Dry Bias and Variability in Vaisala RS80-H Radiosondes: The ARM Experience

D. D. TURNER

Pacific Northwest National Laboratory, Richland, Washington, and University of Wisconsin—Madison, Madison, Wisconsin

B. M. LESHT

Argonne National Laboratory, Argonne, Illinois

S. A. CLOUGH

Atmospheric and Environmental Research, Inc., Cambridge, Massachusetts

J. C. LILJEGREN

Argonne National Laboratory, Argonne, Illinois

H. E. REVERCOMB AND D. C. TOBIN

University of Wisconsin—Madison, Madison, Wisconsin

(Manuscript received 25 February 2002, in final form 28 June 2002)

ABSTRACT

Thousands of comparisons between total precipitable water vapor (PWV) obtained from radiosonde (Vaisala RS80-H) profiles and PWV retrieved from a collocated microwave radiometer (MWR) were made at the Atmospheric Radiation Measurement (ARM) Program's Southern Great Plains Cloud and Radiation Testbed (SGP CART) site in northern Oklahoma from 1994 to 2000. These comparisons show that the RS80-H radiosonde has an approximate 5% dry bias compared to the MWR. This observation is consistent with interpretations of Vaisala RS80 radiosonde data obtained during the Tropical Ocean Global Atmosphere Coupled Ocean–Atmosphere Response Experiment (TOGA COARE). In addition to the dry bias, analysis of the PWV comparisons as well as of data obtained from dual-sonde soundings done at the SGP show that the calibration of the radiosonde humidity measurements varies considerably both when the radiosondes come from different calibration batches and when the radiosondes come from the same calibration batch. This variability can result in peak-to-peak differences between radiosondes of greater than 25% in PWV. Because accurate representation of the vertical profile of water vapor is critical for ARM's science objectives, an empirical method for correcting the radiosonde humidity profiles is developed based on a constant scaling factor. By using an independent set of observations and radiative transfer models to test the correction, it is shown that the constant humidity scaling method appears both to improve the accuracy and reduce the uncertainty of the radiosonde data. The ARM data are also used to examine a different, physically based, correction scheme that was developed recently by scientists from Vaisala and the National Center for Atmospheric Research (NCAR). This scheme, which addresses the dry bias problem as well as other calibration-related problems with the RS80-H sensor, results in excellent agreement between the PWV retrieved from the MWR and integrated from the corrected radiosonde. However, because the physically based correction scheme does not address the apparently random calibration variations observed, it does not reduce the variability either between radiosonde calibration batches or within individual calibration batches.

1. Introduction

The vertical distribution of water vapor strongly affects the magnitude of radiative fluxes in the atmosphere, both at the surface and at the top of the atmosphere, and thus plays a vital role in energy transport and cloud processes. The vertical distribution of water vapor in the atmosphere is quite variable, decreasing several orders of magnitude from the surface to the tropopause. Measurements of the water vapor distribution are critical for a variety of research efforts including boundary layer studies, polar meteorology, cloud processes, atmospheric chemistry, climate studies, and severe weather prediction (Weckworth et al. 1999). However, obtaining highly accurate water vapor measurements, especially in operational settings, has been a long-standing, but unsolved, problem.

Corresponding author address: David D. Turner, CIMSS, University of Wisconsin—Madison, 1225 W. Dayton St., Madison, WI 53706.
E-mail: dturner@ssec.wisc.edu

Water vapor profiles are typically obtained from radiosonde observations, both in the global observation network and in many field experiments. Several studies have reported on the importance of the inconsistency in radiosonde data in global analyses (Elliott and Gaffen 1991; Parker and Cox 1995; Soden and Lanzante 1996). While these studies have described the impact of biases between different sensor types, other studies have pointed to the importance of understanding the bias in the response of single sensors. Barr and Betts (1997) and Fleming (1998), for example, both describe a dry bias in the Vaisala RS80 humidity sensor observed during field campaigns. More recently, Zipser and Johnson (1998) and Guichard et al. (2000) explored the consequences of a dry bias in radiosonde relative humidity measurements on interpretation of data from Tropical Ocean Global Atmosphere Comprehensive Ocean–Atmosphere Response Experiment (TOGA COARE). Lesht (1998) discusses some implications of the dry bias in radiosonde measurements made by the Atmospheric Radiation Measurement (ARM) program.

