of anomaly ratios) was also computed for comparison with FMJ.

[7] We use five radiosonde datasets created using four different methods: RATPAC [*Free et al.*, 2005], HadAT

<sup>1</sup>Air Resources Laboratory, NOAA, College Park, Maryland, USA.

<sup>2</sup>Universities Space Research Association, Columbia, Maryland, USA.

Corresponding author: D. J. Seidel, Air Resources Laboratory, NOAA, 5830 University Research Ct., College Park, MD 20740, USA. (dian.seidel@noaa.gov)



**Figure 1.** Locations of tropical ( $20^{\circ}\text{N}$ – $20^{\circ}\text{S}$ ) radiosonde stations with symbols indicating which adjusted datasets include each station. See dataset references for station network details.

[Thorne *et al.*, 2005], RAOBCORE [Haimberger, 2007], and RICH [Haimberger *et al.*, 2008]. (See references for dataset descriptions and acronym explanations.) All are adjusted datasets from which time-varying biases have been removed, but each involves a unique adjustment method and a particular subset of radiosonde stations (Figure 1). Note that RAOBCORE and RICH use information from a reanalysis to locate data discontinuities, and RAOBCORE incorporates reanalysis information to make adjustments. We use two versions of RATPAC: RATPAC-B station data with no adjustments after 1997 and RATPAC-A zonal-mean products adjusted for the full record. We use the most recent versions (1.5) of RAOBCORE and RICH [Haimberger *et al.*, 2012] and note that previous related studies employed earlier versions (e.g., Santer *et al.* [2008] used versions 1.2, 1.3, and 1.4), and that there are large trend differences between versions 1.4 and 1.5 (not shown here). For RAOBCORE, RICH and HadAT, tropical averages were calculated from area-weighted means from four  $10^{\circ}$  latitude bands derived from gridded data. Tropical averages from RATPAC-B were calculated by combining 0000 and 1200 UTC data (when available) for each station, averaging time series for all stations in each  $10^{\circ}$  band, and computing the weighted average of the bands. RATPAC-A averages station data in three longitudinal regions that are then combined to obtain a tropical mean.

[8] Following FMJ, simulations are from Phase 3 of the Coupled Model Intercomparison Project [Meehl *et al.*, 2007] from 36 runs of 19 climate models for 1979–2010, derived by merging 20th Century simulations with 21st Century projections. Multi-model means are computed by averaging ensembles to form means for each model, then averaging (and calculating standard deviations from) the 19 model means.

### 3. Results

#### 3.1. Tests of Tropical Sampling Differences

[9] Compared with MSU observations, radiosonde data have significant gaps due to irregular station spacing and irregular observing programs. We explore the resulting uncertainty in tropical-average temperature trends by (1) subsampling MSU data at radiosonde locations, (2) imposing temporal sampling requirements on station data, and (3) comparing three methods of combining station data to form tropical averages. In each approach, we test the consistency of tropical-average upper-minus-lower-tropospheric temperature trends from differently sampled data. In brief, all three sets of tests showed no statistically significant ( $p = 0.05$ ) trend differences. The second and third sets of tests were performed on four of the five radiosonde datasets, but not on RATPAC-A, for which individual station data are not provided.

[10] Subsampling MSU  $T_{24}$  and  $T_{2LT}$  data [Mears and Wentz, 2009] at the radiosonde station locations of the sparsest dataset, RATPAC, yielded positive trends in  $T_{24}$ ,  $T_{2LT}$  and  $T_{24}$ -minus- $T_{2LT}$  from both the complete and subsampled MSU data. Moreover, the trend differences between the complete and subsampled results were not statistically significant, indicating that the irregular spatial distribution of the radiosonde network does not bias tropical-average trend estimates.

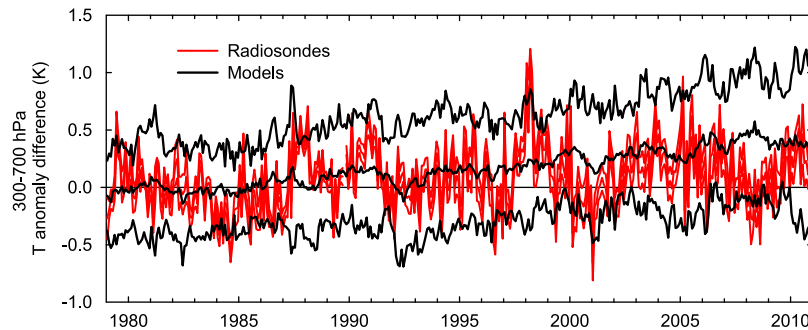
[11] Some radiosonde station records have large gaps during 1979–2010. To test their impact, we compared trends based on all of the stations in each dataset to trends based on subsamples of stations having at least 10 daily observations for a given month, pressure level, and observation time (0000 or 1200 UTC) for at least two thirds of the months during 1979–2010. This reduced the number of stations from 192 to 37 in RAOBCORE and RICH, from 192 to 31 in HadAT, and from 21 to 13 in RATPAC. For all four datasets, trends in tropical-average 300-minus-700 hPa temperature differences, and in 200-minus-700 hPa differences, from the full and reduced station networks showed no statistically significant differences.

