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

WILLIAM J. RANDEL AND FEI WU

National Center for Atmospheric Research, Boulder, Colorado*

(Manuscript received 29 June 2005, in final form 7 October 2005)

ABSTRACT

Temperature trends derived from historical radiosonde data often show substantial differences compared to satellite measurements. These differences are especially large for stratospheric levels, and for data in the Tropics, where results are based on relatively few stations. Detailed comparisons of one radiosonde dataset with collocated satellite measurements from the Microwave Sounding Unit reveal time series differences that occur as step functions or jumps at many stations. These jumps occur at different times for different stations, suggesting that the differences are primarily related to problems in the radiosonde data, rather than in the satellite record. As a result of these jumps, the radiosondes exhibit systematic cooling biases relative to the satellites. A large number of the radiosonde stations in the Tropics are influenced by these biases, suggesting that cooling in the tropical lower stratosphere is substantially overestimated in these radiosonde data. Comparison of trends from stations with larger and smaller biases suggests the cooling bias extends into the tropical upper troposphere. Significant biases are observed in both daytime and nighttime radiosonde measurements.

1. Introduction

Accurate knowledge of past temperature trends is an important aspect of detecting and attributing decadal climate variability and change, and stratospheric changes are a key aspect of the vertical profile. Lower stratospheric temperature changes contribute to fingerprint patterns used in climate change studies (e.g., Tett et al. 1996; Santer et al. 1996, 2005), and are also key for understanding decadal trends within the stratosphere itself (Shine et al. 2003). A current outstanding uncertainty for stratospheric temperature trends regards the magnitude of recent decadal-scale cooling in the tropical lower stratosphere, and agreement (or lack thereof) between radiosonde and satellite observations. Briefly, historical radiosonde data show substantially larger cooling than satellite data in this region. This is seen in comparison of trends from radiosondes and satellites calculated for various periods (e.g., Lanzante et al. 2003b; Seidel et al. 2004), although the issue is confused by the fact that satellite comparisons typically involve a vertical integral over a layer spanning the upper troposphere and lower stratosphere. More pointedly, both the vertical and latitudinal structure of lower stratospheric temperature trends derived from radiosondes (Lanzante et al. 2003b; Thompson and Solomon 2005) disagree with results derived from satellite data [viz. trends derived from the Microwave Sounding Unit (MSU) and Stratospheric Sounding Unit (SSU) shown in Ramaswamy et al. (2001), and updated in the World Meteorological Organization (2003)]. The satellite data show a relative minimum in cooling trends for the tropical lower stratosphere, whereas trends derived from the historical radiosonde record show a maximum in this region. The tropical cooling maximum derived from radiosondes is also in disagreement with current model simulations of past stratospheric temperature changes (Shine et al. 2003; Santer et al. 2005).

Our objective in this paper is to demonstrate that there are significant biases in temperature trends derived from historical radiosonde data, and to explore the magnitude and structure of these biases in one radiosonde dataset. Direct comparisons between collo-

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Corresponding author address: William J. Randel, NCAR, P.O. Box 3000, Boulder, CO 80307.
E-mail: randel@ucar.edu

cated satellite measurements and vertically integrated radiosonde data show that many stations suffer from discontinuities or jumps, probably attributable to changes in instrumentation or radiation corrections that are known to cause long-term cooling biases (Gaffen 1994; Parker and Cox 1995; Gaffen et al. 2000; Elliott et al. 2002; Angell 2003; Lanzante et al. 2003a,b; Christy et al. 2003). These biases are particularly large for stratospheric levels. Although several recent analysis of radiosonde data attempt to account for these changes (Lanzante et al. 2003a; Parker et al. 1997; Thorne et al. 2005), the problems are subtle, pervasive, and difficult to correct accurately. We furthermore examine the vertical structure of radiosonde biases in the Tropics by comparing trend statistics between groups of stations with larger and smaller biases, highlighting the influence on both tropospheric and stratospheric levels. Additionally, the effects of day versus night measurements on bias statistics are briefly examined in the appendix.

2. Data and analyses

The radiosonde data used here are based on the temporally homogenized dataset described in Lanzante et al. (2003a). These data (hereafter called LKS) are based on a near-global network of 87 stations, providing monthly temperature anomalies on standard pressure levels covering the period 1948–97. We focus here on the period after 1979, which overlaps the satellite record. The LKS data are based on careful analysis of measurements at each station, with adjustments made for changes in instrumentation or data processing, as such effects are evident in deseasonalized time series. The LKS data were updated to 2004 using the Integrated Global Radiosonde Archive (IGRA), and we focus on stations with long records of stratospheric data (to at least 50 hPa) that have measurements extending to 2004. We obtained IGRA data for each of the LKS stations, and matched deseasonalized time series at each pressure level to corresponding time series from LKS data. Special care was taken to assure continuity between the LKS and IGRA data across 1997–98, and this was evaluated using comparisons with MSU satellite data as discussed below. We maintain separate 0000 and 1200 UTC datasets, in order to analyze local time behavior (see appendix), but most of our results are based on combined 0000 and 1200 UTC data for stations that have both [for two stations we choose to use only one measurement time, based on our analysis of data quality: Santa Cruz (using 1200 UTC data) and Bet Dagan (0000 UTC)]. Overall the combined LKS + IGRA time series analyzed here are very similar to the Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC) data described in Free et al. (2005).[1]

Satellite observations from the MSU instrument provide measurements of mean temperatures over vertical layers, covering approximately 13–24 km altitude for channel 4 (termed TLS) and 0–12 km for channel 2 (termed TMT). The TLS weighting primarily covers the stratosphere in the extratropics, but spans the upper troposphere and lower stratosphere in the Tropics (the weighting function peaks near the tropical tropopause). TMT also has a contribution from the stratosphere, with ~10% of the integrated signal arising from levels above 15 km. There is also a satellite product derived by differencing the near-nadir and off-nadir scans from MSU channel 2, termed TLT, which has an effective weighting function somewhat lower (~0–8 km) than TMT. TLT has the advantage of having minimal contributions from the stratosphere. The MSU time series are derived from a series of operational satellite instruments, and continuous time series are generated by considering overlaps between adjacent satellites, taking into account orbital changes and calibration effects. Here we consider results based on two different MSU datasets, obtained from Remote Sensing System (RSS) analysis, described in Mears et al. (2003) and Mears and Wentz (2005), and from the University of Alabama at Huntsville (UAH), as described in Christy et al. (2003). The MSU2LT data from UAH are from a recently updated version (v5.2). We use two different satellite datasets because there are some important differences in trends derived between the two, especially for TMT and TLT (see Mears et al. 2003; Santer et al. 2005), and

