

## Discontinuity Issues with Radiosonde and Satellite Temperatures in the Australian Region 1979–2006

JOHN R. CHRISTY AND WILLIAM B. NORRIS

*Earth System Science Center, The University of Alabama in Huntsville, Huntsville, Alabama*

(Manuscript received 13 February 2008, in final form 26 August 2008)

### ABSTRACT

The temperature records of 28 Australian radiosonde stations were compared with the bulk-layer temperatures of three satellite products of The University of Alabama in Huntsville (UAH) and Remote Sensing Systems (RSS) for the period 1979–2006. The purpose was to use the satellite data as “reference truth” to quantify the effect of changes in station equipment, software, and operations on the reported upper air temperatures and resulting trends. The products are lower troposphere (LT), midtroposphere (MT), and lower stratosphere (LS).

Four periods of significant shifts in temperatures were found in the radiosondes relative to both satellite datasets. In the first two shifts—around 1982/83 and 1987/88—the radiosondes experienced an accumulated LT and MT warming shift of 0.5 K on average. These shifts coincided with equipment changes. If unadjusted for these shifts, the radiosondes report spurious tropospheric warming of almost 0.2 K decade<sup>-1</sup>. For LS in the first period, there is relative warming but in the second, cooling. If unadjusted, the radiosondes overstate LS cooling by about -0.15 K decade<sup>-1</sup>.

The third (early 1990s) and fourth (1998 LT and MT and 2002 LS) shifts are less robustly connected to changes in the radiosondes. Errors in the construction methodology of the satellite products likely account for at least part of the discrepancies but cannot be attributed with confidence to a specific cause. Having opposite signs in the two periods, the last two discrepancies tend to cancel each other. The net effect of these last two shifts on the overall LT and MT trends of  $\pm 0.03$  K decade<sup>-1</sup> is small.

### 1. Introduction

The present study, an analysis of 28 yr of temperature records from Australian radiosondes (sondes), is the second phase of an effort to identify and remove discontinuities in the records of the worldwide sonde network and, when possible, to determine the causes of the discontinuities. The first, described in Christy and Norris (2006, hereafter CN06), chose a family of sonde stations in the United States for analysis. Partitioning the sondes by national region before analysis tends to simplify the inhomogeneity analysis, since the sondes in a region are typically controlled by a single political entity and, therefore, more likely to have experienced relatively uniform changes in hardware, software, or operational procedures. A subset of the U.S. sondes were chosen in

CN06 for these reasons as well as for their high degree of record completeness.

The Australian network was chosen for this study not only for the first reason and for its high degree of completeness of record but also because of a discovery made by Christy and Norris (2004) in an analysis of 89 sondes in the Southern Hemisphere (SH). We found in that study that the shifts at the discontinuities in the Australian sondes were mostly positive, whereas those of the rest of the SH tended to be negative. The Australian network is important because its sondes tend to dominate Southern Hemisphere sonde compilations. This study takes a more thorough look at the records from this group.

The impetus for these studies is a desire to determine long-term temperature trends in the atmosphere. Sondes are a valuable tool for such work. By documenting more than 50 yr of temperature changes in the atmosphere from near the surface to above the tropopause, sondes contribute to an understanding of the sensitivity of the climate system to different forcings. Through sonde

---

*Corresponding author address:* John R. Christy, ESSC, The University of Alabama in Huntsville, Cramer Hall, Huntsville, AL 35899.  
E-mail: christy@nsstc.uah.edu

observations, the high correlation of temperature with energy content provides the means to examine the accumulation or depletion of energy in the bulk atmosphere. Significantly, the sonde record includes the period starting in the mid-1970s when surface temperatures began to rise at a rate that many climate models have not been able to simulate without including the modeled influence of increasing concentrations of man-made greenhouses gases.

Although valuable, the role of sondes for such climate studies is limited. Ground stations are not optimally located to capture all the spatial degrees of freedom in the global temperature field. Also, discontinuities occur because of changes in instrumentation, ground station hardware or software, or operational procedures. These nonclimatic changes often introduce shifts in temperatures that cause calculated trends to differ from the true trend by amounts greater than or equal to the magnitude of the true trend.

Several research groups have developed strategies to detect and remove these inhomogeneities. Lanzante et al. (2003) selected 87 stations with reasonably complete data and fairly representative spatial distribution over the globe. Each station was examined individually through 1997. Using statistical analyses and human judgment, they determined adjustments to reduce discontinuities and produce the Lanzante, Klein, and Seidel (LKS) dataset. Among the factors they considered were vertical coherence, day–night differences, sudden shifts, and reports of instrument changes. Free et al. (2005) have appended the post-1997 data—adjusted through objective criteria—to 85 of the LKS stations to produce the Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC) datasets.

Christy and Norris (2004) compared Southern Hemisphere sonde time series with those derived from Microwave Sounding Unit (MSU) satellite data (version 5.1) compiled by the University of Alabama in Huntsville (UAH). The sonde and satellite data for the troposphere were directly compared station by station to determine the effects of known major changes in instrumentation. For Australia they found that the most prominent change was caused by the switch from Philips Mark III to Vaisala sondes, which led to spuriously warmer temperatures in the troposphere. Other SH instrument changes, for example, from either Vaisala RS-21 or VIZ-B to Vaisala RS-80, caused shifts toward temperatures cooler than those of the satellite values. The cumulative effect of the adjustments on the tropospheric trend of the 89-station composite time series was small ( $\pm 0.02$  K decade<sup>-1</sup>). This result was implied in earlier comparisons (Parker et al. 1997).

Thorne et al. (2005) amassed data from about 650 stations [the Hadley Centre Atmospheric Temperature (HadAT2) dataset]. A common mean period was established for the stations from which anomalies were generated. They identified discontinuities in each station (the target) by comparing its anomalies with those of a reference series. The reference series was obtained from the nontarget stations whose anomalies had high temporal coherence with those of the target. Statistical tests were applied to discover discontinuities in the target. Metadata were also consulted to build confidence in the statistically identified breakpoints.

Haimberger (2007) used the first-guess output of a global reanalysis or forecast model to supply the reference time series. A statistical analysis determined whether shifts or changes in trends appeared over time and if so, the sonde data were adjusted accordingly. (The documented breakpoints from this procedure will be noted later.)

Sometimes, systematic problems with sondes have been discovered that cannot always be eliminated by detecting and removing discontinuities. Parker et al. (1997) demonstrated that trends of sonde temperatures were clearly too negative in the stratosphere. Sherwood et al. (2005) found that trends of temperatures from daytime releases tended to be significantly more negative than those from nighttime releases, especially in the stratosphere. However, Christy and Spencer (2005) responded by showing that other fundamental issues, such as the documented positive shifts in Australian sondes, must be taken into account before conclusions about tropospheric trends can be made. Randel and Wu (2006) compared microwave satellite temperatures with a limited set of sondes and concluded that time series of composites of sondes were likely to be too negative, especially at the highest altitudes.

The approach used in this study is similar to that of CN06 in which the analysis was restricted to a family of 31 U.S. sondes that had largely experienced common changes. The stations were composited and compared with both the UAH MSU data and the MSU data of Remote Sensing Systems (RSS). When a discontinuity was identified by both sets of satellite data, more confidence could be placed in the genuineness of the result. When a discontinuity was identified by one set of satellite data but not by the other, suspicions were raised that unresolved errors in the satellite datasets existed. CN06 concluded that for the U.S. family of sondes studied, the tropospheric shifts tended to balance each other and have little effect on the 26-yr (1979–2004) trend.

Certain apparent breaks in the sonde records may have been due to spurious shifts in both satellite datasets. This was found in Christy et al. (2007, hereafter

C07), through performing direct comparisons between all available sondes in the tropics and both UAH and RSS data for the lower troposphere. CN06 and C07 found a warming shift in the satellite datasets around 1993 that pervasive radiosonde changes, either in the U.S. VIZ network or in the tropical stations as a whole, could not explain. The shift was more prominent in RSS data but was evident in UAH data as well. Because this was a time when the MSU on the National Oceanic and Atmospheric Administration (NOAA) satellite *NOAA-12* needed special attention and the diurnal drift corrections on *NOAA-11* were becoming quite large, the source of the sonde versus satellite discrepancy could be inadequate satellite corrections.

For the present study, only stations with sufficient data to form meaningful time series were selected from the Australian network. Each station released sondes in the daytime (0000 UTC). Some also made nighttime releases (1200 UTC) with enough data to form useful time series.

## 2. Data

The temperature soundings for the Australian stations were accessed from the quality-checked Integrated Global Radiosonde Archive of the National Climatic Data Center (Durre et al. 2006). As in CN06 and C07, the daily radiosonde profiles were converted to weighted-average temperatures of the three atmospheric layers defined below using full radiation calculations. Only those stations with at least 180 months of usable monthly averages were retained. This criterion eliminated all but 28 daytime stations and 5 nighttime stations.

