Comments on "Biases in Stratospheric and Tropospheric Temperature Trends Derived from Historical Radiosonde Data"

MELISSA FREE AND DIAN J. SEIDEL

NOAA/Air Resources Laboratory, Silver Spring, Maryland

(Manuscript received 29 June 2006, in final form 19 September 2006)

Randel and Wu (2006, hereafter RW) examine time series of differences between microwave sounding unit (MSU) channel-4 [lower stratosphere (LS)] equivalent radiosonde temperature observations and actual MSU LS satellite data and find many abrupt shifts in those difference series. They conclude that the radiosonde data, although previously adjusted for homogeneity by Lanzante et al. [2003a (hereafter LKS), 2003b], still have large inhomogeneities after 1979 that produce unrealistically large cooling trends in the tropical stratosphere. They also suggest that those inhomogeneities extend into the troposphere and could significantly affect tropospheric trends in the Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC; Free et al. 2005) dataset, which is based on the LKS adjustments.

Their paper raises several questions addressed in the following sections.

1. Are the discontinuities in the difference series in the stratosphere due to errors in the sonde data or the satellite data?

RW argue that the discontinuities in the difference series are not due to problems with the satellite data because 1) they are present in the difference series using both the University of Alabama (UAH; Christy et al. 2003) and the Remote Sensing Systems (RSS; Mears et al. 2003; Mears and Wentz 2005) versions of the satellite data, and 2) they are not present at the same times at all stations as would be expected for satellite errors. The first reason is somewhat convincing, but it is likely that both satellite datasets have errors at similar times, as for example, at the times of transitions between satellites (see Christy and Norris 2006). We discuss the second argument in more detail in this section.

Many of the difference time series (3 out of 6 in RW's Fig. 2 for 1987, 4 out of 6 for 1993–96) show shifts around 1987–88 or 1993–96. From subjective examination of the time series, no other dates show similarly large numbers of shifts in satellite–radiosonde difference series for the LS. This suggests possible common timing for some of the discontinuities. Of the nine difference series that are not classified as having large shifts, four (Darwin, Brownsville, Rio, and Norfolk Island) also show small downward shifts in the difference series somewhere between 1993 and 1995. These times correspond roughly to transitions between the *NOAA-9* and *-10* (1987) and *NOAA-11–NOAA-14* (1995) satellites, when possible problems with the MSU temperatures in the LS have been noted (Christy and Norris 2006). We would need a more detailed study using objective changepoint detection methods (beyond the scope of this comment) to be certain of the significance of this apparent coincidence in timing. It is therefore not immediately clear whether the differences shown in RW are due predominantly to problems with radiosondes or satellites. The absence of shifts in sonde–satellite differences in the stratosphere at some stations like San Juan (Puerto Rico) and Hilo (Hawaii), which are believed to have relatively few changes in instruments or procedures in the satellite period, does nevertheless argue in favor of radiosonde rather than satellite problems.

RW's argument assumes that errors or changed biases in the satellite data due to transitions between satellites will occur uniformly among all locations, and indeed there is little evidence for large geographical variations in satellite biases in the stratosphere. Differences between trends in RSS and UAH MSU LS temperatures at the locations of the tropical LKS stations

---

Corresponding author address: Melissa Free, NOAA/Air Resources Laboratory, 1315 East–West Highway, Silver Spring, MD 20910.
E-mail: melissa.free@noaa.gov

DOI: 10.1175/JCLI4210.1

vary by no more than ~0.2 K decade^{-1}, which is small in comparison with the radiosonde-satellite trend biases of up to 0.7 K decade^{-1}. It is, of course, still possible that both versions of the satellite data contain similar biases that vary strongly with location.

Assuming, as seems reasonable, that many of the large discontinuities shown in RW arise from radiosonde biases, there remains the question of the net effect of such problems on trends. Not all of the trends in the sonde–satellite difference series are due to abrupt shifts, and the comparison in RW's Fig. 8 includes effects of gradual as well as sudden changes. Radiosonde biases could, of course, change gradually due to a series of small shifts that would be difficult to detect individually, but satellite data could be subject to similar gradual shifts. Since the estimates of biases shown in RW could include effects of satellite errors as well as radiosonde problems, they may overstate at least partially the effect of inhomogeneities on radiosonde trends in the stratosphere.