[1] The version of RATPAC data described in Free et al. (2005) was based on combined 0000 and 1200 UTC observations. Subsequently, an updated version of RATPAC with separate 0000 and 1200 UTC data has become available (with plans for regular updates), and we have made detailed comparisons between our LKS + IGRA data and these RATPAC data. For the majority of the stations the datasets are very similar, and results shown in this paper are nearly identical to those derived from RATPAC. There are some differences in specific time series for a few stations where LKS time series ended before 1997 (due to choices regarding homogeneity rather than data availability). The time series for these stations were not updated beyond 1997 in RATPAC, but were updated to 2004 in our dataset. These stations include: Verkhoyansk, Saint Paul, Annette, Petropavlovsk, Tripoli, Santa Cruz, Dakar, Durban, and Gough Island. The different end points in time series for these stations make trend results substantially different between our LKS + IGRA data and RATPAC. There are also some differences in the "gappiness" of our dataset versus RATPAC for a few stations where we have filled in short gaps in the pre-1997 time series with IGRA data; these stations include: Bermuda, Hilo, Bogota, Nairobi, Manaus, Ascension Island, and Martin de Vivies. Trend results for these stations are very similar between our data and RATPAC.

we wish to demonstrate that our results are not dependent on the specific satellite dataset used.

To directly compare the radiosonde and MSU satellite data, we vertically integrate the radiosonde temperatures using the TLS, TMT, and TLT weighting functions, approximated for the data on standard pressure levels (see Fig. 1 of Seidel et al. 2004). For the TLS comparisons, we integrate the radiosonde data to the 30- or 50-hPa level, depending on data availability at each station; for stations where observations do not reach 50 hPa, the TLS calculations are not performed (data are considered missing). Note that the upper data level is constant for each station, so that the radiosonde minus satellite changes identified here are not a result of variations in upper-level data availability. We focus on LKS + IGRA radiosonde stations over ~70°N–S that have at least 70% of monthly data for the time period 1979–2004; this amounts to 66 stations listed in Table 1.

3. Results

The overall variability observed by satellites and (vertically integrated) radiosonde data agree very well at each station, but time series of differences can reveal small, systematic changes. Figure 1 compares RSS TLS measurements with vertically integrated radiosonde data for Hilo (20°N) and Majuro (7°N), illustrating this behavior. The differences at Hilo show relatively small values throughout the record, with no obvious large changes or discontinuities. In contrast, the time series of radiosonde minus satellite differences at Majuro show an obvious jump of approximately 1 K occurring in ~1989. Note that this jump is not apparent within the variability of the original time series, but is clear in the differences. Consistent with this change, the vertically integrated radiosonde data at Majuro show substantially larger negative trends than the satellite measurements (as shown below).

The discontinuous jumps in satellite–radiosonde differences revealed at Majuro (Fig. 1b) are evident at many stations, and are particularly evident at a number of tropical stations with long-term data records. Figure 2 shows the differences at six additional stations over ~30°N–S showing discontinuous behavior at each station. The exact timing of the jumps is different among the different stations, suggesting the differences are not a result of changes in the satellite data, but rather in the individual radiosonde records. Similar calculations based on UAH data show very similar results, with nearly identical timing to the jumps seen in Fig. 2. The overall effect of the changes is similar at each of the stations in Fig. 2, with a net cooling in the radiosonde–

TABLE 1. LKS + IGRA radiosonde stations shown in Fig. 4, together with the TLS difference trend with respect to RSS data at each station (as in Fig. 4b). Trend results are shown for separate 0000 and 1200 UTC data for stations where both measurement times are available. The bracketed suffixes denote that these stations are combined within the larger (L) and smaller (S) bias tropical station groups, shown in Fig. 8.

Station	Lat	Difference trend for TLS (K decade^{-1})
Syowa (0000/1200 UTC)	−69.0	−0.25/−0.38
Macquarie Island (0000/1200 UTC)	−54.5	−0.30/−0.15
Marion Island (0000/1200 UTC)	−46.8	−0.26/−0.45
Invercargill (0000/1200 UTC)	−46.4	−0.60/−0.62
Chatham Island (0000 UTC)	−43.9	−0.85
Puerto Montt (1200 UTC)	−41.4	−0.59
Gough Island (0000/1200 UTC)	−40.3	−0.28/−0.05
Martin de Vivies (1200 UTC)	−37.8	−0.31
Adelaide (0000/1200 UTC)	−34.9	−0.35/−0.20
Buenos Aires (0000/1200 UTC)	−34.8	−0.61/−0.57
Capetown (0000 UTC)	−33.9	−0.15
Perth (0000/1200 UTC)	−31.9	−0.52/−0.60
Durban (0000 UTC)	−29.9	−0.31
Norfolk Island (0000 UTC) (S)	−29.0	0.14
Antofagasta (1200 UTC) (L)	−23.4	−0.83
Rio de Janeiro (1200 UTC) (S)	−22.8	−0.08
Townsville (0000 UTC) (S)	−19.2	−0.21
Tahiti (0000 UTC)	−17.5	−0.34
Nandi (0000 UTC) (L)	−17.4	−0.65
Darwin (0000 UTC) (S)	−12.4	−0.23
Ascension Island (1200 UTC) (L)	−7.9	−0.80
Manaus (1200 UTC) (S)	−3.1	−0.23
Nairobi (0000 UTC) (S)	−1.3	−0.18
Singapore (0000 UTC) (L)	1.3	−0.72
Bogota (1200 UTC) (L)	4.7	−0.51
Majuro (0000 UTC) (L)	7.0	−0.52
Truk (0000 UTC)	7.4	−0.41
Bangkok (0000 UTC)	13.7	−0.21
Dakar (1200 UTC)	14.7	−0.78
San Juan (0000/1200 UTC) (S)	18.4	−0.12/−0.21
Hilo (0000/1200 UTC) (S)	19.7	−0.20/−0.04
Jeddah (0000/1200 UTC)	21.6	−0.30/−0.17
Hong Kong (0000/1200 UTC) (L)	22.3	−0.60/−0.49
Minamitorishima (0000/1200 UTC)	24.3	−0.47/−0.31
Brownsville (0000/1200 UTC) (S)	25.9	−0.24/−0.19
Santa Cruz (1200 UTC)	28.4	−0.20
Kagoshima (0000/1200 UTC)	31.6	−0.42/−0.24
Bet Dagan (0000 UTC) (S)	32.0	0.14
Bermuda (1200 UTC) (L)	32.3	−0.80
Tripoli (1200 UTC)	32.6	−0.10
Miramar (0000/1200 UTC)	32.8	−0.22/−0.08
North Front (0000/1200 UTC)	36.2	−0.06/−0.16
Dodge City (0000/1200 UTC)	37.7	−0.20/−0.14
Minqin (0000/1200 UTC)	38.7	−0.39/−0.42
Kashi (0000/1200 UTC)	39.4	−0.29/−0.43
Wakkanai (0000/1200 UTC)	45.4	−0.17/0.01
Rostov (0000 UTC)	47.2	0.10
Great Falls (0000/1200 UTC)	47.4	−0.30/−0.42
Torbay (0000/1200 UTC)	47.6	0.04/−0.43
Munchen (0000/1200 UTC)	48.2	−0.06/−0.18
Moosonee (0000/1200 UTC)	51.2	0.10/−0.04
Petropavlovsk (0000/1200 UTC)	53.0	0.35/0.06