The absolute accuracy of the water vapor profiles is especially important to ARM because the radiative transfer models upon which much ARM research is based are very sensitive to the amount and distribution of atmospheric water vapor (Revercomb et al. 2003). Modelers who have attempted to calculate the downwelling longwave radiance at the surface over the Southern Great Plains (SGP) site have found much more variability in the observed versus calculated longwave radiance residuals than was expected on the basis of similar studies done in other experiments (Clough et al. 1999). Analysis of this variability suggested that it was due to the uncertainty in the water vapor profiles used in the model calculations. The uncertainty in the atmospheric water vapor profiles used in these calculations, which was provided solely by radiosondes, limited the ability of ARM scientists to improve longwave radiative transfer models. Because of its importance in radiative transfer modeling and for understanding cloud processes, improvement of atmospheric water vapor observations has been made a priority by the ARM program.

The purpose of this paper is to describe ARM's approach to improving the accuracy of radiosonde measurements of water vapor. The approach is based on analysis of the extensive collection of soundings done since 1992 at ARM's SGP Cloud and Radiation Testbed (CART) site in north-central Oklahoma (Stokes and Schwartz 1994). By comparing the water vapor measurements from the radiosondes with observations obtained from collocated water vapor remote sensing instruments, we have found that scaling the radiosonde absolute humidity measurements can greatly improve the accuracy of the radiosonde water vapor profile. These improvements are quantified by reductions in the residuals between model calculations and observations of longwave radiance at the surface, and by qualitative measures of water vapor profile shape in the planetary boundary layer.

We briefly describe the instruments used in this study, which include radiosondes, a two-channel microwave radiometer, and a high-resolution infrared interferometer in section 2. In section 3 we use these long time series of coincident data, together with data collected during three intensive observation periods (IOPs) that focused on water vapor, to characterize the Vaisala radiosondes launched by the ARM program at its SGP CART site. Section 4 describes an empirical correction scheme that is applied routinely to ARM radiosonde data, which removes the diurnal difference discovered in radiosonde data as well as significantly reduces the amount of scatter in the radiosonde data. We compare the ARM correction to a physically based correction developed by the manufacturer in section 5. Finally, the paper concludes with a summary of the results and of the remaining issues that need to be addressed.

2. Instrumentation

a. Radiosondes

The ARM program has launched thousands of radiosondes at the SGP CART site since the program began in 1992. ARM has used the Vaisala[1] RS80-15LH radiosonde for all its soundings. This radiosonde incorporates the H-humicap capacitive moisture sensor that is more sensitive and stable than the more commonly used A-humicap (Antikainen and Paukkunen 1994). The 1.5-s raw data sent from the radiosonde are processed with the standard ground-station software, and quality controlled (i.e., filtered, edited, and interpolated) before being output with 2-s resolution. All the soundings are done with 350-g balloons and have an approximate ascent rate of 5 m s^{-1}. ARM currently launches four radiosondes per day at the SGP central facility (near Lamont, Oklahoma). During IOPs, which typically occur 3–5 times per year, radiosondes are launched 8 times per day at the central facility and at each of four boundary facilities located approximately 150 km away from the central site. Thus our dataset spanned all seasons and all times of day. (Additional information on radiosondes used in the ARM program may be found online at http://www.arm.gov/docs/instruments/static/bbss.html.)