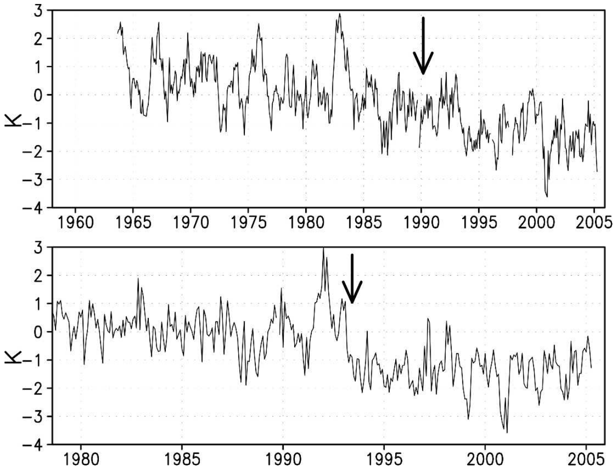

FIG. 1. MSU channel-4 equivalent temperature anomalies (K) at (top) Majuro from RATPAC radiosonde data and (bottom) Hong Kong from RATPAC radiosonde data (0000 and 1200 UTC combined). Only daytime data were used for LKS at Majuro. Arrow shows time of apparent downward shift in the satellite–sonde difference series shown in Randel and Wu (2006).

2. If the RATPAC data contain large remaining biases in the stratosphere, why were these problems not adjusted by LKS? Could they be fixed now?

We have examined data from some of the stations used in RW as examples of inhomogeneities and found two possible reasons.

a. *No apparent change in temperatures*

At several stations in the western tropical Pacific (e.g., Majuro and Truk), shifts in the sonde–satellite difference series occur around 1989–90, coinciding with changes in ground equipment and, at some stations, with the resumption of regular nighttime soundings, which had been sparse or nonexistent since the 1970s. These coincident changes, although not providing a clear reason for a shift, suggest that the shift in the differences could have come from a change in radiosonde observations. As RW pointed out, radiosonde temperatures do not shift noticeably around this time (top of Fig. 1). Since night observations were not available, and since day–night differences were an important aspect of the LKS analysis method, the LKS team had no reason to suspect an inhomogeneity. Furthermore, because the adjustment method depends on differencing the temperatures before and after a change point, it cannot adjust effectively for discontinuities that are masked by natural changes. Without a reference time series, there is no way to reconstruct the true temperature history when a change point coincides with a real change in temperature. LKS deliberately chose not to use the satellite data as a reference so as to produce an independent time series that could be compared to the satellite temperature series, but this choice may have made it harder to detect or adjust for some changes.

The shifts in LS satellite–sonde difference series around 1988 shown in RW at Singapore, Antofagasta, and Ascension similarly show no obvious changes in the LKS time series (not shown). Again, there were little or no night data for comparison.

b. *Coincident natural temperature changes*

Some of the shifts shown in RW occur around the times of large, widespread natural temperature changes such as those after volcanic eruptions. In the presence of a large natural shift, it is difficult to recognize an artificial shift, or to adjust for one if it is seen, without using a reference series that also includes the signature of the natural shift. LKS were aware of this problem and in several cases declined to adjust change points around the Pinatubo eruption, choosing instead to delete data after the change point. In other cases, downward shifts around 1993–95 were assumed to be real. For example, at Hong Kong, LKS noted the drop in 1993 (see bottom of Fig. 1) in the stratosphere but considered it a natural response. Shifts in difference time series at Singapore and Tahiti may also have been considered natural. Again, without an external reference series, it is difficult or impossible to adjust for inhomogeneities at these times. The only apparent way to fix

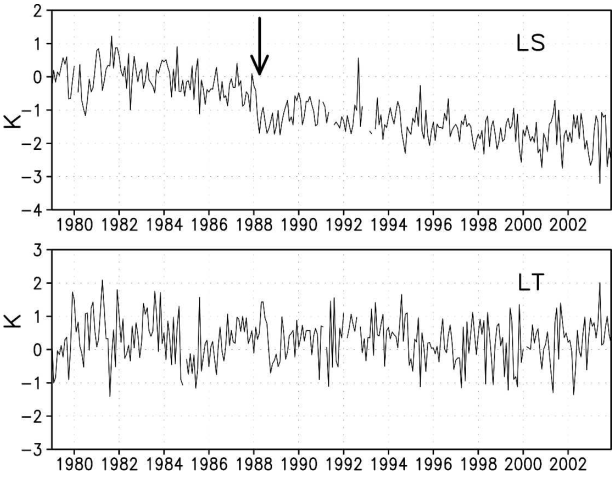

FIG. 2. The LS and LT MSU equivalent temperature anomalies (K) from RATPAC − RSS MSU temperatures at Antofagasta (daytime observations only).