TABLE 1. (*Continued*)

Station	Lat	Difference trend for TLS (K decade^{-1})
Omsk (0000/1200 UTC)	54.9	0.20/0.18
Annette Island (0000/1200 UTC)	55.0	−0.17/−0.46
Saint Paul Island (0000/1200 UTC)	57.1	−0.16/−0.56
Kirensk (0000/1200 UTC)	57.7	0.05/0.01
Lerwick (0000/1200 UTC)	60.1	0.20/0.05
Keflavik (0000/1200 UTC)	64.0	−0.08/−0.02
Baker Lake (0000/1200 UTC)	64.3	−0.15/−0.12
Pechora (0000 UTC)	65.1	0.06
Angmagssalik (0000/1200 UTC)	65.6	−0.52/−0.59
Turuhansk (0000/1200 UTC)	65.8	0.16/0.11
Sodankyla (0000/1200 UTC)	67.4	−0.58/−0.51
Verkhoyansk (0000/1200 UTC)	67.6	−0.03/−0.43
Jan Mayen (0000/1200 UTC)	70.9	−0.44/−0.57
Point Barrow (0000/1200 UTC)	71.3	−0.47/−0.48

satellite differences of order 1–2 K over the 1979–2004 record. This artificial cooling is consistent with previous analyses of radiosonde data, which show that changes (improvements) in instrumentation over time can result in spurious cooling (e.g., Elliott et al. 2002). While the LKS analyses have attempted to correct for such factors, Fig. 2 is evidence that small systematic shifts can still remain uncorrected.

For comparison, Fig. 3 shows radiosonde–satellite differences at a number of low latitude stations that do not exhibit large systematic changes over time. While there can be substantial variability in the difference fields at some stations (e.g., Nairobi), the lack of systematic changes at a number of stations is evidence that there are not sizeable inhomogeneities in the RSS TLS satellite data (which would appear as jumps at each station), and further supports that the jumps in Fig. 2 are related to radiosonde data.

The fact that the radiosonde changes often occur as jumps suggests an association with specific changes in instrumentation or data analysis. Sometimes such changes are documented in station metadata (often with inexact dates for changes), but for other stations metadata is absent or incomplete. For example, the cooling at Kagoshima after 1993 (top, Fig. 2) is probably associated with incorporating a new radiation correction in January 1993 and a change in radiosondes (from MEISEI RSII-80 to R-91) in May 1994. At Singapore there was a change from Vaisala RS21 to RS80 during 1982–89 (exact date unknown), and a change in radiation correction in December 1992, and instruments at Ascension Island changed from Vaisala to Space Data to VIZ during 1986–92 (exact dates unknown). The cooling at Antofagosta in 1987 (bottom, Fig. 2) is likely associated with a change from VIZ to Vaisala RS80 radiosondes in June 1987. At Perth, radiosonde type changed from Philips MkIII to Vaisala during 1987–89, and there was a change in radiation correction in December 1992 (near the time of the change observed in Fig. 2). These associations strongly suggest that the jumps are associated with radiosonde-related changes, although the exact dates may be un-

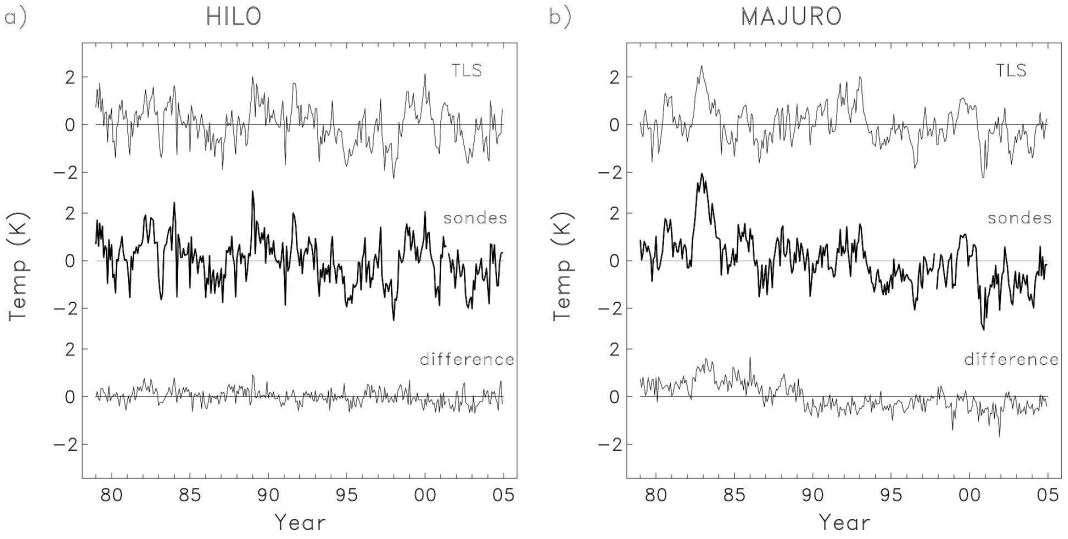

FIG. 1. Top curves show deseasonalized temperature anomalies from RSS TLS observations at (a) Hilo (20°N) and (b) Majuro (7°N). Middle curves show the respective time series from radiosonde observations, vertically integrated with the TLS weighting function. Bottom curves show the integrated radiosonde minus satellite observations at each station.