The satellite data used for comparison are the MSU brightness temperatures from UAH (CN06) and RSS (Mears et al. 2003; Mears and Wentz 2005). These products are best thought of as bulk-layer temperatures. The three products and their approximate vertical extents are lower troposphere (LT; surface to 300 hPa), midtroposphere to lower stratosphere (MT; surface to 70 hPa), and lower stratosphere (LS; 120–20 hPa).

UAH and RSS merge satellite data differently. The effect can be seen in the overall global trends of the three layers for 1979–2006. The trends ( $\text{K decade}^{-1}$ ) differ as follows: LT by 0.05, MT by 0.08, and LS by 0.12. In each case the RSS trends are warmer (see the U.S. Climate Change Science Program, Synthesis and Assessment Product 1.1 report for further details). The versions used here are UAH 5.2 for LT, UAH 5.1 for MT and LS, and RSS 3.0 for all three layers. The recent correction in RSS LT version 3.1 was applied to post-2006 data and not relevant to this comparison study.

For a further source of information to check the discovered differences between the sondes and both UAH and RSS MT in the latter part of the time series, we have also included the data from Zou et al. (2006, hereafter Z06), which only covers the period 1987–2005. The main differences between the Z06 merging methodology and UAH and RSS are 1) Z06 lacks adjustments for spacecraft drift through the diurnal cycle, and 2) Z06's intersatellite bias calculation procedure relies exclusively on simultaneous nadir overpasses (SNO) in the north polar region rather than all latitudes. Both differences likely lead to the result shown later that Z06 has considerably greater error relative to the Australian sondes than either UAH or RSS. Evidence is clear that (i) spacecraft drift introduces errors to trends and (ii) there are latitudinal dependencies in intersatellite biases.

## 3. Method

As a starting point, the sonde and satellite records were processed into series of monthly averages. All subsequent comparisons were made with the series having this averaging time. The time series of each sonde station was matched with the satellite data corresponding to the  $2.5^\circ$  grid cell in which the station resides. An important series for analyzing a sonde is the difference between its series and that of its satellite counterpart. Forming the difference assumes, at least tentatively, that the satellite series represents ground truth. In this study both UAH and RSS satellite series are used as references for the three main products and Z06 for part of the MT time series.

Discontinuities in a sonde series were detected by considering its difference series relative to one of the satellites. Shifts were identified by computing adjusted 36-month means on either side of a target month in the series and subtracting the earlier mean from the later. Before the subtraction, adjustments were made to correct for the autocorrelation of the data in each 36-month period. A mean was divided by  $\sigma/\sqrt{N}$ , where  $N$  ( $\leq 36$ ) is the number of data points in the averaging period, and  $\sigma$  is the standard deviation. Typically, the autocorrelation adjustment reduced the number of degrees of freedom in the mean from 36 to about 15. A test statistic or the magnitude of the shift computed in this manner is the  $z$  score for the difference of means.

For each difference series, this approach produced a  $z$  score for each month based on the latter minus the earlier 36-month means (the length was shortened to 24 months at the beginning and end of the series). Because noise exists in the reference series, shifts within the noise level are assumed to be undetectable. Therefore, a

breakpoint was considered to exist in the sonde series wherever the magnitude of a  $z$  score exceeded a selected threshold. The  $z$  scores selected for this study were 2.5, 3.5, and 4.5. CN06 provides more details.

A separate time series of sonde data was prepared relative to each satellite dataset. Once significant shifts were identified relative to a satellite dataset, the magnitude of the discovered shift was removed from the respective sonde series then further comparisons were performed. Because of possible erroneous shifts in the satellite series, we recognize shifts may be spuriously attributed to the sonde series with this method. Therefore, it is more accurate to say that the sonde series have been adjusted to agree with the respective satellite time series at the breakpoints rather than corrected.

To test for the robustness of the method, we constructed 100 synthetic time series of noisy data to which specified  $z$  score shifts of  $\Delta = 5, 3,$  and  $0$  (control) were inserted in one place for each test. Using the three detection thresholds ( $z = 2.5, 3.5,$  and  $4.5$ ), we detected the  $\Delta = 5$  event of 100%, 91%, and 63%, respectively, with an root-mean-square (RMS) error of 6.5 months. Again with  $\Delta = 3$ , we detected the event using our three thresholds of 86%, 46%, and 19%, respectively, of the cases with an RMS error of 9 months. Finally, for the control case ( $\Delta = 0$ ), we detected “significant” shifts of either sign at the three thresholds of 83%, 14%, and 2%, respectively, that is, a high false detection rate for threshold  $z = 2.5$  but low for 3.5 and 4.5. For a single station, therefore, the detection threshold of 3.5 and 4.5 produce the most credible “hits” with fewest false positives.

This test illustrates the importance of the composite results. When 10 of the synthetic time series are composited, the result is consistent from test to test, always identifying the breakpoint within two months of its insertion. In our project here, we have 28 stations, so that the composited time series in effect virtually eliminates the random false positive “hits,” which is very important for the trend analysis. In summary, thresholds of  $z = 3.5$  and  $4.5$  are most useful for identifying breakpoints for individual stations, while all three thresholds are informative for the composite analysis. And, the influence of the individual false positives, because of their randomness and low magnitude, is essentially washed out in the composite results.

In CN06, the main comparisons between the sonde and satellite datasets were performed using the composite anomalies of the 31 U.S. VIZ stations. This was possible because the institutional changes to the sondes were mostly simultaneous. This approach cannot be applied to the Australian sondes because of the significant differences in the times at which instruments or

software were changed at the stations. Also, there were basic software differences among the stations. Under these circumstances the breakpoint methodology of CN06 was applied to each station individually before composite series were formed and analyses of breakpoint magnitudes were performed.

Another type of series used in the analysis, constructed to assist in understanding the effect of the shifts on the composite Australian time series, was the sum of accumulated shifts. For a given month, all station shifts exceeding a selected threshold were averaged. This average was added to the accumulated total from the previous month.

Among the metrics used to assess the similarity of series are the median of the individual station correlation of anomalies, correlation of composite anomalies, RMS differences in trends, and difference in trend of the sonde and satellite composites.

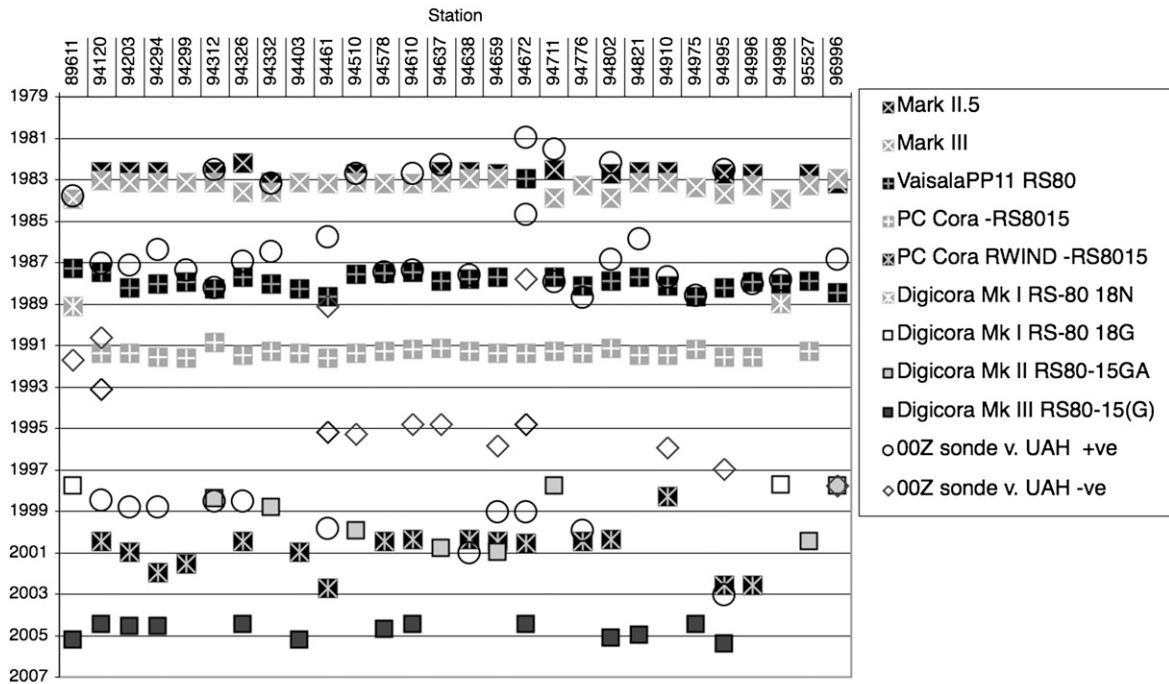
#### 4. Results

As an example of the methodology employed, Fig. 1 shows the breakpoints discovered for each sonde relative to UAH (Fig. 1a) and RSS (Fig. 1b) for the layer MT when  $z = 3.5$ . As will be shown, UAH comparisons with the individual stations tend to have higher correlations and lower RMS error than RSS. Because of the lower noise, UAH tends to identify more breakpoints for a given  $z$  threshold: on average, a  $z$  threshold for UAH detects the same number of breaks as a threshold 0.5 smaller for RSS for LT and MT.