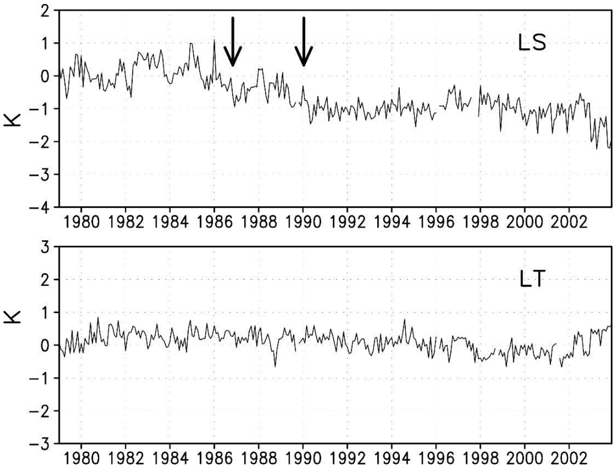

FIG. 3. Same as in Fig. 2, but for Majuro.

these problems would be to alter the LKS method to introduce another reference series such as one derived from neighbor station data (Thorne et al. 2005).

3. Do biases seen in the stratosphere affect tropospheric trends?

a. Discontinuities in the lower-tropospheric temperature difference series

RW separated the tropical stations into "L" stations, with large biases, and "S" stations, with smaller biases, based on the difference between radiosonde and satellite trends in the stratosphere. They then compared vertical profiles of mean trends in radiosonde data for the two sets of stations to estimate the biases in the radiosonde data. We examined time series of radiosonde equivalent channel 2LT (lower troposphere; LT) temperatures made from RATPAC data minus RSS and UAH LT temperatures for the RW L and S stations. In most cases, the major discontinuities found in the LS difference series were not present for the LT. Of 13 downward shifts identified subjectively in the LS at the 7 tropical L stations, 3 corresponded to apparent downward steps in the LT, 8 showed no apparent LT change, and 2 showed upward LT shifts. Conversely, some downward shifts in the LT difference series occur at times that do not show shifts in the LS series. For example, at Antofagasta (see Fig. 2), a large drop occurs in the stratospheric difference series in 1988, but the LT difference series shows no obvious change. At Majuro (Fig. 3), the LT difference series shows a small drop in 1996, in contrast to the LS where drops appear in 1986 and late 1989. Of the S stations, several (Townsville, Norfolk Island, and Nairobi; see Fig. 4) have what appear to be upward discontinuities in the LT difference series. (Upward jumps are also visible in some of the S station difference series for the LS, e.g., Bet Dagan, Rio de Janeiro, and Norfolk Island, and in some of the series at stations not included in either the S or L group.) Often, both upward and downward jumps are visible in the same difference series for LT, with a net result of little or no change.

Since shifts in the tropospheric difference series that coincide with the LS shifts could be too small to see clearly by visual inspection, we took the difference between mean temperatures for 36 months before and after the years of the largest stratospheric shifts seen in the 7 tropical L stations from RW. In the stratosphere,

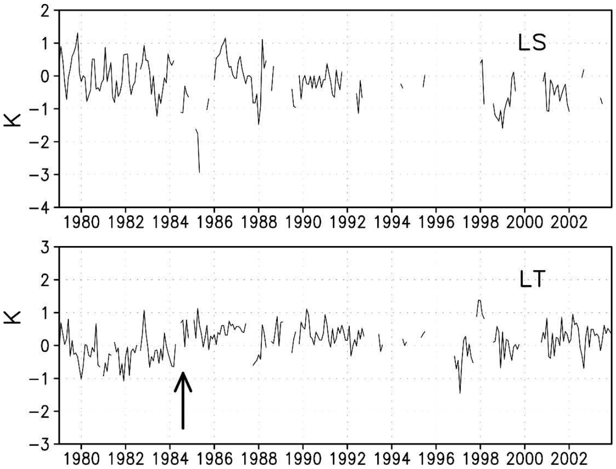

FIG. 4. Same as in Fig. 2, but for Nairobi (0000 UTC observations only). Arrow indicates time of abrupt rise in temperature difference in LT series.