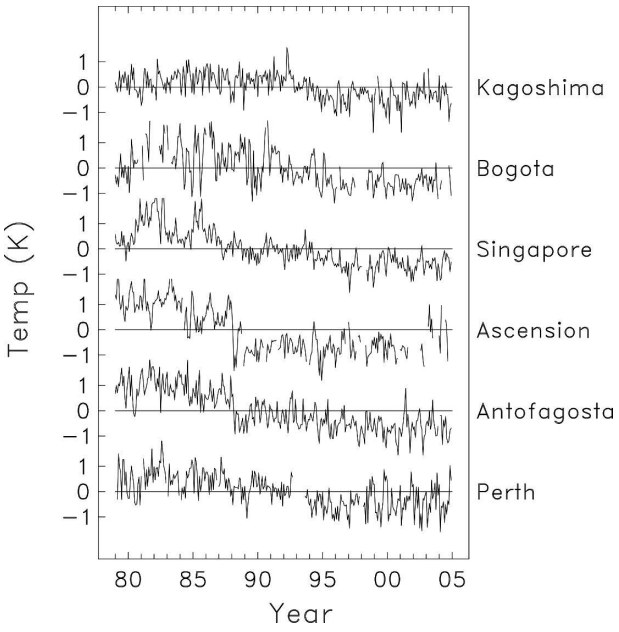

FIG. 2. Time series of vertically integrated radiosonde minus RSS TLS satellite temperature observations at each of the listed stations. Note the discontinuities and long-term cooling evident at each station.

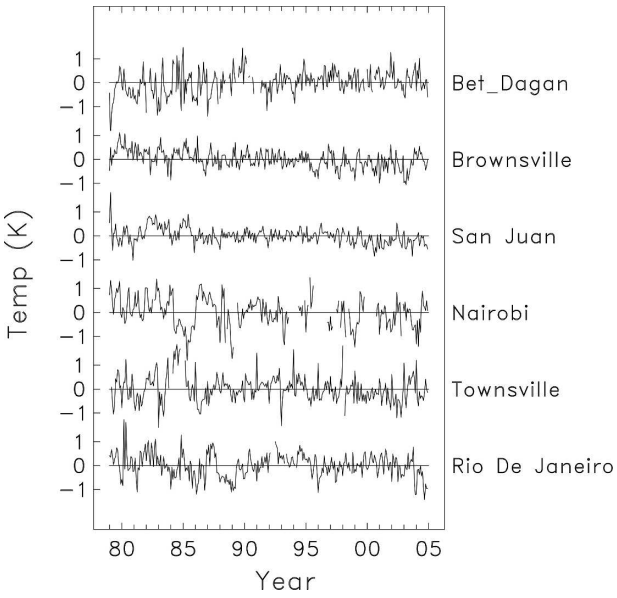

FIG. 3. Time series of vertically integrated radiosonde minus RSS TLS satellite temperature observations at the listed stations. Each of these stations shows relatively stable differences in time with respect to TLS measurements.

known. On the other hand, there are no documented instrument or data handling changes at Majuro associated with the cooling after ~1989 (Fig. 1b), or for similar cooling observed at Truk or Bermuda (not shown here).

To quantify the cooling biases in the vertically integrated radiosonde and satellite datasets, we calculate linear trends for the period 1979–2004. Trends are derived from a standard linear regression model, with statistical uncertainty estimates evaluated via resampling techniques (which include the effects of serial autocorrelation; Efron and Tibshirani 1993). Also, in order to minimize the effects of volcanic eruptions on the stratospheric trend results, we simply omit two years of data following the eruptions of El Chichon (April 1982) and Pinatubo (June 1991); note the transient stratospheric warming for these periods seen in Fig. 1. We show two sets of trend results, namely (i) directly comparing trends from the zonal mean satellite data and radiosonde datasets, and (ii) calculating trends for the (radiosonde minus satellite) differences at each station. Statistical uncertainties for (i) include a large component associated with "natural" climate variability (common to both datasets, as seen in Fig. 1), and so overestimate the uncertainties of radiosonde minus satellite differences; these trend results are useful, however, to understand the context and relative size of the radiosonde biases. Uncertainties calculated for the difference trends [(ii) above] more directly quantify the statistical uncertainty of biases between the radiosondes and collocated satellite data.

Figure 4a compares trend results for TLS and integrated radiosonde data; we include satellite results from both the RSS and UAH TLS data in Fig. 4; the UAH data give slightly larger cooling trends, but with very similar overall structure. Both the satellite data and radiosondes exhibit cooling over much of the globe, but the trends are substantially larger in the radiosonde data. The differences are especially large in the Tropics, where the satellite trends show a relative minimum, but the radiosondes show maximum trends. Figure 4b shows the corresponding plot of trends in the radiosonde minus satellite time series at each station; in this case the error bars reflect the statistical uncertainty of trend biases (i.e., stations where error bars do not cross zero highlight significant trend biases). The difference trends shown in Fig. 4b are tabulated in Table 1, allowing identification of biases at each individual station. The differences in Fig. 4b (and Table 1) are based on RSS data, but are very similar to results based on UAH data. A large fraction of the radiosonde stations exhibit significant biases in Fig. 4b, and biases are especially large in the Tropics (~30°N–S) and at a number of stations in Southern Hemisphere (SH) midlatitudes. Most of the trend biases isolated in Fig. 4b are associated with obvious discontinuities in the difference time series, as illustrated in Fig. 2. Thus we believe that