Because radiosondes are designed for operational use in weather forecasting, they often are not considered to be ''research-grade'' instruments. Since its inception, ARM has tried to minimize the uncertainty in radiosonde data that might be attributed to operator effects. All launches are done following standardized procedures. Ancillary data, which includes surface reference values of temperature, pressure, and relative humidity, are logged by the operators, and the operators minimize

[1] Mention of commercial products does not imply endorsement by the U.S. Department of Energy or its contractors.

both the time they spend handling the radiosondes and the time the radiosondes spend in the launcher before the balloon is released. To further minimize any potential effects of solar heating of the radiosonde before launch, ARM developed a system in September 1996 by which the radiosondes are aspirated with ambient air in a shaded chamber until a few seconds before launch (Richardson et al. 2000). The temperature and relative humidity of the ambient air entering the chamber as well as the radiosonde readings are recorded for each sounding.

b. Microwave radiometers

The microwave radiometers (MWRs) used at all of the ARM facilities are Radiometrics WVR-1100 radiometers. These instruments are two-channel systems that measure downwelling radiation at 23.8 and 31.4 GHz. These two frequencies allow simultaneous determination of the water vapor and liquid water burdens along a selected path (nominally zenith at the ARM sites). Atmospheric water vapor observations are made near the ''hinge point'' of the 22.2-GHz water vapor line where the vapor emission does not change with pressure and hence is altitude independent. [The actual hinge point for the 22.2-GHz water vapor line is approximately 24.4 GHz (S.A. Clough, 2001, personal communication).] Water vapor emission dominates the 23.8-GHz observation, whereas cloud liquid water, which emits in a continuum that increases with frequency, dominates the 31.4-GHz signal. The temporal resolution of the sky brightness temperature data observed by the instrument is 20 s. By observing these two frequencies, the water vapor and liquid water signals can be separated using a retrieval model based on the Liebe87 model (Liebe and Layton 1987) to retrieve total precipitable water vapor (PWV) and the integrated cloud liquid water path. Details of the retrieval of these parameters are given in Liljegren and Lesht (1996). It should be noted that although the retrieval algorithm was developed by using radiosonde profiles to drive the Liebe model, the MWR retrievals are independent of the ARM radiosonde measurements because both the retrieved variable (PWV) and the observed variables (brightness temperatures) were calculated by the model.

As with any instrument, calibration of the MWR, that is, the transformation of the raw voltage measurements made by the instrument to brightness temperature, is critical. The ARM MWRs are calibrated by the tipping curve method (Han et al. 1994). An automated procedure has been developed that permits the calibration to be automatically and continuously maintained (Liljegren 2000). This calibration algorithm is able to maintain the MWR's calibration to 0.2–0.3 K rms. This corresponds to an uncertainty in PWV due to MWR observation uncertainty of less than 0.3 mm. This autocalibration algorithm was installed on the MWRs in October 1998. Data collected before this point in time have been postprocessed to ensure a consistent, high-quality, calibrated dataset.

c. Atmospheric emitted radiance interferometer

The Atmospheric Emitted Radiance Interferometer (AERI) is a fully automated ground-based passive infrared interferometer. This instrument measures downwelling atmospheric radiance spectra from 3.3 to 18 μm (3000–550 cm^{-1}) with a spectral resolution of approximately 1 cm^{-1}. A typical measurement cycle consists of a 3-min sky dwell period, followed by 2-min dwell periods on each of the two calibration blackbodies. One blackbody is fixed at 60°C while the other is at ambient temperature, and these temperatures are monitored with high-precision thermistors. This measurement strategy allows the AERI to be self-calibrated for each observation cycle, resulting in absolute calibration accuracy better than 1% of the ambient radiance (Revercomb et al. 1993). This measurement strategy also ensures that the AERI instrument is very stable over time, making it an excellent reference to judge the stability of other instruments. More details on the AERI are given in Revercomb et al. (1993).