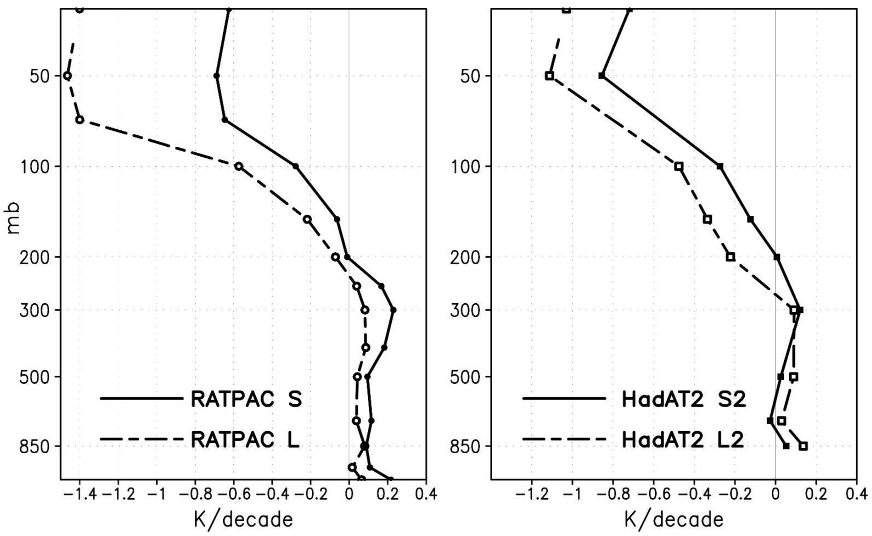

FIG. 5. (left) Trends (K decade^{-1}) in mean temperature for 1979–2004 for stations designated as S and L in RW that fall between 30°N and 30°S, using RATPAC. (right) Least squares linear trends (K decade^{-1}) in mean temperature for 1979–2004 from HadAT2 for tropical stations with smaller (S2) and larger (L2) differences between HadAT2 and RSS MSU lower-stratospheric trends.

median and mean values of these shifts were both about −0.7 K. In the lower troposphere, the changes at these times ranged from +0.4 (Bogota) to −0.2 (Nandi), with a median of +0.02 and a mean of +0.06. Although there were some downward shifts, they were relatively small and balanced by a comparable number of upward shifts in other series. This supports our impression that the large downward jumps shown in RW for the LS do not coincide in time with important downward biases in the LT in the stations considered here.

It seems that problems in the stratosphere are not necessarily reliable predictors of problems in the troposphere, particularly in data like LKS's/RATPAC or HadAT2 (an alternative adjusted radiosonde dataset; Thorne et al. 2005) that have been adjusted by individual level. Our conclusion is consistent with other evidence that artificial shifts in radiosonde data can have complex vertical structures (Nash et al. 2006; Thorne et al. 2005). Furthermore, there are possible positive as well as negative biases in the sonde series, including at least three of the S stations, that could affect the composite trends shown in RW.

b. Effects on trends

RW did not use RATPAC data directly (because the station data for separate observation times were not available at the time of their work) but instead constructed a similar series by supplementing LKS station data with soundings from the Integrated Global Radiosonde Archive (Durre et al. 2006). In most cases the results from RATPAC are similar. We repeated the comparison of trends in the means of the L and S stations for 1979–2004 (RW, their Fig. 8) and got qualitatively similar results for our RATPAC data (Fig. 5, left). While RW omitted postvolcanic eruption periods in calculating trends, we use the full period.

In our data, MSU LT equivalent temperature trends for the L station group range from −0.07 to +0.21 K decade^{-1}, compared to a range from −0.03 to +0.27 K decade^{-1} for the S stations. The mean trend for the L stations for MSU LT is 0.07 K decade^{-1} less than that for the S stations. Since some of the S stations seem to have upward jumps relative to RSS for LT, the comparison between S and L stations in the lower troposphere may overstate the effects of cooling biases in the RATPAC data. If the 3 stations with apparent positive trend biases are removed, the mean LT trend for the remaining S stations goes from 0.15 to 0.11 K decade^{-1}, only 0.03 K decade^{-1} greater than the mean trend for the L stations. Although the total number of tropical stations in the RATPAC dataset is probably adequate for trend analysis (Free and Seidel 2005), there are only seven tropical stations in the L group and nine in the S group, so it is likely that the differences between mean trends for the two groups in the lower troposphere are quite sensitive to the choice of stations. Given these issues, the comparison of trends in the L and S stations is probably not a very reliable measure of the effects of remaining radiosonde biases in the lower troposphere.