The additional noise in RSS relative to UAH probably relates to the construction methodology of the gridded maps. There are gaps between the 14 daily north–south-oriented satellite swaths in the raw data. RSS treats the gaps as missing data for the day, whereas UAH interpolates between the swaths to fill the gaps day by day as a result of the high spatial coherency of the deep-layer temperature field. Thus, when calculating the monthly mean, UAH essentially has temporally complete daily data for each grid cell, whereas RSS averages only those days when direct observations were available.

We note here that an early concern of the UAH and RSS datasets was the different method each used in merging multiple satellite time series together in the period of *NOAA-9* around 1985. This is of small influence here. During operation of *NOAA-9*, the differences between UAH and RSS shifted by a few hundredths in one direction then back within 2 yr. The small magnitude of the shifts and their length relative to the averaging period for breakpoint determination (36 months) introduced no significant effect. Notice, for example, in the LT accumulated adjustments the differences between

UAH 00UTC MT Events Z > 3.5



RSS Events Z > 3.5

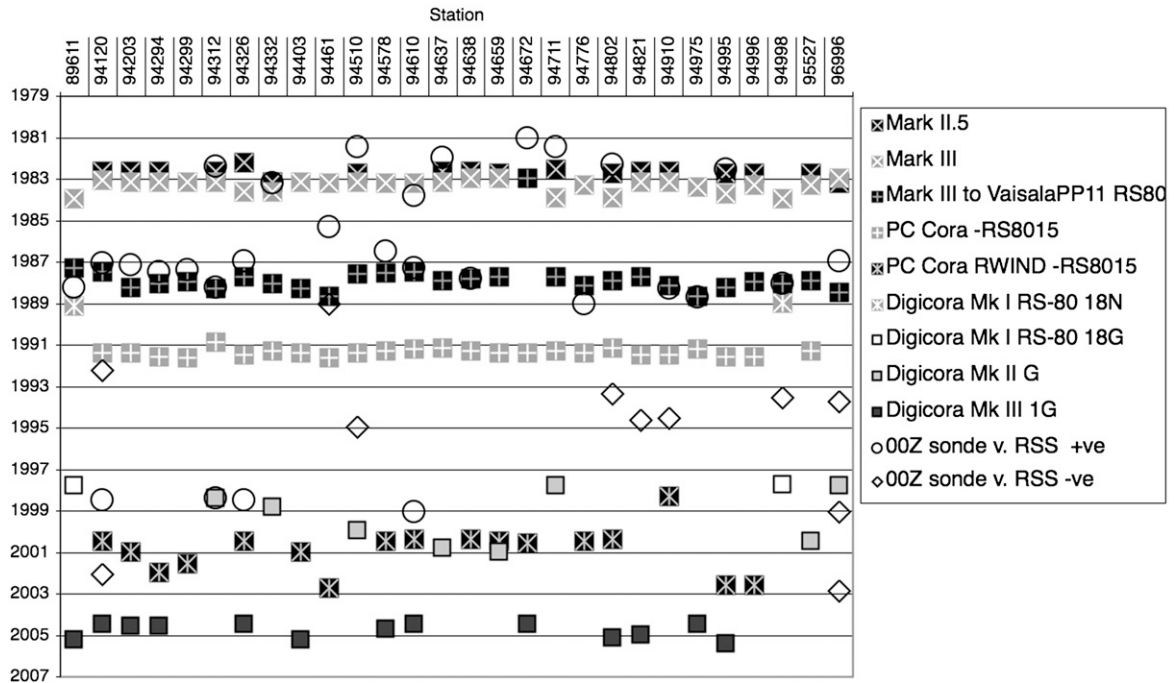


FIG. 1. Key dates in the analysis of discontinuities in the temperature records of 28 Australian radiosonde stations vs (a) UAH and (b) RSS satellite data. Stations are designated by World Meteorological Organization (WMO) identifier. Square symbols indicate the occurrence of a change in sonde instrumentation or software. Circles indicate the detection of a statistically significant positive shift between the sonde data for layer MT and the UAH satellite data at a z-score threshold of  $z = 3.5$ . Diagonal squares indicate a significant negative shift. [Source for (b) is Michael Joyce, Australian Bureau of Meteorology.]



## Australian Radiosonde Anomalies (RLT) with differences vs. UAH and RSS LT

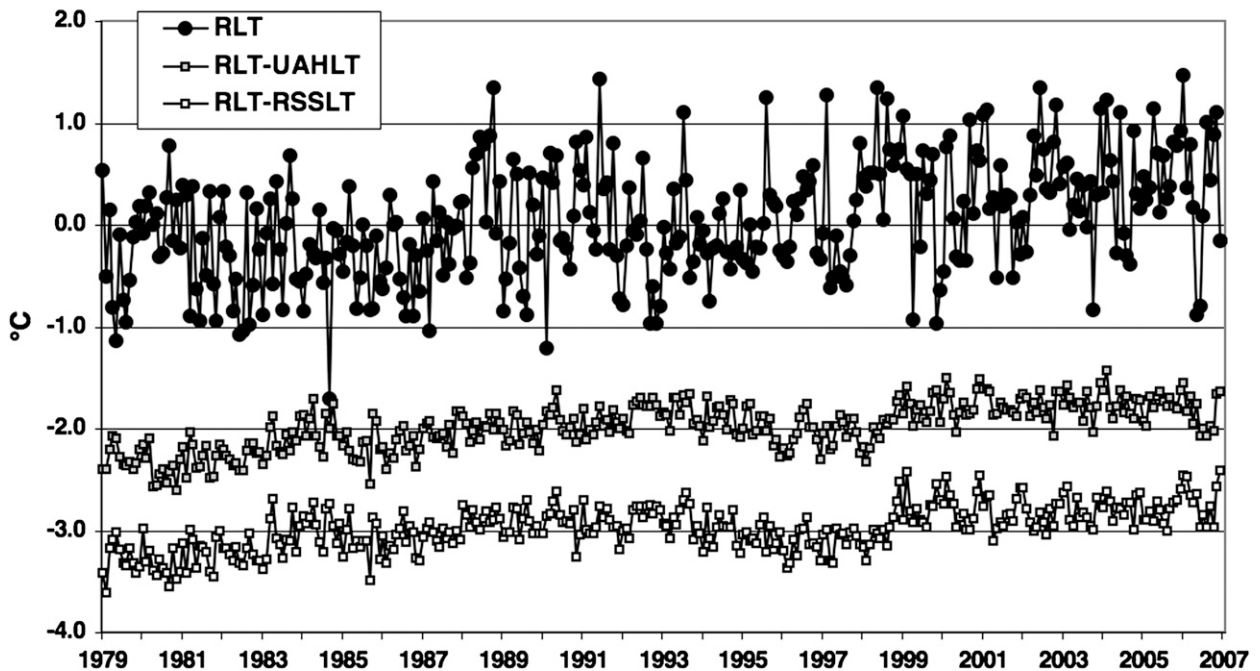


FIG. 2. Time series of the composite monthly averaged LT anomalies (K) and LT anomaly differences (K) based on the 28 daytime Australian sonde stations at 0000 UTC. (top) Series are derived from the unadjusted station series. Series are the differences of the top series and the corresponding (middle) UAH and (bottom) RSS satellite LT products (sonde minus satellite) and have been shifted by  $-2.0$  and  $-3.0$  K, respectively, for clarity.

UAH and RSS do not appear until 1991, well after the NOAA-9 period.

Despite these differences, the clear pattern in both figures is the tendency for shifts toward warmer sonde temperatures (the earlier two sets of shifts) to coincide with sonde changes from Mark II/II.5 to Mark III (from around 1982 to 83) and from Mark III to Vaisala RS80 (around 1987–89). This has been noted elsewhere (Christy and Norris 2004; Christy and Spencer 2005).

The shifts toward cooler sonde temperatures identified around 1993–95 are not associated with documented changes in the sonde equipment. These shifts may be the result of spurious warming in the satellite data. The last main area of breakpoints—around 1999—is only loosely associated with sonde changes. Each atmospheric layer will now be examined.