3. Characterizing the RS80-H radiosonde

Several studies intended to characterize the operational performance of the RS80-H radiosonde have been conducted by ARM since the program began in 1992. These studies included comparisons of radiosonde readings with stable reference values at the surface and of data collected during ascending and descending phases of radiosonde flights (Lesht 1995). Subsequent work (Lesht and Liljegren 1997) focused on comparison of PWV measurements obtained from the MWRs deployed at the CART site with PWV calculated from concurrent radiosonde soundings. One result of this comparison was the discovery that the discrepancies between the two measurements had a strong correlation with the radiosonde's manufacturing date. This result was particularly important because it revealed that variations in the manufacturer's calibration process were a possible source of measurement error and that, as a result, the uncertainty associated with the radiosonde measurements of relative humidity was greater than had previously been supposed. Comparisons between observed downwelling infrared radiance and calculations using radiosondes to define the atmospheric state also highlighted issues with the radiosonde moisture profile (Clough et al. 1996). Miloshevich et al. (2001) presented a detailed discussion of calibration issues affecting the Vaisala RS80 humidity sensors.

a. Long-term comparison with collocated microwave radiometer

The discrepancies between the PWV from the MWR and from the radiosondes launched by ARM between

April 1994 to July 2000 are shown as a function of manufacture date (hereafter referred to as calibration batch[2]) in Fig. 1. The MWR data used here were averaged over a 40-min window centered on the radiosonde launch time. The comparisons were limited to flights during which the radiosondes reached at least 10 km (and thus profiled approximately 99.9% of the water vapor in the column) and during which the retrieved integrated cloud liquid water amount from the MWR was less than 50 μm. This latter condition was intended to avoid possible retrieval errors due to thick liquid water cloud or precipitation from contaminating the MWR results. Over 4000 flights satisfied these criteria.

The distributions of the ratio of PWV values (MWR/radiosonde) showed that the radiosondes almost always measured less water vapor than did the MWR. The mean value of the PWV ratio from batches 5456 to 8284, after which Vaisala changed the radiosonde packaging in an effort to address the dry bias problem, is 1.05 with a standard deviation of 0.06. The batch-to-batch variability represented by the differences in the median value of this ratio for each batch (dark horizontal lines in Fig. 1) is immediately apparent. Batch-median ratios ranged from approximately 0.98 (batch 6372) to 1.24 (batch 8233). The variation within each batch, as represented by the error bars that are ± 1 standard deviation about the mean, averages approximately 7%. These variations, both within and between calibration batches, can result in peak-to-peak differences in PWV between any two radiosondes of greater than 25%. Although some of this variability may be due to the ''age'' of the radiosonde, that is, the interval between the date of manufacture and launch, a substantial portion is otherwise unexplained as demonstrated below.

In addition to the overall dry bias, analysis of these data suggests that there is a diurnal dependence in the magnitude of the bias. We find that the PWV ratio is typically 3%–4% higher for soundings done during the daytime than soundings done using radiosondes from the same batch launched during the night; that is, that daytime radiosondes are drier than nighttime radiosondes. The standard Vaisala processing software does apply a solar radiation correction to the radiosonde temperature measurements in the daytime that might account for the difference, but this correction is applied only to temperatures obtained at altitudes above 500 mb and the magnitude of the correction is not sufficient to account for the discrepancy. Another possible cause is diurnal variation in the MWR calibration, but careful radiometric intercomparisons suggest that the MWR calibration is very stable. Comparisons of downwelling infrared radiance observations (AERI) and calculations with a line-by-line model also support the conclusion that the diurnal difference is due to the radiosonde (sec-

[2] The appendix describes how the radiosonde's manufacture date, or its calibration batch, may be determined from the radiosonde's serial number.