[12] To examine the sensitivity of trends to the method of computing tropical averages, recognizing that some stations take one and others take two observations daily, we combined station data three ways: (a) equally weighting all time series within each of three  $120^{\circ}$  longitude regions, allowing two time series at stations with both 0000 and 1200 UTC data, then computing the average of the regions, (b) creating 0000 and 1200 UTC averages for each region, averaging these, then averaging the three regions, and (c) equally weighting time series from all tropical stations to create 0000 and 1200 UTC averages, then averaging these. These methods were compared with each other and with the standard RATPAC-B method described above. As with the tests of temporal sampling effects, there were no statistically significant upper-minus-lower-tropospheric temperature trend differences associated with averaging method.

[13] In summary, for each radiosonde dataset and for MSU data, our tests demonstrated that neither spatial nor temporal sampling choices result in significant differences in tropical-average tropospheric temperature difference trends. For the rest of this paper, we report results from each radiosonde dataset based on its full station network and the RATPAC-B averaging method.

#### 3.2. Upper and Lower Tropospheric Temperature Trends and Difference Trends

[14] FMJ reported 1979–2010  $T_{24}$  and  $T_{2LT}$  warming trends from three and two MSU datasets, respectively, of  $\sim 0.1$  to  $0.2$  K/decade. The climate models also showed  $T_{24}$



**Figure 2.** Monthly 300-minus-700 hPa tropical ( $20^{\circ}\text{N}$ – $20^{\circ}\text{S}$ ) temperature anomalies for 1979–2010, based on mean  $\pm 2$  standard deviations of 5 radiosonde datasets (red) and 36 climate model simulations (black).

and  $T_{2\text{LT}}$  warming but larger than in MSU observations (Table 1 in FMJ). The radiosonde datasets yield a wider range of 1979–2010 tropical tropospheric temperature trends than MSU, and not all are warming trends. Trend ranges are:  $-0.01$  (HadAT) to  $0.15$  K/decade (RICH) at 200 hPa;  $0.11$  (HadAT) to  $0.20$  K/decade (RICH) at 300 hPa; and  $0.06$  (RAOBCORE) to  $0.14$  K/decade (RICH) at 700 hPa. We find these trend ranges in the models:  $0.10$  to  $0.69$  K/decade at 200 hPa;  $0.10$  to  $0.67$  K/decade at 300 hPa; and  $0.06$  to  $0.40$  K/decade at 700 hPa. Thus some model simulations show greater lower tropospheric (700 hPa) warming than any of the radiosonde datasets, and upper-tropospheric trends are more consistent between 300 and 200 hPa in the simulations than in the radiosonde observations.

[15] Figure 2 shows the evolution of upper-minus-lower-tropospheric temperature anomaly differences for the tropics, from radiosondes and models (mean  $\pm 2$  standard deviations). The small spread indicates good agreement among the radiosonde time series, which show strong interannual variability associated with ENSO (FMJ), an expected consequence of the vertical amplification of tropical temperature anomalies (the stronger ENSO signal in the upper troposphere remains in the upper-minus-lower-tropospheric difference [Santer *et al.*, 2005]). The model results show larger spread because of inconsistent timing of ENSO events in these coupled atmosphere-ocean models. *Po-Chedley and Fu* [2012] have recently replicated the main findings of FMJ using atmospheric model simulations with observed sea surface temperature boundary conditions, suggesting that the timing of ENSO events is not the main source of discrepancy between model and MSU trends. Radiosonde temperature anomaly differences have a range of  $\sim 2$  K at any point in time, about five times the range of the MSU differences (compare our Figure 2 with Figure 3 in FMJ), because the amplification is more pronounced between the vertically-separated radiosonde pressure levels than between the broad MSU layers.