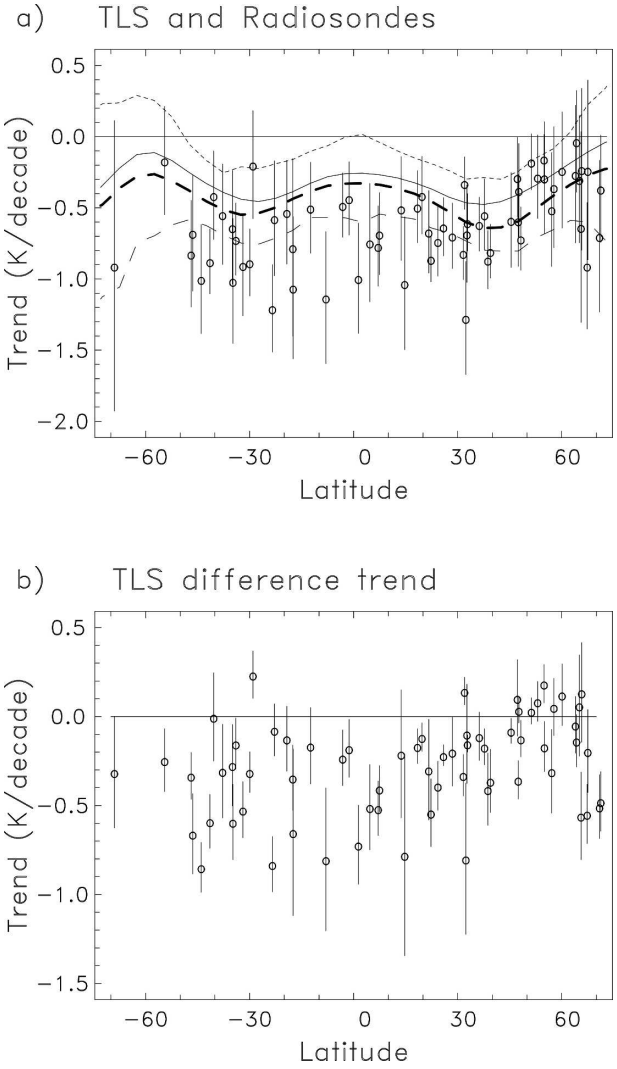

FIG. 4. (a) Latitudinal profile of temperature trends over 1979–2004 from zonal mean TLS and vertically integrated radiosonde data. The solid and dark dashed lines show TLS trends derived from RSS and UAH data, respectively, and the light dashed lines indicate two-sigma uncertainty levels (RSS plus two-sigma, and UAH minus two-sigma). The circles with error bars show the corresponding trends and uncertainties from each radiosonde station. (b) Corresponding trends of the radiosonde minus satellite TLS time series at each station derived using RSS data. Error bars denote two-sigma uncertainty levels for the difference trends.

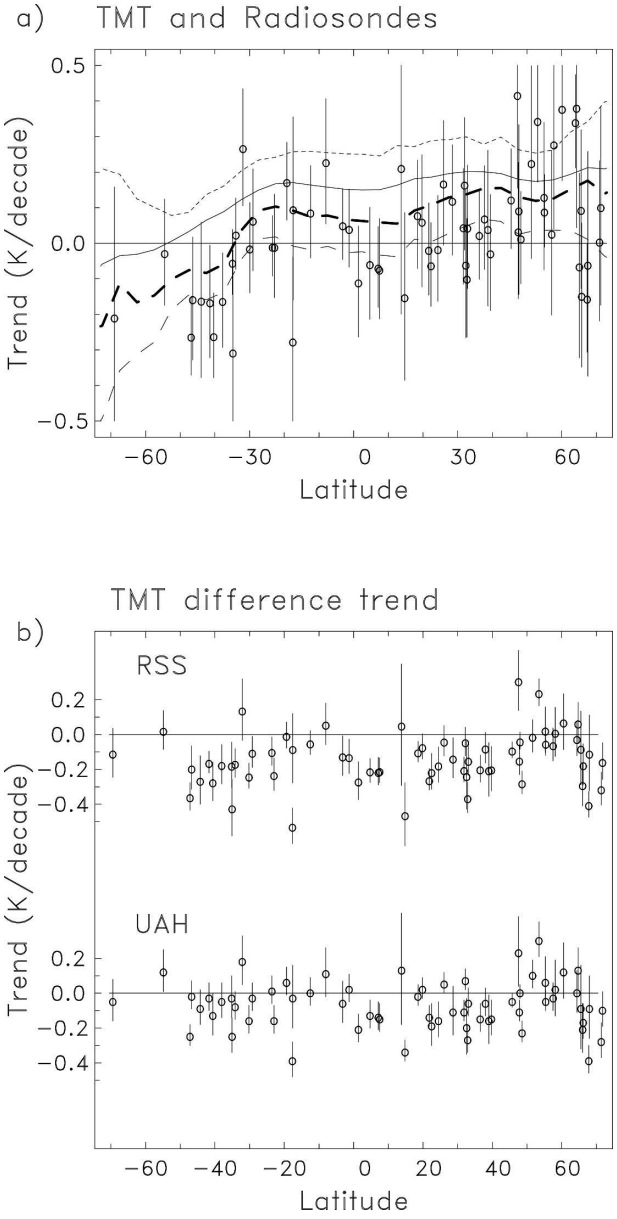

FIG. 5. (a) Latitudinal profile of temperature trends over 1979–2004 from zonal mean TMT and vertically integrated radiosonde data. The plotting convention for trends and uncertainties is the same as in Fig. 4a. (b) Corresponding trends in the radiosonde minus satellite TMT time series at each station. Separate results are shown based on the RSS and UAH satellite datasets.

the jumps, often associated with radiosonde instrumentation changes, are responsible for the significant cooling biases in these radiosonde data.

Similar discontinuities and differences in trends are evident in comparisons between radiosondes and TMT. Figure 5a shows a comparison of zonal mean trends for 1979–2004 from TMT with vertically integrated radiosonde data, and Fig. 5b shows corresponding difference trends calculated RSS and UAH data (these are shown separately, because the detailed results depend on which satellite dataset is used). Note the trend values in Fig. 5a are relatively small for TMT (of order 0.1 K decade^{-1}), and the differences between the RSS and UAH datasets are of this same magnitude (with UAH giving smaller trends). The radiosonde–satellite trend comparisons in Fig. 5a show systematic patterns, including (ii) an overall similar latitudinal structure, with

trends becoming more negative in each dataset south of ~30°S, and (ii) an overall systematic negative bias of the radiosonde trends compared to the satellite data. This is quantified in Fig. 5b, where differences with respect to the RSS data show significant negative biases for the majority of stations. Differences calculated with respect to UAH show smaller overall biases, but there are still a large number of stations with significant negative values. Time series of TMT satellite–radiosonde differences shows that these biases are often associated with discontinuities, similar to TLS. An example comparison for TMT is shown for Singapore data in Fig. 6. Here the satellite TMT datasets show warming trends, while the radiosondes give a (small) cooling trend. Difference time series show a jump between the first and second half of the time series, and similar results are found using both RSS and UAH satellite data. In general the radiosonde biases for TMT (Fig. 5b) are less than for TLS (Fig. 4b), and this is consistent with larger radiosonde trend biases in the lower stratosphere (as shown below).

The corresponding radiosondes minus satellite difference trends for TLT are shown in Fig. 7, and overall the results are similar to those for TMT (Fig. 5b). The satellite–radiosonde difference trends based on RSS data show significant negative values for a large number of stations, suggesting that the biases are not solely a stratospheric level effect. Smaller biases are found based on UAH data, but the overall tendency is for generally negative biases.