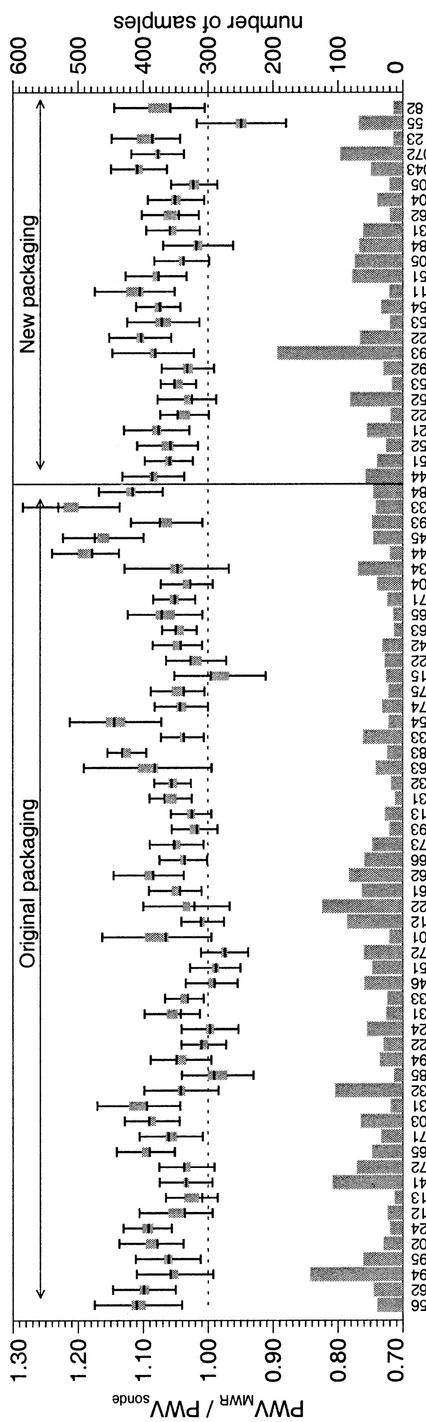

FIG. 1. Comparisons of the ratio of PWV from the MWR to the radiosonde from Apr 1994 to Jul 2000 separated by calibration batch. Data are included only if the radiosonde achieved a height of at least 10 km, the liquid water path retrieved from the MWR was less than 50 μm, and the number of such samples in the calibration batch is greater than 10. The histogram at the bottom indicates the number of MWR/sonde comparisons in each batch. The error bars represent ± 1 std dev about the mean ratio for the calibration batch, where the gray boxes indicate ± 1 standard error of the mean. The dark horizontal line indicates the median value of the batch. The period with the original packaging, which is hypothesized to be responsible for the contamination dry bias, as well as the period where the packaging was changed in an attempt to eliminate the dry bias, are indicated.

TABLE 1. Mean (std dev) of the ratio of PWV (MWR to radiosonde). After 11 Sep 1996 all radiosonde sensor packages were aspirated before launch. Only radiosondes prior to batch 8344 (original packaging) are included in this table.

	MWR / radiosonde	MWR / corrected radiosonde	No. of points
All data	1.052 (0.061)	0.992 (0.054)	2195
All nighttime data	1.031 (0.055)	0.973 (0.049)	910
All daytime data	1.067 (0.064)	1.008 (0.058)	1285
Nightime data before 9/11/96	1.047 (0.053)	0.980 (0.045)	263
Daytime data before 9/11/96	1.084 (0.056)	1.015 (0.046)	342
Nightime data after 9/11/96	1.024 (0.054)	0.971 (0.051)	647
Daytime data after 9/11/96	1.062 (0.066)	1.005 (0.062)	943

tion 4). Sensor arm heating (Cole and Miller 1995; Wang et al. 2002) is the most likely reason for these differences, but we found that the daytime–nighttime dependence of the PWV ratios shows no significant change after the radiosonde aspiration procedure was begun on 11 September 1996 (Table 1). Thus, if sensor arm heating is the source of the bias, it must result either from heating that occurs during the brief interval (<60 s) between the time that the radiosonde is removed from the aspirator and the time it is launched, or from heating during flight. Both possibilities are currently under investigation.

b. Intensive observation periods

In order to focus effort on water vapor measurement issues, ARM conducted three IOPs during which the standard CART instrument suite was augmented by additional ARM and visiting investigator instrumentation and the frequency of soundings was increased to 8 soundings per day for a 3-week period. These IOPs, which were conducted in the fall of 1996 and 1997 and again in the fall of 2000, are described in detail by Revercomb et al. (2003).