### *a. Lower troposphere*

The uppermost time series in Fig. 2 is the composite of LT anomalies for the 28 Australian stations, unadjusted for discontinuities. The next (last) series in the figure is the sonde minus UAH (RSS) satellite differ-

ences. Figure 3 displays the cumulative sum of the station shifts exceeding  $z = 2.5$  to show how the sondes and satellites differ in the composite. As would be expected, the character of the accumulated breakpoint adjustments mimics that of the difference time series in Fig. 2. Table 1 provides statistical information on the direct comparisons of the individual pairs of time series and the comparisons of the station composites for the LT product.

The first two columns in Table 1 indicate 1) the median of the variance explained by the individual comparisons—that is, a grid-level statistic—and 2) the variance explained by the 28-station composite (i.e., a multigrad statistic, which is informative for large-scale averages.) The resulting improvement in the second column indicates a level of independence among the stations, with random errors being reduced in the mean.

The Comp Trend column shows that UAH and RSS differ in their trends of the 28-station daytime composite by less than  $0.04 \text{ K decade}^{-1}$ , which is within the error associated with each dataset. (For comparison CN06, which is based on the 31-station U.S. VIZ network, estimates the 95% error in the 1979–2006 trend to

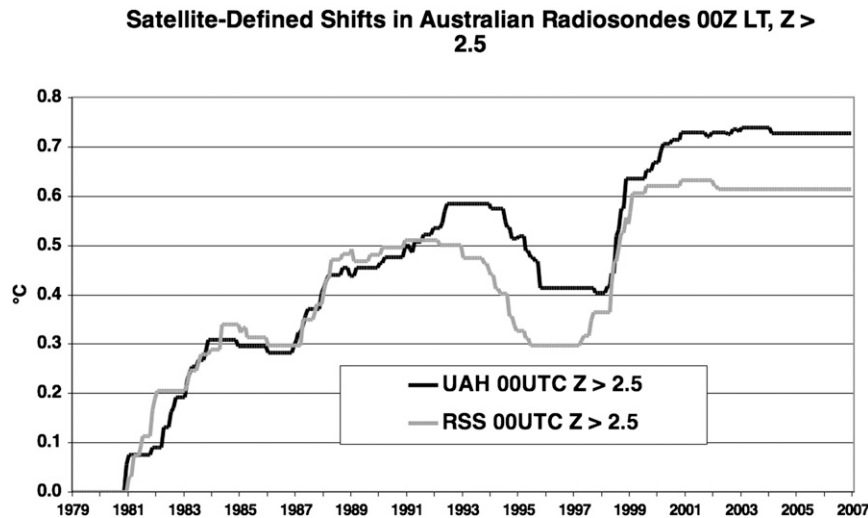


FIG. 3. Time series of the accumulated values (K) of the significant LT 0000 UTC shifts detected by comparison against UAH (black) and RSS (gray) for a  $z$ -score threshold of 2.5. The character of the time series mimics the features in the difference time series of Fig. 2.

be  $\pm 0.06 \text{ K decade}^{-1}$ .) However, the difference between the unadjusted sonde data and both satellite results is substantial.

According to the Pos and Neg columns in Table 1, both UAH and RSS “detect” many more positive than negative breakpoints in the daylight sondes. Table 1 (Comp Diff Trend column) also shows improvement in the comparisons, as the sonde data are adjusted for greater numbers of breaks as the  $z$ -score threshold is lowered. Obviously, adjusting for the breaks forces the sondes to agree more closely with the satellites. The original sonde trend of  $+0.29 \text{ K decade}^{-1}$  declines to  $+0.05$  (UAH) and to  $+0.11$  (RSS) when the threshold is lowered to  $z = 2.5$ . Both values are less positive than the respective satellite trends. The RMS of the individual 28-station trend differences indicates that as the threshold is tightened, the satellites and sondes agree to about  $0.1\text{--}0.2 \text{ K decade}^{-1}$  at the individual station level while the composite trend shows agreement to within a few hundredths  $\text{K decade}^{-1}$ . (We use RMS here as a more descriptive metric of the comparison of individual trend differences as it includes the mean bias, whereas the standard deviation does not.)

According to Fig. 3, the accumulated adjustments as a result of the comparison of the sondes with UAH and RSS are almost identical from the start to 1992. After 1992, UAH and RSS diverge as UAH continues a small rise, whereas RSS begins a downward drift. By 1994, both RSS and UAH begin a parallel sequence of adjustments. [This difference in the 1992/93 period is the main source driving the trend difference between UAH and RSS, as it occurs near the center of the time series

to have its greatest influence on the linear trend (C07).] In 1998, a sudden increase in the accumulated error time series indicates either a spurious warming in the sonde data or a spurious cooling in the satellite data of up to  $0.3 \text{ K}$ .

Results for the five nighttime stations are also shown in Table 1. The much smaller sample is associated with higher noise levels. At these few stations, the composite UAH trend is now slightly more positive than RSS ( $0.07$  versus  $0.05 \text{ K decade}^{-1}$ ). The sondes and satellites agree best when  $z = 3.5$  is the breakpoint threshold rather than  $z = 2.5$ , as might be expected from the experimental results with synthetic time series discussed earlier—that is, a small sample allows random errors to be influential. In particular, UAH adjustments to station 89611 (Casey at  $-66$  latitude), which reported data from 1989 onward only, caused the trend to change from  $-0.12$  to  $+0.50 \text{ K decade}^{-1}$ . The change is likely an overcorrection and is the main reason the UAH-adjusted sonde composite in Table 1 is  $+0.13 \text{ K decade}^{-1}$  warmer than the actual UAH composite.

#### *b. Midtroposphere*

In general, the MT product is more stable than the LT product because, unlike LT, its construction does not depend on small differences of different satellite scan angles (Christy et al. 2003). Typically, the standard errors are twice as small for MT than LT. The results for MT are given in Figs. 4 and 5 (includes Z06) and Table 2. An abbreviated version of Table 2, which includes the shorter period of 1987–2005 and all three satellite datasets for 0000 UTC, is presented in Table 3.

TABLE 1. Comparison of adjusted and unadjusted Australian sonde LT temperatures with UAH and RSS satellite temperatures for 1979–2006. Column heads are defined as follows:  $z = m.n$ : the  $z$ -score threshold of shifts in sonde anomaly series (relative to satellite anomaly series) such that shifts larger than this value are removed from the sonde series. Med  $r^2$ : the median of the squared correlations of 28 pairs of sonde anomalies with the corresponding satellite anomalies. Comp  $r^2$ : the squared correlation of station-composited sonde anomalies with the corresponding composite of satellite anomalies. Std dev diff: the standard deviation of a difference series (the station-composited sonde anomalies minus the corresponding composite of satellite anomalies). Comp trend: trend of station-composited anomalies (either sonde or satellite). Comp diff trend: the trend of a difference series (the station-composited sonde anomalies minus the corresponding composite of satellite anomalies). RMS trend: RMS of the trend differences of station anomalies (either sonde or satellite). Pos (Neg): the number of positive (negative) shifts detected by comparison with satellite data.

Daytime (0000 UTC)	Med $r^2$	Comp $r^2$	Std dev diff	Comp trend	Comp diff trend	RMS trend	Pos	Neg
Sonde unadjusted				0.289				
UAH	0.734	0.837	0.231	0.091	0.198	0.253		
RSS	0.687	0.854	0.220	0.130	0.161	0.211		
Sonde adjusted								
UAH: $z = 4.5$	0.743	0.897	0.174		0.066	0.195	21	1
RSS: $z = 4.5$	0.687	0.874	0.200		0.107	0.197	9	3
UAH: $z = 3.5$	0.762	0.918	0.156		0.027	0.148	39	11
RSS: $z = 3.5$	0.719	0.899	0.177		0.072	0.168	23	11
UAH: $z = 2.5$	0.783	0.920	0.153		-0.045	0.119	71	27
RSS: $z = 2.5$	0.728	0.914	0.162		-0.019	0.143	55	23
Nighttime (1200 UTC)	Med $r^2$	Comp $r^2$	Std dev diff	Comp trend	Comp diff trend	RMS trend	Pos	Neg
Sonde unadjusted				0.139				
UAH	0.785	0.792	0.304	0.071	0.068	0.205		
RSS	0.792	0.806	0.299	0.049	0.09	0.107		
Sonde adjusted								
UAH: $z = 4.5$	0.785	0.806	0.294		0.09	0.274	1	1
RSS: $z = 4.5$	0.792	0.806	0.299		0.09	0.107	0	0
UAH: $z = 3.5$	0.799	0.837	0.266		-0.006	0.2	4	2
RSS: $z = 3.5$	0.789	0.819	0.287		-0.03	0.142	3	0
UAH: $z = 2.5$	0.799	0.814	0.296		0.131	0.187	5	8
RSS: $z = 2.5$	0.787	0.826	0.284		0.039	0.17	5	3