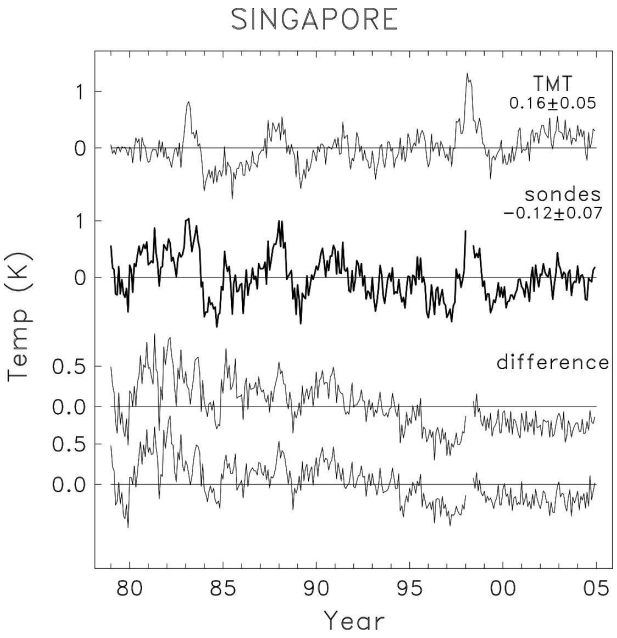

FIG. 6. Top curve shows time series of deseasonalized temperature anomalies from RSS TMT data at Singapore (1°N), and the darker next curve shows time series from vertically integrated radiosonde observations. Linear trends and uncertainties for these data are also shown, with units of K decade^{-1}. The lower curves shows the radiosonde minus satellite differences, based on RSS (upper) and UAH (lower) satellite data.

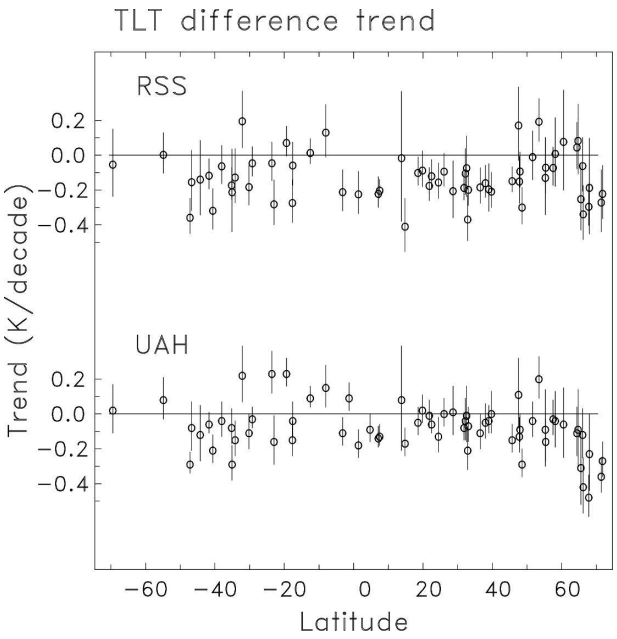

FIG. 7. Latitudinal profile of trends in radiosonde minus satellite TLT time series at each station. Separate results are shown based on the RSS and UAH satellite datasets.

A simple estimate of the vertical structure of tropical radiosonde biases can be generated by comparing stations with larger and smaller trend biases (identified from Fig. 4b). We identify stations over ~30°N–S where the difference trends from Fig. 4b are larger than −0.5 K decade^{-1} (larger biases), and smaller than −0.3 K decade^{-1} (smaller biases). This yields eight stations for the larger biased group, and 10 for the smaller biased group; the individual stations are noted in Table 1. Note that although Bangkok (latitude 14°N) has small trend biases, the time series of differences (not shown) reveals large variations of random sign that suggest data homogeneity problems and this station is not included in the smaller bias group. Figure 8a shows ensemble temperature trends derived from these groups of stations; that is, trends are calculated for each station and averaged. Uncertainties for the ensemble trends are calculated treating each station as being independent. While this is a small data sample with limited geographical coverage, clear systematic differences in trends are evident. Stronger negative trends are evident in the larger bias group, with a factor of two difference in the lower stratosphere (100–30 hPa). Significant trend differences extend to lower levels (~300 hPa).

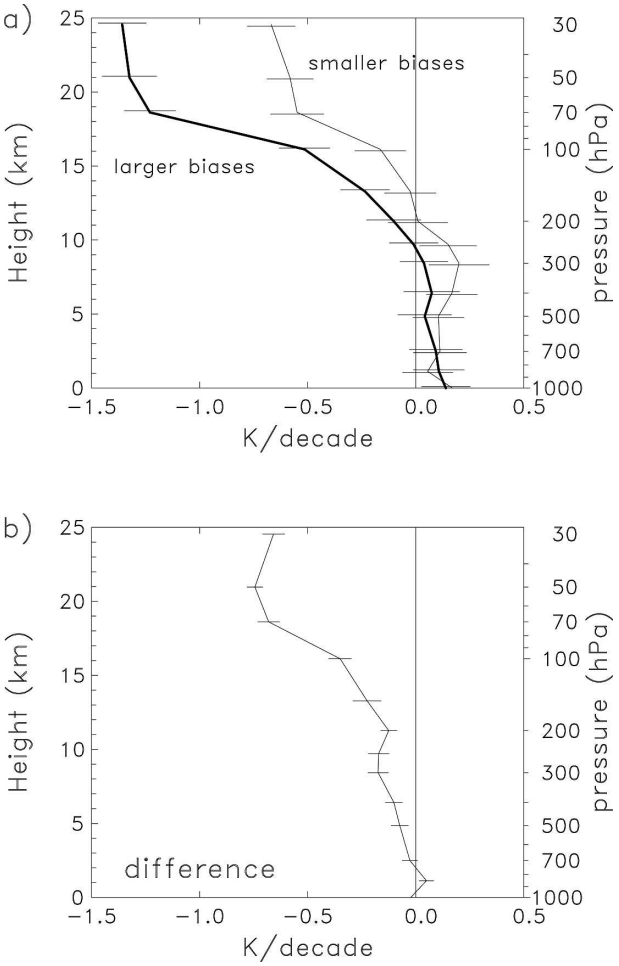

FIG. 8. (a) Vertical profile of temperature trends over 1979–2004 derived from the groups of larger bias and smaller bias tropical radiosonde stations, discussed in the text and identified in Table 1. The uncertainty estimates assume that trends from each station are independent. (b) Trends of the time series of temperature differences between the groups of larger bias and smaller bias tropical stations.

Note that for the smaller bias ensemble, positive temperature trends increase with height in the troposphere, consistent with model and theoretical expectations for near-surface warming (Santer et al. 2005); this increase is not found in the larger bias ensemble.

Vertical profile trend differences are examined further by first averaging time series within each group of stations, then calculating trends of the difference time series; this increases statistical significance by first removing variability that is common between the datasets, such as ENSO variability. The difference trends are shown in Fig. 8b, showing statistically significant differences that span the lower stratosphere and extend throughout the upper troposphere, down to near the 500-hPa level. Note that although the difference trends are small in the troposphere (~0.1 K decade^{-1}), this is the same order as the trend signal itself (i.e., Fig. 8a). The relatively deep vertical structure of the trend differences in Fig. 8 is consistent with a strong influence on TLS, with smaller effects on TMT and TLT (as shown above).