Because measurement of atmospheric state profiles are so important to ARM, detailed characterization of radiosonde observations was a primary objective of the first two IOPs. Dual-sonde soundings, in which two radiosondes (of the same model/manufacturer) were flown on the same balloon, were a major component of these experiments. The goal for the dual-sonde soundings was to quantify sonde-to-sonde variation both within and between calibration batches. Because both radiosondes would be simultaneously sampling the same environment, any differences in measurement likely reflect the operation uncertainty of the sensor. Furthermore, because the dual-sonde approach involved only comparisons between radiosondes, we avoided complications resulting from the lack of an accepted absolute standard for the measurements. Details of the dual-sounding procedures and overview of the results may be found in Lesht (1998, 1999).

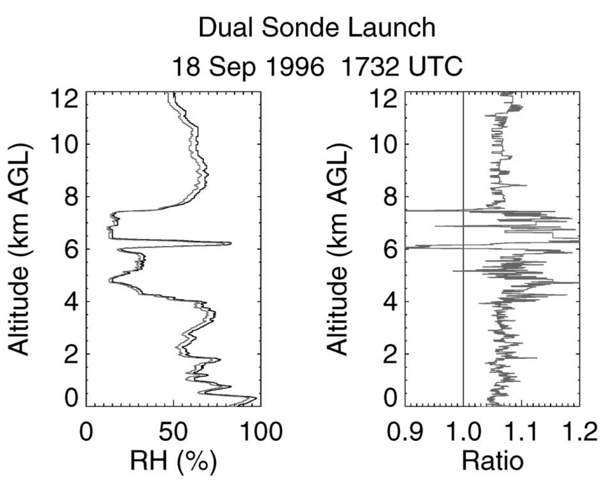

FIG. 2. An example of a dual-sonde launch at 1732 UTC 18 Sep 1996. Note the height-independent nature of the ratio between the two moisture profiles in the troposphere.

Over 90 of the dual-sonde soundings that were conducted during the 1996 and 1997 WVIOPs satisfied our minimum comparison criterion (good data obtained from both radiosondes above 10-km altitude). A typical example of the differences in RH obtained from two radiosondes flown in a dual-sounding is shown in Fig. 2. This example shows that although both radiosondes seem to measure the same vertical structure of the RH distribution, the absolute RH difference between the two is on the order of 5% RH throughout the sounding (as indicated by the relatively constant ratio of the humidity as a function of height in the right-hand panel of Fig. 2). The manufacturer cites an accuracy of ±2% for the RH sensor used in the RS80 (Vaisala 1996). This value, however, is the standard deviation of a series of differences obtained by comparing radiosonde readings made in a calibration chamber to an accurate reference. Assuming a normal distribution and no bias, this suggests that approximately 95% of the radiosonde RH readings should be within 4% of the true value. Our results (Fig. 1) demonstrate that the radiosonde RH values are not unbiased and that the bias depends on the calibration batch. The dual-sounding experiments done during the IOPs were designed to explore this dependence.

During each of the 1996 and 1997 WVIOPs we used radiosondes from two distinct calibration batches. In 1996 the radiosondes came either from batch 6231XXXXX, which were calibrated in June 1996, or from batch 6322XXXXX, which were calibrated in August 1996. In 1997, we used radiosondes from batch