The pattern of results is similar to that of LT: sondes before 1990 experience relative warm shifts; in the mid 1990s, a downward shift; and in 1998, an upward shift but with less amplitude than LT. In Table 2 the higher correlations and lower RMS values relative to those in Table 1 demonstrate the lower noise of the MT comparisons. The unadjusted sonde trend of  $+0.19 \text{ K decade}^{-1}$  is reduced to  $+0.06$  and  $+0.10 \text{ K decade}^{-1}$  (UAH and RSS, respectively) when the threshold is reduced to  $z = 2.5$ . The discrepancy indicates relative shifts accumulating to about 0.5 K in the years prior to 1990, being similar to the LT result. If the post-1990 shifts are a result of satellite problems, the net error on the satellite trends (UAH and RSS only) would imply an introduction of spurious warming of  $+0.03$  to  $+0.04 \text{ K decade}^{-1}$  as a result of the situation where the assumed spurious warming in the mid-1990s is larger than the assumed spurious cooling in 1998–99. Thus, the shape of the MT discrepancies in the 1990s is slightly different from that of the LT product.

For the early 1990s discrepancy, the information contained in Z06 indicates a somewhat similar shift ( $\sim 0.15 \text{ K}$ )

to that determined by UAH and RSS. The last (1999–2002) discrepancy though, as determined by Z06 comparisons, implies a significant spurious warming ( $\sim 0.3 \text{ K}$ ) of the sondes versus that of  $\sim 0.1 \text{ K}$  for UAH and RSS. We believe the evidence shows that the true discrepancy is the smaller amount and the larger magnitude is likely due in Z06 to a lack of diurnal drift and latitudinal bias adjustments for *NOAA-14*. Note that both UAH and RSS apply these adjustments and that they also use the coincident observations of *NOAA-15*, which experienced virtually no drift during this period. The statistical comparisons in Table 3 indicate Z06 to be the least consistent product in comparison with the sondes.

The overall magnitude of the MT trend is less positive than LT as a result of the stratospheric component of MT, where broad cooling is occurring. The smaller MT relative shifts after 1993 could result from the smaller diurnal corrections for MT, if the satellites are the source of the discrepancies. Also, because the individual station correlations with satellite data are higher for UAH than for RSS (and Z06), UAH detects more breakpoints and thus produces larger accumulated shifts.



**Australian Radiosonde Anomalies (RMT)  
with differences vs. UAH, RSS and ZOU MT**

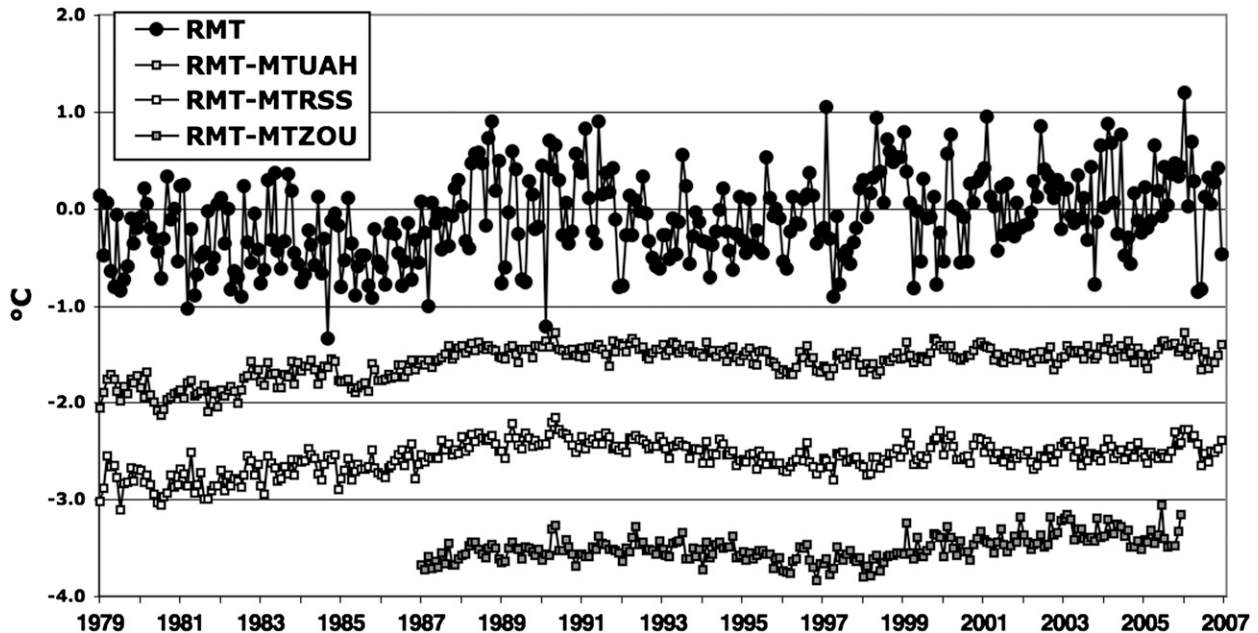


FIG. 4. Same as in Fig. 2, but for MT and including Z06 data.

The composite MT trend for the unadjusted 5 nighttime sondes is similar to that of the 28 daytime sondes ( $+0.22 \text{ K decade}^{-1}$ ). The resulting adjustments are a bit more stable than those of LT and generally support trends of around  $+0.06$  and  $+0.11 \text{ K decade}^{-1}$  (UAH and RSS, respectively), being virtually the same as the 28-daytime sondes. These trends strongly point to a spurious warming since 1979 in the unadjusted tropospheric sonde time series for both daytime and nighttime releases.

### c. Lower stratosphere

The time series of anomalies and accumulated breakpoints for LS are displayed in Figs. 6 and 7 with statistical results in Table 4. There are four main sections of detected change in LS, just as in LT and MT. However, the sign of the second one is now reversed, and the timing of the fourth is a few years later. Table 4 shows that the unadjusted sonde data have a significant negative trend of  $-0.68 \text{ K decade}^{-1}$ . This compares with trends of  $-0.46$  (UAH) and  $-0.39$  (RSS)  $\text{K decade}^{-1}$  for these 28 stations. The cause of such a large discrepancy is seen in Fig. 7, where the accumulated differences reveal a large upward shift ( $\sim 0.4 \text{ K}$ ) around 1982 followed by a very large downward shift ( $\sim 0.6 \text{ K}$ ) around 1987. This positive “hump” in the differences early in the time series translates into a significant negative trend in the sondes rela-

tive to the satellites. The more minor shifts in the mid-1990s and early 2000s add to the magnitude. In LT and MT, the accumulation of shifts based on UAH data exceeded in magnitude those based on RSS data. Here, the reverse is true despite the lower noise in the sonde-versus-UAH comparisons.

In Table 4 the differences in the unadjusted daytime composite sonde trend and the satellite data ( $>0.2 \text{ K decade}^{-1}$ ) are reduced by breakpoint adjustments to less than  $0.07 \text{ K decade}^{-1}$  at  $z = 3.5$  and less than  $0.02 \text{ K decade}^{-1}$  at  $z = 2.5$ . The reductions are associated with significant shifts of more than  $1.0 \text{ K}$  detected by the satellites at several stations.

The trend of the nighttime composite sondes is a surprising  $-0.112 \text{ K decade}^{-1}$ —substantially more positive than the daytime composite of  $-0.667 \text{ K decade}^{-1}$ . Two of the five stations—94975 (Hobart) and 94998 (Macquarie Island)—had positive trends, which is loosely consistent with a region of near-positive stratospheric trends that has been observed in satellite data. As with MT, the adjustments based on  $z = 3.5$  produce the best agreement between the sondes and each satellite dataset, reducing composite trend differences to less than  $0.02 \text{ K decade}^{-1}$ . The magnitude of the four positive adjustments (one each for four of the five stations) that were detected for  $z = 3.5$  and  $4.5$  was large,

Satellite-Defined Shifts in Australian Radiosondes MT, Z > 2.5

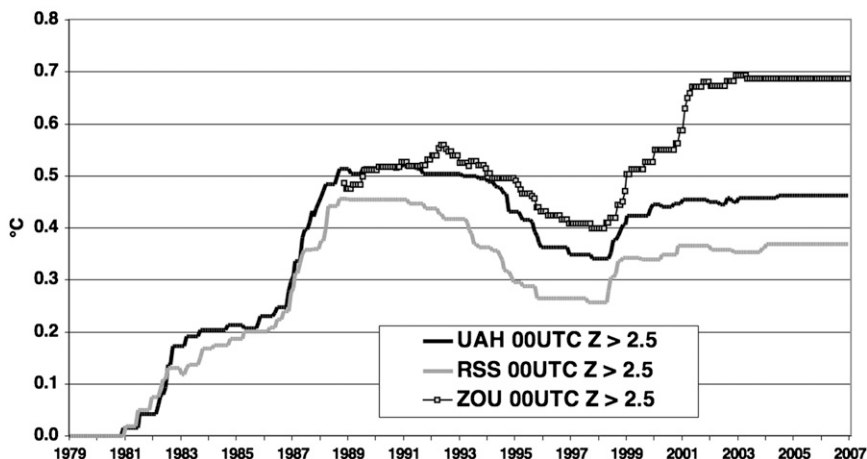


FIG. 5. Same as in Fig. 3, but for MT and including Z06 data.

averaging 1.28 K and coinciding with the change from Philips Mark III to Vaisala RS80 in the 1980s.