The results in Fig. 8 are based on a small sample of stations within each group, with limited geographical coverage. To determine if this sampling can significantly influence the results, we compare satellite-derived trends and differences at the corresponding station locations. Results show no significant trend differences between the data sampled at the larger bias locations versus the smaller bias locations; for example, TLS trends are -0.26 ± 0.08 (-0.28 ± 0.09) K decade^{-1} for the larger (smaller) bias groups, and there is a similar lack of significant differences for the TMT and TLT calculations. This sampling exercise suggests that the radiosonde trend differences highlighted in Fig. 8 are in fact due to differences in the radiosonde biases between the groups, rather than spatial sampling effects.

4. Summary and discussion

Calculating differences between collocated satellite measurements and vertically integrated radiosonde temperatures can reveal jumps and discontinuities in individual station records. These jumps are not evident within the variability of the individual datasets, but are apparent when differences with satellite measurements are taken. We find that many stations of the LKS + IGRA radiosonde dataset exhibit such discontinuities, and that a large number of the problematic stations occur in the Tropics. The fact that the jumps occur at different times for different stations (Fig. 2) suggests that the problems originate primarily with the radiosondes rather than the satellite data; this is corroborated by station metadata (available for most locations, but likely incomplete), and also by the lack of changes at a number of other stations (Fig. 3). The net effect of the jumps is a systematic tendency for spurious cooling in the radiosonde data at each of the identified stations. This tendency for artificial cooling in the radiosonde data is likely associated with changes (improvements) in radiosonde instrumentation over the last several decades; for many stations the observed jumps are coincident with changes in the radiation correction applied to the data. While the LKS data have undergone adjustments to minimize such changes (which reduces the tropical stratospheric trends compared to unadjusted data; Lanzante et al. 2003b), the detailed comparisons here show that problems persist at many stations. As discussed in the appendix, cooling biases are observed

for both daytime and nighttime radiosonde measurements, suggesting the problem is not due solely to solar shortwave effects.

It is important to stress the importance of station metadata for both identifying and correcting changes in historical radiosonde data. Historically there has been less effort directed to metadata retrieval than data retrieval, although notable efforts were undertaken by Gaffen (1996) and subsequent updates by the National Climate Data Center (NCDC). Without metadata investigators must rely on the use of statistical inference to identify and adjust for nonclimatic changes. Metadata describing changes in equipment and procedures allows investigators to identify the likely timing, sign, and magnitude of any nonclimatic effects, and hence apply more optimal adjustments. It should be noted that metadata completeness was a criteria in station selection in LKS and that the rest of the historical network typically has much worse metadata.

There are relatively few radiosonde stations in the Tropics with continuous long-term records of lower stratospheric temperatures (24 stations in the LKS + IGRA dataset over 30°N–S. Our results (Fig. 4) show that about half of these LKS + IGRA tropical stations exhibit spurious cooling trends associated with artificial jumps in the data. The results are most problematic for the near-equatorial region, where most of the LKS + IGRA stations have substantial cooling biases (see Fig. 4). A comparison of the vertical structure of trends from the biased and unbiased tropical stations (Fig. 8) suggests that the biased radiosonde records can overestimate lower stratospheric trends (over 100–30 hPa) by a factor of 2. Significant differences extend to the upper troposphere. The fact that a large fraction of the tropical stations with long stratospheric records are significantly biased suggests that trend results derived from an ensemble of these data should be viewed skeptically. In particular, these results call into question the reality of a cooling trend maximum in the tropical lower stratosphere derived from the LKS + IGRA radiosonde data (Thompson and Solomon 2005).

We also find discontinuities and systematic trend biases for radiosonde comparisons with TMT, and these TMT differences can be substantial in the Tropics. As with the TLS comparisons, the differences with TMT can be seen as jumps in the difference time series (Fig. 6). Our work does not distinguish between the tropospheric and stratospheric contributions to the TMT differences, but we find similar biases in comparisons with TLT (Fig. 7), which argues that radiosonde biases in fact influence tropospheric levels. The profile trend comparisons (Fig. 8) suggest that radiosonde trend biases extend to the upper or middle troposphere, and this argues that there are substantial uncertainties in tropical radiosonde trends over much of their domain [note a similar conclusion was reached by Lanzante et al. (2003b)].

The biases highlighted here suggest that the LKS + IGRA radiosonde data are problematic for analysis of the vertical profile of tropical temperature trends. While the homogenization procedures applied to the LKS dataset have substantially reduced artificial cooling biases (Lanzante et al. 2003b), large biases persist at many tropical stations. We note that other datasets that rely heavily on historical radiosonde observations [viz. the Hadley Center datasets, Parker et al. (1997), Thorne et al. (2005), and National Centers for Environmental Prediction (NCEP) reanalysis] show large tropical stratospheric cooling trends, similar to the LKS + IGRA or RATPAC data (Seidel et al. 2004; Free et al. 2005; Thompson and Solomon 2005). This suggests that the problems identified here may also influence other datasets, although this needs to be confirmed by further analyses. It will be important to minimize such biases in future analyses of the historical radiosonde data record.

Acknowledgments. The LKS and RATPAC radiosonde data sets were obtained courtesy of John Lanzante and Melissa Free, the UAH MSU data were obtained courtesy of John Christy, and the RSS MSU satellite data was obtained from Carl Mears and from the RSS Web site: http://www.remss.com. We thank Dian Seidel and Dave Thompson for several discussions during the course of this work; Kevin Trenberth, John Lanzante, Carl Mears, Junhong Wang, and Melissa Free for comments on the manuscript; and Peter Thorne and an anonymous reviewer for suggestions that helped clarify and improve the paper. This work is partially supported under the NASA ACMAP program.