[16] Figure 3 (left) presents trends in the vertical differences, for both 300-minus-700 and 200-minus-700 hPa, for all 36 model simulations (and the multi-model mean) and all 5 radiosonde datasets. All model trends are positive, the majority of trends have 95% confidence intervals that do not span zero, and model trends are very similar for the two pairs of pressure levels. Model difference trend values range from  $0.04$  to  $0.28$  K/decade for 300-minus-700 hPa, and from  $0.00$  to  $0.29$  K/decade for 200-minus-700 hPa (vertical lines in Figure 3), substantially greater than the model difference trend range for  $T_{24}$  minus  $T_{2\text{LT}}$ ,  $\sim 0.015$  to  $0.090$  K/decade

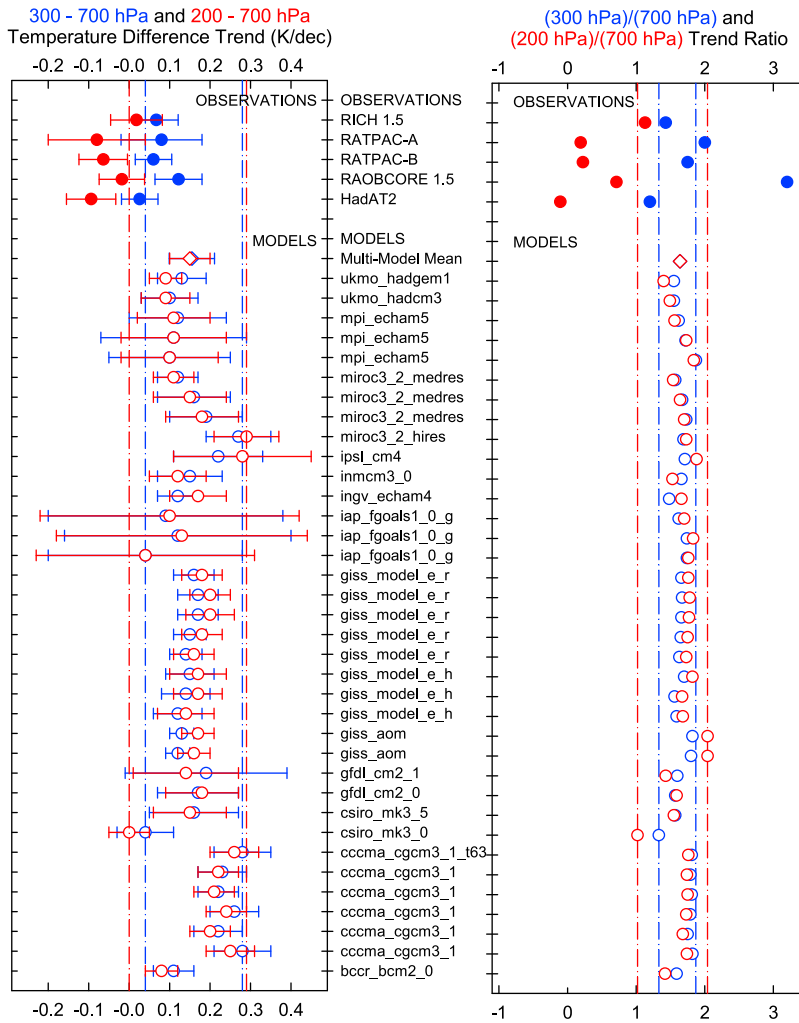
(FMJ). The multi-model mean difference trend is  $0.161 \pm 0.060$  K/decade for 300-minus-700 hPa, and  $0.150 \pm 0.050$  K/decade for 200-minus-700 hPa, both exceeding the  $T_{24}$ -minus- $T_{2\text{LT}}$  trend of  $0.051 \pm 0.007$  K/decade (FMJ).

[17] Ranges of radiosonde difference trends are smaller than the model ranges and shifted toward lower values ( $0.026$  to  $0.122$  K/decade for 300-minus-700 hPa, and  $-0.094$  to  $0.018$  for 200-minus-700 hPa) with systematically more positive difference trends for 300 than 200 hPa. This might be due to remaining time-varying biases in the radiosonde data (which we expect contribute to spurious cooling that is larger at 200 than 300 hPa); alternatively, model trends at 200 hPa might be biased by tendencies to overestimate the heights to which radiative-convective equilibrium dominates radiative equilibrium and to underestimate stratospheric influences on the upper troposphere.

[18] Table 1 compares each radiosonde difference trend estimate with the other four observational results and with the 36 model simulations. Trend estimates and their confidence intervals (K/decade) from each radiosonde dataset (columns 2 and 5) are compared with estimates from 4 other radiosonde datasets (columns 3 and 6) and with estimates from 36 model simulations (columns 4 and 7). Entries in columns 3, 4, 6 and 7 are the percentage of statistically significant trend differences, evaluated with two-sample *t* tests based on  $\pm 2$  standard errors and accounting for time series autocorrelation.