5. Discussion

The goal of this section is to attempt to attribute and quantify probable errors in the datasets used in the analysis. The results indicate four basic periods when significant differences between the sondes and the sat-

ellites occur. The causes of the discrepancies of the first two are easily explained, but the causes of the last two are more difficult to attribute with confidence.

a. 1982/83

The first major discrepancy between the two observing systems occurs in 1982/83 at a time when the sondes were being upgraded from Philips Mark II and II.5 to Mark III. The coincidence and significance of these

TABLE 2. Same as in Table 1, but for MT comparisons.

Daytime (0000 UTC)	Med $r^2$	Comp $r^2$	Std dev diff	Comp trend	Comp diff trend	RMS trend	Pos	Neg
Sonde unadjusted				0.192				
UAH	0.797	0.859	0.168	0.060	0.132	0.164		
RSS	0.697	0.870	0.161	0.098	0.097	0.133		
Sonde adjusted								
UAH 4.5	0.837	0.935	0.109		0.040	0.119	23	7
RSS 4.5	0.702	0.889	0.148		0.032	0.101	11	1
UAH 3.5	0.856	0.951	0.094		-0.010	0.085	41	14
RSS 3.5	0.729	0.929	0.113		-0.013	0.118	29	11
UAH 2.5	0.859	0.958	0.087		-0.001	0.074	60	36
RSS 2.5	0.741	0.943	0.101		-0.001	0.092	48	30
Nighttime (1200 UTC)	Med $r^2$	Comp $r^2$	Std dev diff	Comp trend	Comp diff trend	RMS trend	Pos	Neg
Sonde unadjusted				0.218				
UAH	0.726	0.854	0.315	0.018	0.200	0.219		
RSS	0.672	0.697	0.317	0.052	0.165	0.165		
Sonde adjusted								
UAH 4.5	0.801	0.832	0.217		0.012	0.121	4	0
RSS 4.5	0.731	0.711	0.304		0.065	0.165	2	1
UAH 3.5	0.841	0.841	0.212		0.037	0.135	5	2
RSS 3.5	0.734	0.755	0.272		0.008	0.072	4	2
UAH 2.5	0.893	0.828	0.224		0.049	0.042	7	5
RSS 2.5	0.755	0.753	0.2782		0.085	0.132	6	6

TABLE 3. Abbreviated version of Table 2 for 1987–2005 0000 UTC MT to allow for direct comparison with Z06 data.

Daytime (0000 UTC)	Med $r^2$	Comp $r^2$	Std dev diff	Comp trend	Comp diff trend	RMS trend
Sonde unadjusted						
UAH	0.878	0.958	0.087	+0.109	-0.015	0.103
RSS	0.769	0.937	0.107	+0.159	-0.065	0.150
Z06	0.734	0.914	0.124	-0.018	+0.112	0.324
Sonde adjusted						
UAH 3.5	0.891	0.964	0.082		-0.010	0.088
RSS 3.5	0.773	0.943	0.101		-0.041	0.158
Z06 3.5	0.780	0.945	0.153		+0.012	0.139
UAH 2.5	0.891	0.966	0.079		+0.031	0.085
RSS 2.5	0.783	0.956	0.090		+0.001	0.141
Z06 2.5	0.796	0.951	0.148		-0.001	0.127

changes makes it highly likely that the sonde network experienced a spurious shift to warmer temperatures with the introduction of the newer Mark models. A further confirmation is found in CN06 in which a comparison of UAH and RSS data versus 31 VIZ sondes shows no differences in the pre-1985 period. The figures suggest that the network-wide effect of these shifts (K) was about +0.2, +0.3, and +0.4 for the bulk averages of LT, MT, and LS respectively.

A check of breakpoints documented in Radiosonde Observation Correction using Reanalyze (RAOBCORE)

version 1.4 (Haimberger 2007) and HadAT2 (Thorne et al. 2005) indicates that 15 of these 28 stations experienced a shift for 0000 UTC in the 850–300 (850–150)-hPa layer, where the bulk of LT (MT) signals occur on average of +0.10 (+0.07) K for RAOBCORE and +0.02 (+0.05) K for HadAT2 relative to the remainder of the time series; and in the 100–50-hPa layer (where LS peaks) +0.11 K in RAOBCORE and +0.20 K in HadAT2. These are generally less than those of the satellite-determined values. (Note HadAT2 uses a more conservative detection scheme and thus will

### Australian Radiosonde Anomalies (RLS) with differences vs. UAH and RSS LS

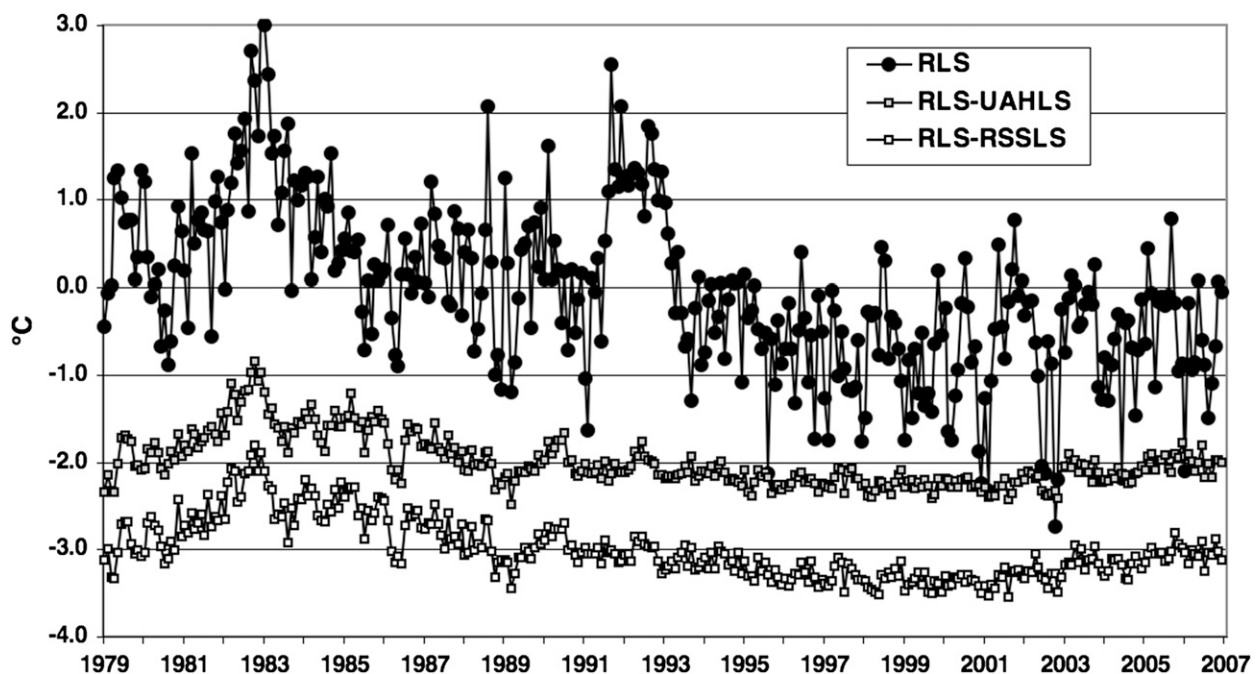


FIG. 6. Same as in Fig. 2, but for LS.