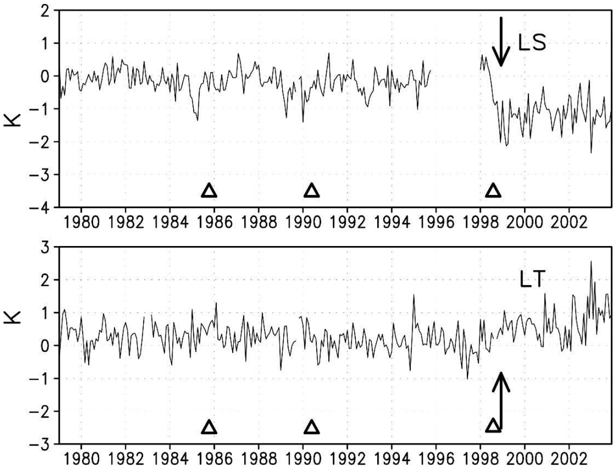

FIG. 6. (top) LS and (bottom) LT MSU equivalent temperature anomalies (K) from HadAT2 − RSS MSU temperatures at Hilo, HI (0000 and 1200 UTC combined). Arrows indicate times of apparent shifts in temperature differences. Triangles indicate times of homogeneity adjustments made in HadAT2.

4. Do similar biases affect HadAT2?

The Met Office has recently created HadAT2 (Thorne et al. 2005) using a different method that relies heavily on near-neighbor comparisons. RW do not address possible biases in that dataset.

Examining the time series for HadAT2 MSU equivalent temperatures minus RSS temperatures shows that some of the same discontinuities shown in RW using a RATPAC-like dataset are also present for HadAT2, but with smaller shifts. In all five of the largest shifts at the L stations in the LS, the size of the shift in the satellite minus radiosonde time series as measured by the change in 3-yr means before and after the shift was reduced in HadAT2 in comparison to RATPAC. The mean shift was ~0.3 K less than in RATPAC. On the other hand, some stations (e.g., Hilo, San Juan, and Nandi) show large discontinuities in difference plots for HadAT2 (see Fig. 6) when little or no discontinuity exists in the RW or RATPAC plots.

HadAT2 station groups with larger and smaller differences in trends from the satellite data in the stratosphere are different from the L and S stations chosen by RW from the RATPAC data. We combined data from the six tropical LKS stations with RSS–HadAT2 trend differences of less than 0.20 K decade^{-1} (S2 stations: Norfolk Island, Nandi, Bangkok, Hong Kong, Brownsville, and Durban) and from the 6 stations with differences of more than 0.40 K decade^{-1} (L2 stations: Easter Island, Manaus, Nairobi, San Juan, Truk, and Hilo). Least-square linear trends (Fig. 5, right) for these two groups of stations in HadAT2 show little difference below 200 mb, with the L2 stations having somewhat larger trends than the S2 stations. Again, this suggests that biases in the stratosphere may not be useful predictors of tropospheric biases in adjusted radiosonde temperature data.

To summarize, the examination of sonde–satellite time series for HadAT2 suggests possible remaining discontinuities in both the stratosphere and troposphere, some of which occur at different times, stations, and levels from those shown in RW for RATPAC. It thus appears that HadAT2 shares at least some of the problems seen in RATPAC.

5. What are the implications for reconciling surface and troposphere or radiosonde and model trends?

Climate model simulations of twentieth-century temperatures show tropospheric temperatures that increase with altitude in the Tropics (Santer et al. 2005), and the trends for the RW S stations in the RATPAC data show a similar increase within the free troposphere. However, for 1979–2004, even in the S stations, the trends in the upper troposphere are no greater than those at the surface in RATPAC (see Fig. 5, left). For 1979–99 (not shown), the period covered by the model simulations shown in Santer et al. (2005), the S stations do show the expected amplification of surface trends in the upper troposphere, but trends in the lower troposphere are still smaller than those at the surface. Trends in the mean tropical surface–troposphere difference series for the S stations are significantly greater than zero for levels below 400 mb for 1979–2004. Thus even the S stations do not show the expected amplification of surface warming in the Tropics for all levels and time periods. This is related to the larger mean surface trends found at S versus L stations. Whether the differences between the observed and modeled trend profiles can be explained by model or observational uncertainties is beyond the scope of this comment.