[19] The observational trend estimates tend to show few statistically significant trend differences with the other observational estimates. Not shown in Table 1 is the degree of agreement among the model runs. Briefly, comparisons of pairs of model trends show no statistically significant differences for most of the runs. But comparisons involving 9 of the 36 runs do show significant 200-minus-700 hPa temperature difference trend differences from the majority of other runs. For 300-minus-700 hPa temperature difference trends, 8 of 36 runs show significant differences from the rest.

[20] More relevant to this study are the comparisons of observed and modeled difference trends in Table 1, a large fraction of which show statistically significant differences. For 200-minus-700 hPa, all five radiosonde datasets have significantly smaller trends than the majority (25 to 33) of the 36 model simulations. The disagreement between observations and models is less striking for 300-minus-700 hPa; only 42% of the 180 comparisons (5 radiosonde estimates times 36 model estimates) show significant differences, compared with 84% of the 200-minus-700 hPa comparisons. Overall, there is more disparity between model simulations and observations than within either the group of simulations or



**Figure 3.** (left) Trends in 300-minus-700 hPa (blue) and 200-minus-700 hPa (red) temperature difference and their 95% confidence intervals during 1979–2010 from 36 climate model experiments (open symbols) and 5 radiosonde observations (filled symbols). (right) Ratios of 300-to-700 hPa (blue) and 200-to-700 hPa (red) temperature trends, from the same model runs and radiosonde datasets. Vertical lines indicate model ranges, and the trend in the multi-model mean is shown as an open diamond. After Figure 2 in *Fu et al.* [2011].

the group of observations, and more disparity using 200 hPa than 300 hPa as representative of the upper troposphere.

### 3.3. Temperature Trend Ratios

[21] FMJ found the ratio of  $T_{24}$  to  $T_{2LT}$  trends to be larger in the model simulations (1.20) than in two versions of MSU observations (1.10 and 1.06). Figure 3 presents analogous

300-to-700 hPa, and 200-to-700 hPa trend ratios. The range of model values is smaller for 300-to-700 hPa than for 200-to-700 hPa, but the multi-model means are the same, 1.64, and larger than for the MSU layers, again highlighting the greater amplification obtained using distinct pressure levels. Radiosonde data yield much larger ranges of trend ratios. The range of 300-to-700 hPa ratios easily encompasses all

**Table 1.** Comparison of Observational and Model Estimates of 1979–2010 Trends in Upper (300 or 200 hPa) Minus Lower (700 hPa) Tropical Tropospheric Temperature

Dataset	300 Minus 700 hPa			200 Minus 700 hPa		
	Trend (K/decade)	Obs. (%)	Models (%)	Trend (K/decade)	Obs. (%)	Models (%)
RICH	$0.067 \pm 0.054$	0	44	$0.018 \pm 0.064$	25	69
RATPAC-A	$0.080 \pm 0.100$	0	19	$-0.080 \pm 0.120$	0	89
RATPAC-B	$0.060 \pm 0.045$	0	53	$-0.064 \pm 0.060$	0	89
RAOBCORE	$0.122 \pm 0.058$	25	19	$-0.018 \pm 0.056$	0	83
HadAT	$0.026 \pm 0.045$	25	75	$-0.094 \pm 0.061$	25	92

the model results, but for 200-to-700 hPa ratios, four of the five observational results fall outside the model range.

#### 4. Discussion and Summary

[22] Our tests using radiosonde rather than MSU observations lend some support to FMJ's finding that climate models exaggerate the vertical amplification of tropical tropospheric warming during 1979–2010. However, our results are not as clear, due to two main factors: the wide range of trends in datasets using different methods of adjusting time-varying observational biases, and the dependence of the findings on the discrete pressure level (i.e., 200 vs. 300 hPa) chosen to represent the tropical upper troposphere. Had our scope been more similar to FMJ's (i.e., had we examined only two radiosonde datasets and considered only 200 hPa) the range of results might have been smaller and more supportive of their conclusion. Taken together, the two studies highlight the challenges inherent in using imperfect observations of tropical tropospheric temperature over a few decades to assess climate model performance, and the value of independent satellite and *in situ* observations.

[23] **Acknowledgments.** J.S.W. was supported by the NOAA Air Resources Laboratory through a National Research Council Research Associateship. We are grateful to Qiang Fu, Syukuro Manabe, and Celeste Johanson for generously providing processed model results and for helpful discussions. We thank Yehui Zhang for computer code, and Leopold Haimberger for RICH data. Qiang Fu, Jim Angell and two anonymous reviewers provided constructive review comments.