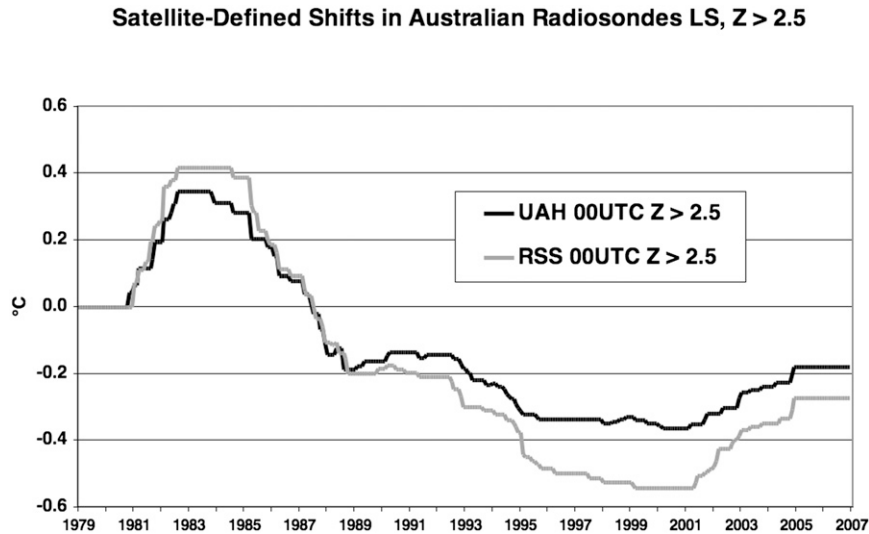


FIG. 7. Same as in Fig. 3, but for LS.

generally have fewer detection events or smaller magnitudes.)

#### b. 1987–89

The second major discrepancy appears in 1987–89 and has a different character from the first. The LT and MT sonde series display warming shifts of about +0.2 and +0.25 K relative to UAH and RSS, respectively. However, in LS the shift is dramatic and of the opposite sign, which for the network as a whole is about  $-0.6$  K. These changes correspond exceptionally well with the replacement of the Philips equipment with Vaisala RS80 (Christy and Norris 2004). Again, RAOBCORE (27 stations) and HatAT2 (23 stations) calculate adjustments for the 850–300-hPa layer to be +0.17 and +0.14 K, respectively; and for 850–150 hPa to be +0.14 and +0.15 K, respectively. RAOBCORE (26 stations) and HadAT2 (25 stations) indicate 100–50-hPa shifts of  $-0.94$  and  $-0.49$  K, respectively.

For the stratosphere this likely represents a significant improvement to the sonde observations through a combination of better instrumentation and the application of correction tables that take into account the heating of the sensor in direct sunlight. However, the same sharp temperature shift was observed in the nighttime LS sonde data, indicating that a more fundamental change in equipment is likely the major factor and also implies that day-versus-night comparisons would miss this large shift. Thus, the first two discrepancies are highly consistent with the hypothesis that the discrepancies between sonde and satellite data may be attributed to changes in the sondes and their associated software.

#### c. 1993–95

In the third period of discrepancies, the sonde temperatures become cooler than those of the satellites for all three products. The magnitudes of the relative declines are about  $-0.2$ ,  $-0.15$ , and  $-0.2$  K for LT, MT, and LS, respectively, although with regard to LS, the shift relative to RSS is close to  $-0.3$  K, whereas UAH is about  $-0.1$  K. Quantifying error in the sonde records becomes more difficult in this period because of the divergence in the UAH and RSS products. However, when the MT breakpoint threshold is  $z = 2.5$ , 22 (20) of the UAH (RSS) comparisons out of 28 indicate a significant shift—all of the same sign—so this is a pervasive event.

For each product, RSS warms relative to UAH. Figures 3, 5, and 7 show that the accumulated breakpoint magnitudes for RSS dip further downward than those of UAH in this period. Evidence has been presented elsewhere that suggests RSS produces a more positive movement in the early 1990s than several other temperature products, including surface data at this point in time, and thus is more likely the source of part of the larger discrepancy (C07; Randall and Herman 2008). However, like RSS, UAH reveals the same tendency relative to the sondes in both Australian and U.S. sondes, though of a lesser magnitude.

This is a period when significant adjustments related to the east–west drifting of the *NOAA-11* satellite are applied, adjustments made differently by UAH and RSS (CN06; C07). *NOAA-12* was also on-station at this time, and its MSU was subject to empirical adjustments applied after launch when anomalous performance was discovered (Mo 1995; Christy et al. 1998).

TABLE 4. Same as in Table 1, but for LS comparisons.

Daytime (0000 UTC)	Med $r^2$	Comp $r^2$	Std dev diff	Comp trend	Comp diff trend	RMS trend	Pos	Neg
Sonde unadjusted				-0.677				
UAH	0.852	0.920	0.296	-0.460	-0.217	0.286		
RSS	0.771	0.887	0.345	-0.391	-0.281	0.347		
Sonde adjusted								
UAH 4.5	0.874	0.953	0.201		-0.063	0.187	8	18
RSS 4.5	0.826	0.937	0.232		-0.114	0.221	6	19
UAH 3.5	0.878	0.958	0.187		-0.052	0.185	16	27
RSS 3.5	0.830	0.947	0.201		-0.068	0.194	18	32
UAH 2.5	0.891	0.962	0.167		-0.014	0.189	37	45
RSS 2.5	0.843	0.958	0.170		0.002	0.136	41	52
Nighttime (1200 UTC)								
Sonde unadjusted				-0.122				
UAH	0.852	0.852	0.497	-0.426	0.304	0.308		
RSS	0.771	0.857	0.507	-0.326	0.203	0.258		
Sonde adjusted								
UAH 4.5	0.912	0.931	0.357		-0.074	0.146	4	1
RSS 4.5	0.832	0.912	0.392		-0.104	0.229	4	2
UAH 3.5	0.908	0.933	0.347		-0.017	0.137	4	3
RSS 3.5	0.830	0.914	0.387		-0.008	0.212	4	4
UAH 2.5	0.903	0.939	0.361		0.151	0.147	6	9
RSS 2.5	0.843	0.914	0.381		0.067	0.186	4	6

No evidence in the sonde record would attribute a shift of this magnitude to changes in sondes or sonde procedures for LT and MT for Australian or VIZ sondes (see below). An upgraded, fast-response coating was applied to the Australian thermistors in 1991, which advanced the response time by 0.2 s, but this would produce a temperature change undetectably small. Also, this change does not coincide with the time of the observed shifts. Another change occurred after 1993 when a new adjustment table caused high-altitude temperatures to decrease slightly (above 100 hPa; M. Joyce 2005, personal communication). The latter is consistent with the changes in LS but its magnitude would have little influence for the MT profile.

A close examination of CN06 reveals that a similar sonde-versus-satellite shift centered in 1993 for all three products using a completely independent set of sondes (VIZ). In that case as well, RSS shows a relative shift to warmer temperatures that is statistically significant. The UAH shifts are also to warmer temperatures but not significantly so. No information exists to suggest that changes in the VIZ sondes could have caused a spurious shift in 1993 (Diamond et al. 2001; CN06).

This result is also consistent with that of C07 and indicates that RSS products experience a relative warming shift around 1993 that is not supported by several other independent data sources. As a check, we examined the differences of two 3-yr periods—1993–95 minus 1990–92—for the 28 Australia stations combined with the 31 VIZ stations of CN06. We are not aware of

equipment or procedural changes that either network underwent in this 6-yr period in the troposphere. Comparison of the composite 3-yr differences for the 59 independent stations indicates that RSS has significantly more relative warming between the two periods than the sondes (LT: 0.09; MT: 0.09 K) or UAH (LT: 0.08; MT: 0.04 K) or Z06 (MT: 0.06 K, Australia only). In summary, the third major tropospheric discrepancy in the satellite-versus-sonde comparison is likely a consequence of spurious warming in the satellite records with more in RSS than UAH, though more research is necessary to provide greater confidence that some event in the sonde record was not responsible. We note that Z06 is broadly consistent with both UAH and RSS here, which suggests a likely satellite warming drift (*NOAA-11* or *NOAA-12*) that has yet to be identified.

We, again, consulted the documented shifts for the various layers. Regarding 850–300 hPa, RAOBCORE (6 stations) and HadAT2 (10 stations) indicate composite shifts of only  $-0.01$  and  $+0.01$  K, respectively. Similarly for 850–150 hPa, RAOBCORE and HadAT2 indicate composite shifts of  $-0.01$  and  $+0.00$  K, respectively. The individual shifts were random in sign and not systematic as implied by the satellite results. These minuscule magnitudes further implicate the satellites as the source of this discrepancy.

In the stratosphere, the change in the Vaisala correction table from RSN86 to RSN93 on 29 Nov 1993, affecting values above 100 hPa, is a possible explanation for at least part of the relative shift shown there. The



main shifts in this period are January 1993 and January 1995 relative to RSS and therefore appear inconsistent with the timing of the table change. The possible influence of the table change at 50 hPa could be 0.2 K (M. Joyce 2005, personal communication). However, a check of ROBCORE indicates a negligible influence of 0.00 K for 100–50 hPa in the period, while in HadAT2 there is a  $-0.11$  K effect (9 stations), consistent in sign with the satellite shift. Thus, at this higher elevation, it appears that the satellites may have a bit of spurious warming drift but that the sondes have a bit of a spurious cooling. These shifts do not appear in the tropospheric temperatures of the sondes.