APPENDIX

Day versus Night Effects in Radiosonde Biases

The radiosonde instrumentation changes probably produce cooling trends because instrument design and the correction for radiation effects (both shortwave and longwave) improve over time. The effects of solar heating on radiosonde measurements are well known (e.g., Luers and Eskridge 1998), and significant changes due to solar heating effects were noted in U.S. stations that changed from VIZ to Vaisala radiosondes in the middle 1990s (Elliott et al. 2002). Such effects are clearly seen in comparing day versus night measurements, and day–night differences are a useful diagnostic to isolate radiosonde changes (and used in constructing the LKS

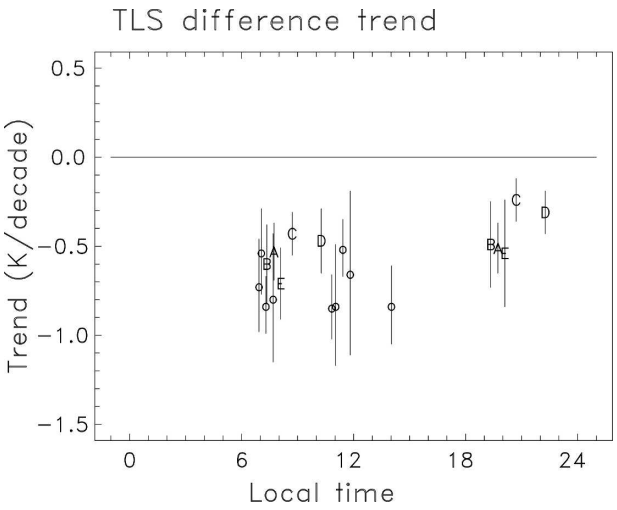

FIG. A1. Trends in radiosonde minus satellite TLS time series (using RSS data), for tropical radiosonde stations with significant biases, plotted as a function of local time of the measurements. Error bars denote two-sigma uncertainty levels. The letters refer to five stations with both 0000 and 1200 UTC measurements (A=Perth, B=Hong Kong, C=Kagoshima, D= Minamitorishima, and E=Buenos Aires).

dataset; Lanzante et al. 2003a). Accordingly, it would be instructive to diagnose day–night differences for the tropical station biases shown here.

Figure A1 shows TLS radiosonde minus satellite difference trends (as in Fig. 4b), but plotted for the individual 0000 and 1200 UTC measurements as a function of local time of the measurements. Results are shown for each of the stations over 35°N–S with large trend biases (from Fig. 4b). The stations in Fig. A1 include five with separate 0000 and 1200 UTC measurements (Perth, Hong Kong, Kagoshima, Minamitorishima, and Buenos Aires), that equate to local morning and evening measurements (the evening measurements are all made in the dark). Results in Fig. A1 show that the largest trend biases occur for daytime measurements, but that significant biases are also observed for nighttime data. Direct comparison of the day–night differences at the five stations with both measurement times shows that the daytime trends are only slightly larger than nighttime (by ~0.1–0.2 K decade^{-1}), and this difference is consistent among four of the five stations (the day–night biases are similar at Perth). These results suggest that solar heating effects contribute somewhat to trend biases (because the daytime trend biases are slightly larger), but are not the only cause of the large biases for the tropical stations studied here. Sherwood et al. (2005) have suggested that daytime biases can account for a substantial component of long-term radiosonde trend biases. The results in Fig. A1 demonstrate that significant biases can also occur for nighttime data.

REFERENCES

Angell, J. K., 2003: Effect of exclusion of anomalous tropical stations on temperature trends from a 63-station radiosonde network, and comparison with other analyses. *J. Climate,* **16,** 2288–2295.

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

Efron, B., and R. J. Tibshirani, 1993: *An Introduction to the Bootstrap*. Chapman and Hall, 436 pp.

Elliott, W. P., R. J. Ross, and W. H. Blackmore, 2002: Recent changes in NWS upper-air observations with emphasis on changes from VIZ to Vaisala radiosondes. *Bull. Amer. Meteor. Soc.,* **83,** 1003–1017.

Free, M., D. J. Seidel, J. K. Angell, J. Lanzante, I. Dume, and T. Peterson, 2005: Radiosonde atmospheric temperature products for assessing climate (RATPAC): A new dataset of large-area anomaly time series. *J. Geophys. Res.,* **110,** D22101, doi:10.1029/2005JD006169.

Gaffen, D. J., 1994: Temporal inhomogenities in radiosonde temperature records. *J. Geophys. Res.,* **99,** 3667–3676.

——, 1996: A digitized metadata set of global upper-air station histories. NOAA Tech. Memo. ERL-ARL 211, Silver Spring, MD, 38 pp.

——, M. Sargent, R. E. Habermann, and J. R. Lanzante, 2000: Sensitivity of tropospheric and stratospheric temperature trends to radiosonde data quality. *J. Climate,* **13,** 1776–1796.

Lanzante, J., S. 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 comparisons. *J. Climate,* **16,** 241–262.

Luers, J. K., and R. E. Eskridge, 1998: Use of radiosonde temperature data in climate studies. *J. Climate,* **11,** 1002–1019.

Mears, C. A., and F. J. Wentz, 2005: The effect of diurnal correction on satellite-derived lower tropospheric temperature. *Science,* **309,** 1548–1551, doi:10.1126/science.1114772.

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

Parker, D. E., and D. I. Cox, 1995: Towards a consistent global climatological rawinsonde data base. *Int. J. Climatol.,* **15,** 473–496.

——, M. Gordon, D. P. N. Cullum, D. M. H. Sexton, C. K. Folland, and N. Rayner, 1997: A new gridded radiosonde temperature data base and recent temperature trends. *Geophys. Res. Lett.,* **24,** 1499–1502.

Ramaswamy, V., and Coauthors, 2001: Stratospheric temperature trends: Observations and model simulations. *Rev. Geophys.,* **39,** 71–122.

Santer, B. D., and Coauthors, 1996: A search for human influence on the thermal structure of the atmosphere. *Nature,* **382,** 39–46.

——, and Coauthors, 2005: Amplification of surface temperature trends and variability in the tropical atmosphere. *Science,* **309,** 1551–1556, doi:10.1126/science.1114867.

Seidel, D. J., and Coauthors, 2004: Uncertainty in signals of large-

scale climate variations in radiosonde and satellite upper-air temperature datasets. *J. Climate,* **17,** 2225–2240.

Sherwood, S. C., J. R. Lanzante, and C. L. Meyer, 2005: Radiosonde daytime biases and late 20th century warming. *Science,* **309,** 1556–1559, doi:10.1126/science.1115640.

Shine, K., and Coauthors, 2003: A comparison of model-simulated trends in stratospheric temperatures. *Quart. J. Roy. Meteor. Soc.,* **129,** 1565–1588.

Tett, S., J. Mitchell, D. Parker, and M. Allen, 1996: Human influences on the atmospheric vertical temperature structure: Detection and observations. *Science,* **274,** 1170–1173.

Thompson, D. W. J., and S. Solomon, 2005: Recent stratospheric climate trends: Global structure and tropospheric linkages. *J. Climate,* **18,** 4785–4795.

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.

World Meteorological Organization, 2003: Scientific assessment of ozone depletion: 2002. Global Ozone Research and Monitoring Project Rep. 47, World Meteorological Organization, Geneva, Switzerland, 410 pp.