[24] The Editor thanks two anonymous reviewers for their assistance in evaluating this paper.

#### References

- Bengtsson, L., and K. I. Hodges (2011), On the evaluation of temperature trends in the tropical troposphere, *Clim. Dyn.*, *36*, 419–430, doi:10.1007/s00382-009-0680-y.
- Climate Change Science Program (2006), Temperature trends in the lower atmosphere: Steps for understanding and reconciling differences, *Synth. Assess. Prod. 1.1*, edited by T. R. Karl et al., U.S. Clim. Change Sci. Program, Washington, D. C.
- Free, M., D. J. Seidel, J. K. Angell, J. Lanzante, I. Durre, and T. C. Peterson (2005), Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC): A new data set of large-area anomaly time series, *J. Geophys. Res.*, *110*, D22101, doi:10.1029/2005JD006169.
- Fu, Q., and C. M. Johanson (2005), Satellite-derived vertical dependence of tropical tropospheric temperature trends, *Geophys. Res. Lett.*, *32*, L10703, doi:10.1029/2004GL022266.
- Fu, Q., C. M. Johanson, S. G. Warren, and D. J. Seidel (2004), Contribution of stratospheric cooling to satellite-inferred tropospheric temperature trends, *Nature*, *429*, 55–58, doi:10.1038/nature02524.
- Fu, Q., S. Manabe, and C. M. Johanson (2011), On the warming in the tropical upper troposphere: Models versus observations, *Geophys. Res. Lett.*, *38*, L15704, doi:10.1029/2011GL048101.
- Gaffen, D. J., B. D. Santer, J. Boyle, J. R. Christy, N. E. Graham, and R. J. Ross (2000), Multi-decadal changes in the vertical temperature structure of the tropical troposphere, *Science*, *287*, 1242–1245, doi:10.1126/science.287.5456.1242.
- Haimberger, L. (2007), Homogenization of radiosonde temperature time series using innovation statistics, *J. Clim.*, *20*(7), 1377–1403, doi:10.1175/JCLI4050.1.
- Haimberger, L., C. Tavalato, and S. Sperka (2008), Toward elimination of the warm bias in historic radiosonde temperature records—Some new results from a comprehensive intercomparison of upper-air data, *J. Clim.*, *21*(18), 4587–4606, doi:10.1175/2008JCLI1929.1.
- Haimberger, L., C. Tavalato, and S. Sperka (2012), Homogenization of the global radiosonde dataset through combined comparison with reanalysis background series and neighboring stations, *J. Clim.*, doi:10.1175/JCLI-D-11-00668.1, in press.
- Lanzante, J. R. (2005), A cautionary note on the use of error bars, *J. Clim.*, *18*, 3699–3703, doi:10.1175/JCLI3499.1.
- Mears, C. A., and F. J. Wentz (2009), Construction of the Remote Sensing Systems v3.2 atmospheric temperature records from the MSU and AMSU microwave sounders, *J. Atmos. Oceanic Technol.*, *26*, 1040–1056, doi:10.1175/2008JTECHA1176.1.
- Meehl, G. A., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. F. B. Mitchell, R. J. Stouffer, and K. E. Taylor (2007), The WCRP CMIP3 multimodel dataset: A new era in climate change research, *Bull. Am. Meteorol. Soc.*, *88*, 1383–1394, doi:10.1175/BAMS-88-9-1383.
- Po-Chedley, S., and Q. Fu (2012), Discrepancies in tropical upper tropospheric warming between atmospheric circulation models and satellites, *Environ. Res. Lett.*, in press.
- Santer, B. D., et al. (2005), Amplification of surface temperature trends and variability in the tropical atmosphere, *Science*, *309*, 1551–1556, doi:10.1126/science.1114867.
- Santer, B. D., et al. (2008), Consistency of modelled and observed temperature trends in the tropical troposphere, *Int. J. Climatol.*, *28*, 1703–1722, doi:10.1002/joc.1756.
- Thorne, P. W., D. E. Parker, S. F. B. Tett, P. D. Jones, M. McCarthy, H. Coleman, and P. Brohan (2005), Revisiting radiosonde upper-air temperatures from 1958 to 2002, *J. Geophys. Res.*, *110*, D18105, doi:10.1029/2004JD005753.
- Thorne, P. W., J. R. Lanzante, T. C. Peterson, D. J. Seidel, and K. P. Shine (2011), Tropospheric temperature trends: History of an ongoing controversy, *Wiley Interdiscip. Rev. Clim. Change*, *2*, 66–88, doi:10.1002/wcc.80.