6. Conclusions

The extent and effect of remaining biases in homogeneity-adjusted radiosonde temperature anomaly data are still not clear and may never be fully understood given the lack of a reliable reference time series. Most of the events shown in the examples in RW do not correspond to shifts in stratospheric temperature in RATPAC data, or else coincide with natural shifts such as the 1994 cooling after the Pinatubo warming. They could easily be inhomogeneities that LKS missed because they were not apparent in the data (and not evident in day–night differences or where night data were not available). On the other hand, many of the shifts cluster around 1987–88 or 1994 (times of satellite transitions that are known to be associated with uncertainties in the satellite series), so at least some could arguably be the result of errors in the satellite rather than sonde data. It is therefore possible that RW overstate the biases in radiosonde data in the stratosphere to some degree.

Assuming that the differences shown in RW for the stratosphere are caused by inhomogeneities remaining in RATPAC, it is still not clear that they affect the troposphere significantly. Discontinuities in the stratospheric difference series do not necessarily coincide with shifts in the lower troposphere, and downward jumps in the stratosphere can be accompanied by upward jumps in lower-tropospheric difference series. In HadAT2, the trends for stations with the largest stratospheric trend biases with respect to satellite data do not differ much below 300 mb from those for stations with less stratospheric bias, yet the tropical mean tropospheric trends in HadAT2 are similar to those in RATPAC. Furthermore, it is likely that the mean lower-tropospheric trends for the small number of stations in the "S" and "L" station groups are sensitive to the exclusion of individual stations and therefore are not a very reliable measure of the effects of inhomogeneities. Thus the existence of large inhomogeneities in the stratosphere provides little useful evidence of potential problems in the mid- and lower troposphere. It seems likely that, as acknowledged in LKS, biases remain in both stratospheric and tropospheric radiosonde datasets, but their extent and significance remain unclear, especially in the lower troposphere.

Acknowledgments. We thank Jim Angell, John Lanzante, Peter Thorne, and Julian Wang for suggestions that improved this comment.

REFERENCES

Christy, J. R., and W. B. Norris, 2006: Satellite and VIZ-radiosonde intercomparisons for diagnosis of non-climatic influences. *J. Atmos. Oceanic Technol.,* **23,** 1181–1194.

——, R. W. Spencer, W. B. Norris, W. D. Braswell, and D. E. Parker, 2003: Error estimates of version 5.0 of MSY/AMSU bulk atmospheric temperatures. *J. Atmos. Oceanic Technol.,* **20,** 613–629.

Durre, I., R. S. Vose, and D. B. Wertz, 2006: Overview of the Integrated Global Radiosonde Archive. *J. Climate,* **19,** 53–68.

Free, M., and D. J. Seidel, 2005: Causes of differing temperature trends in radiosonde upper air data sets. *J. Geophys. Res.,* **110,** D07101, doi:10.1029/2004JD005481.

——, ——, 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.

Lanzante, J. R., S. A. Klein, and D. J. Seidel, 2003a: Temporal homogenization of monthly radiosonde temperature data. Part I: Methodology. *J. Climate,* **16,** 224–240.

——, ——, and ——, 2003b: Temporal homogenization of monthly radiosonde temperature data. Part II: Trends, sensitivities, and MSU comparison. *J. Climate,* **16,** 241–262.

Mears, C. A., and F. J. Wentz, 2005: The effect of diurnal correction on satellite-derived lower-tropospheric temperature. *Science,* **309,** 1548–1550.

——, M. C. Schabel, and F. J. Wentz, 2003: A reanalysis of the MSU channel 2 tropospheric temperature record. *J. Climate,* **16,** 3650–3664.

Nash, J., R. Smout, T. Oakley, B. Pathack, and S. Kurnosenko, 2006: WMO intercomparison of radiosonde systems—Vacoa, Mauritius, 2–25 February 2005. Rep. 83, 115 pp.

Randel, W. J., and F. Wu, 2006: Biases in stratospheric and tropospheric temperature trends derived from historical radiosonde data. *J. Climate,* **19,** 2094–2104.

Santer, B. D., and Coauthors, 2005: Amplification of surface temperature trends and variability in the tropical atmosphere. *Science,* **309,** 1551–1555.

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.