#### d. 1998–99 and 2001–03

The last notable discrepancy occurs in the troposphere in 1998–99 and in the stratosphere during 2001–2003. As with the third discrepancy, the LS shift is much more significant in RSS than UAH comparisons.

Earlier comments on Fig. 1 indicated possible relationships between the sonde software changes around this period (1999–2003) and this shift in MT, but the synchronization is not as robust as for the earlier equipment changes. The satellite datasets began ingesting data from the new Advanced MSU (AMSU) aboard *NOAA-15* in September 1998 and *NOAA-16* (LS only) in February 2001. At present this appears to be a plausible reason for the relative shifts in the satellite products.

The sonde metadata results for both 850–300 and 850–150 hPa indicate composite shifts of  $+0.00$  and  $-0.01$  for RAOCORE and HadAT2, respectively. These small values support the idea that the fourth detected shift is also related completely to satellite errors. If the third and fourth shifts are satellite problems, the net effect on the 28-yr trend would be to introduce a spuriously negative trend of between  $-0.02$  and  $-0.03$  K decade<sup>-1</sup> for LT, and spuriously positive trends of  $+0.02$  for MT, and  $+0.06$  K decade<sup>-1</sup> for LS, though with LS there is indication the sondes are still partially responsible. Though the evidence is relatively strong that the satellites introduce these errors—especially in the case of the third shift—it is possible that the sonde breakpoint schemes are not able to detect what may be real shifts in the sonde temperatures.

Finally, the fact that the LS comparisons show the fourth shift a few years later is confounding. During this period, RAOCORE indicates only two stations with breakpoints introducing a spurious influence of less than 0.01 K. HadAT2 requires a significant period of time to elapse before officially designating breakpoints, so there is no substantial data at this point to detect this shift. This discrepancy could relate to a problem with the inclusion of *NOAA-16*'s AMSU in 2001 in the LS

satellite data, while no problem occurred in 1998 when *NOAA-15*'s AMSU data joined the LS data stream. This does not, of course, rule out a spurious shift in the radiosonde data.

## 6. Summary and conclusions

Four major discrepancies exist in the Australian sonde records when satellite records are used as reference truth. Regarding the first two discrepancies, the relative warm tropospheric shifts during the Philips Mark II (and II.5) to Mark III and Mark III to Vaisala RS80 indicate a total of about 0.5 K of relative warming in the sondes prior to 1992. The shifts in the records of some stations do not achieve the threshold of  $z = 2.5$  perhaps because of noise in both observing systems. Therefore, it is plausible that the composite shifts detected and shown in Fig. 3 underestimate the true, composite shift that would be experienced if the influence on all stations were detected. A check of the  $z = 2.5$  threshold for MT indicates shifts relative to UAH were detected in 24 of the 28 stations at or near the individual Mark III to Vaisala changeovers. We are confident that the first two discrepancies are related to changes in sonde equipment and software.

The remaining two discrepancies are somewhat ambiguous as to their attribution but evidence is presented that both discrepancies—at least in part—may be assigned to satellite errors. The third discrepancy appears to be a consequence of satellite merging problems at all levels—a conclusion supported by evidence from independent sonde comparisons in the Northern Hemisphere (CN06) and the evidence in Haimberger (2007), Thorne et al. (2005), and Randall and Herman (2008). If the third discrepancy alone were a satellite dataset construction problem, the satellites would contain a spuriously positive trend of about  $+0.1$  K decade<sup>-1</sup> in all three products (a bit more so in RSS).

The last discrepancy may result from inadequacies in the merging of the new AMSU data into the time series for LT and MT in 1998 and LS in 2001. If this last discrepancy alone were the only satellite problem, the satellites would contain spuriously negative trends of about  $-0.13$ ,  $-0.05$ , and  $-0.05$  K decade<sup>-1</sup> for LT, MT, and LS, respectively. However, if both the third and fourth discrepancies were dominated by problems in the construction of satellite datasets, which seems the most plausible conclusion, the net error on all three product trends would be much smaller, as the two shifts are compensating in sign in each layer. The net trend influence would be about  $-0.03$ ,  $+0.02$ , and  $+0.06$  K decade<sup>-1</sup> for LT, MT, and LS, respectively—that is, LT too negative and MT and LS too positive in satellite

data. A synthesis of these results suggests the best estimates for 1979–2006 trends for the 28 Australian stations are +0.12 and +0.07 for LT and MT, respectively, being more negative than the unadjusted sonde data; and  $-0.42 \text{ K decade}^{-1}$  for LS, being more positive.

*Acknowledgments.* This research was supported by NOAA Grant NA06NES4400009. We are grateful for input from Michael Joyce, David Jones, and Blair Trewin of the Australian Bureau of Meteorology.

#### REFERENCES

- Christy, J. R., and W. B. Norris, 2004: What may we conclude about global tropospheric temperature trends? *Geophys. Res. Lett.*, **31**, L06211, doi:10.1029/2003GL019361.
- , and R. W. Spencer, 2005: Correcting temperature data sets. *Science*, **310**, 972–973.
- , and W. B. Norris, 2006: Satellite and VIZ–radiosonde intercomparisons for diagnosis of nonclimatic influences. *J. Atmos. Oceanic Technol.*, **23**, 1181–1194.
- , R. W. Spencer, and E. Lobl, 1998: Analysis of the merging procedure for the MSU daily temperature time series. *J. Climate*, **11**, 2016–2041.
- , —, 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.
- , W. B. Norris, R. W. Spencer, and J. J. Hnilo, 2007: Tropospheric temperature change since 1979 from tropical radiosonde and satellite measurements. *J. Geophys. Res.*, **112**, D06102, doi:10.1029/2005JD006881.
- Diamond, H., and Coauthors, 2001: The United States detailed national report on systematic observations for climate: United States Global Climate Observing System (U.S.–GCOS) Program. NOAA Rep. A183593, 159 pp.
- Durre, I., R. S. Vose, and D. B. Wuertz, 2006: Overview of the integrated global radiosonde archive. *J. Climate*, **19**, 53–68.
- 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.
- Haimberger, L., 2007: Homogenization of radiosonde temperature time series using innovation statistics. *J. Climate*, **20**, 1377–1403.
- Karl, T. R., S. J. Hassol, C. D. Miller, and W. L. Murray, 2006: Temperature trends in the lower atmosphere: Steps for understanding and reconciling differences. U.S. Climate Change Science Program, Synthesis and Assessment Product 1.1 Rep., 180 pp.
- Lanzante, J. R., S. A. Klein, and D. J. Seidel, 2003: Temporal homogenization of monthly radiosonde temperature data. Part I: Methodology. *J. Climate*, **16**, 224–240.
- Mears, C. A., and F. J. Wentz, 2005: The effect of diurnal correction on satellite-derived lower-tropospheric temperature. *Science*, **309**, 1548–1551.
- , M. C. Schabel, and F. J. Wentz, 2003: A reanalysis of the MSU channel 2 tropospheric temperature record. *J. Climate*, **16**, 3650–3664.
- Mo, T., 1995: A study of the Microwave Sounding Unit on the NOAA-12 satellite. *IEEE Trans. Geosci. Remote Sens.*, **33**, 1141–1152.
- Parker, D. E., M. Gordon, D. P. N. Cullum, D. M. H. Sexton, C. K. Folland, and N. Rayner, 1997: A new global gridded radiosonde temperature data base and recent temperature trends. *Geophys. Res. Lett.*, **24**, 1499–1502.
- Randall, R. M., and B. M. Herman, 2008: Using limited time period trends as a means to determine attribution of discrepancies in microwave sounding unit–derived tropospheric temperature time series. *J. Geophys. Res.*, **113**, D05105, doi:10.1029/2007JD008864.
- Randel, W. J., and F. Wu, 2006: Biases in stratospheric and tropospheric temperature trends derived from historical radiosonde data. *J. Climate*, **19**, 2094–2104.
- Sherwood, S., J. Lanzante, and C. Meyer, 2005: Radiosonde daytime biases and late-20th century warming. *Science*, **309**, 1556–1559.
- 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.
- Zou, C.-Z., M. D. Goldberg, Z. Cheng, N. C. Grody, J. T. Sullivan, C. Cao, and D. Tapley, 2006: Recalibration of microwave sounding unit for climate studies using simultaneous nadir overpasses. *J. Geophys. Res.*, **111**, D19114, doi:10.1029/2005JD006798.