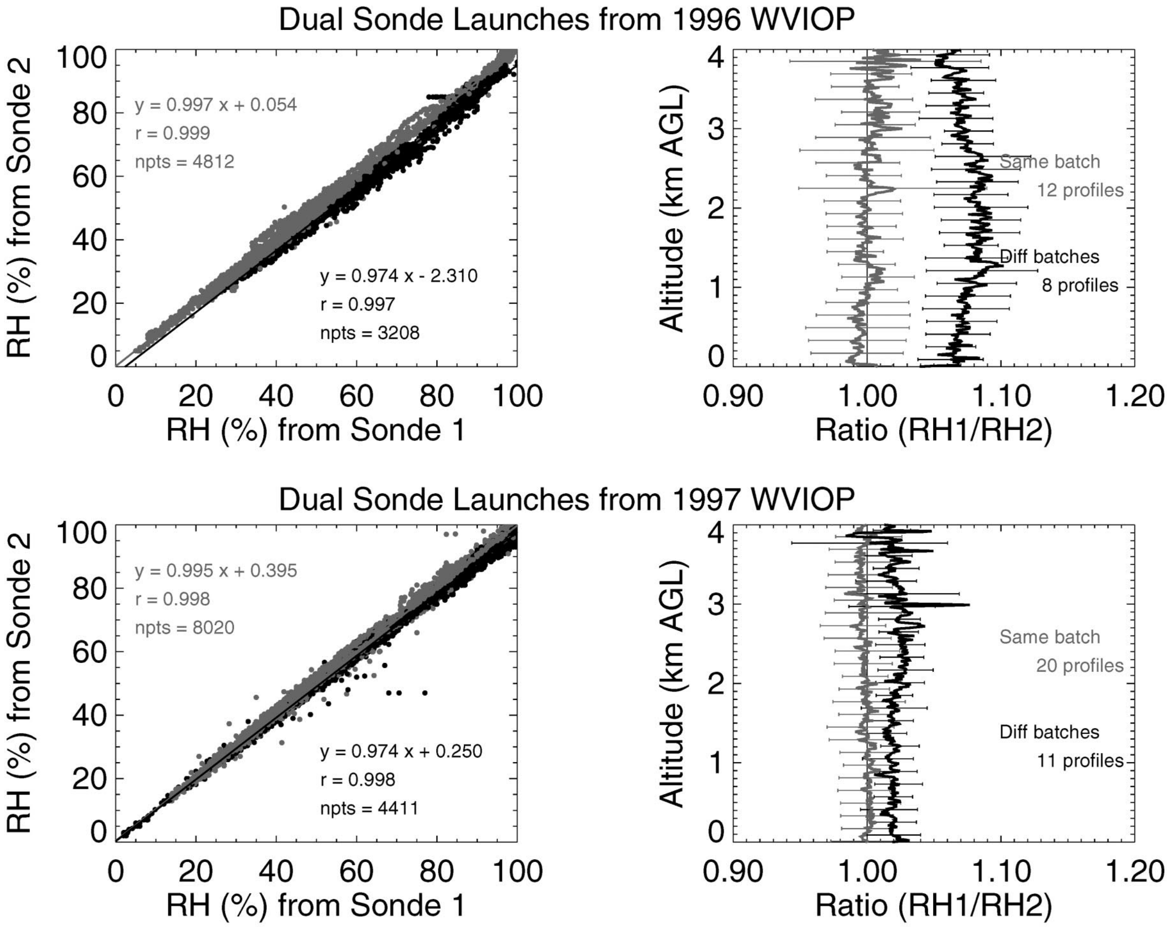

FIG. 3. A statistical analysis of the dual-sonde launches from the 1996 (above) and 1997 (below) water vapor IOPs. When sondes from different calibration batches (shown in black) were analyzed, the newer batch was designated as ''sonde 1.'' The error bars in the right-hand panels represent ± 1 std dev of the RH ratios at 25 vertical levels.

7263XXXXX (June 1997) and batch 7333XXXXX (August 1997). In total, we made 61 successful dual-sonde flights in 1996 (35 mixed-batch and 26 same-batch) and 66 in 1997 (20 mixed-batch and 40 same-batch). Because we further screened the dual-soundings for flights that exceeded 10 km in altitude, the comparisons reported here are based on somewhat fewer flights.

One basic result of the dual-soundings is shown in Fig. 3. The panels on the left are scatterplots of the point-by-point RH values for data below 4 km (since downwelling IR radiation is dictated, for the most part, by these lower levels) obtained by the two radiosondes for both the mixed-batch and same-batch pairs. The most recently calibrated radiosonde in each mixed-pair sounding (e.g., from batch 6322XXXXX in 1996) was designated ''sonde 1'' for this analysis. The mixed-batch comparisons show a definite RH bias between batches for both years. In each case, the ''sonde 2'' (older) radiosondes measure lower values of RH than do the new radiosondes. In addition to the average bias indicated by the intercept of the linear fit to the data, the fact that the slope is less than one suggests an RH dependence as well. When looked at in terms of the RH ratio as a function of altitude (right panels in Fig. 3), we find that the same-batch radiosondes are unbiased and that the RH variability (as opposed to accuracy) is close to $\pm 2\%$ RH. Radiosondes from different calibration batches, however, show a distinct difference, with the average RH ratio of approximately 7% in 1996 and 2% in 1997 (right panels of Fig 3). The average age difference between the two batches of sondes was 109 and 93 days for 1996 and 1997 IOPs, respectively, and this slight difference in age could be responsible for the majority of this change (7% difference in 1996 vs 2% difference in 1997) due to the nature of the contamination of the RH sensor (see section 5). Furthermore, the RH ratio appears to be independent of altitude in the lower troposphere.

The radiosonde is an RH sensor. That is, the fundamental variable measured by the radiosonde is RH as opposed to absolute water vapor density, or dewpoint. Because much of our earlier work was based on comparing the height-integrated absolute water vapor density (i.e., PWV) obtained from the radiosondes with that

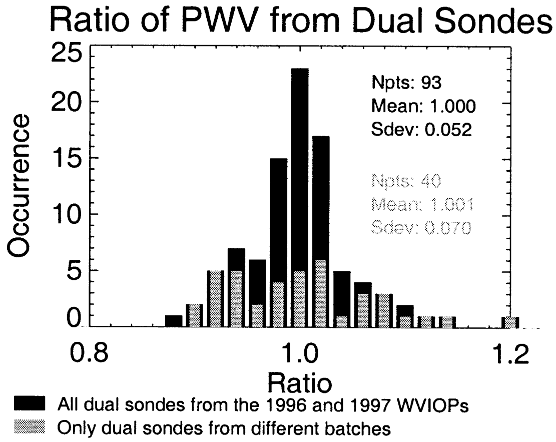

FIG. 4. Histogram of the ratio of PWV from the 93 dual-sonde launches from the 1996 and 1997 WVIOPs. The subset of dual-soundings that were mixed-batch is indicated in gray.

measured by the MWR we calculated the PWV values from the pooled dual soundings and compared the distribution of the PWV ratio from the paired sondes. The distributions for the mixed-batch and same-batch flights are shown in Fig. 4. Although it is not readily apparent in Fig. 3, the results in Fig. 4 show that the range in PWV ratios can be on the order of 10% for both the mixed-batch and same-batch pairs. Clearly, the same-batch pairs are more likely to be closer to each other than are the mixed-batch pairs, but with the exception of a few soundings, the tails of the distributions are very similar indicating that within-batch variations also can result in large relative errors in PWV. It should be noted that this plot does not show any bias because we did not order the ratio calculation by calibration batch.

4. The ARM-scaling correction

Because the observed bias and variability in the radiosonde RH values are greater than can be used to validate radiative transfer models, we have developed a simple method for correcting the radiosonde RH values for bias that also substantially reduces its calibration uncertainty. This correction method is based on two assumptions: 1) that the PWV measured by the MWR can be used as a stable and accurate reference variable, and 2) that the differences in radiosonde absolute humidity are independent of altitude and can thus be corrected by application of a constant scale factor. The first assumption is supported by years of experience with long-term deployments of MWR sensors by ARM (Liljegren 2000). The second assumption is supported by the dual-sonde observations (Figs. 2, 3).

The basic principle of the scaling correction is to adjust the radiosonde RH values to force the PWV obtained from the radiosonde profile to match the PWV obtained simultaneously from the colocated MWR. This is accomplished by calculating a scale factor that is simply the ratio of the MWR PWV to the radiosonde PWV. The radiosonde profile is adjusted by multiplying the water vapor mixing ratio at each level by the scaling factor, and then using the radiosonde temperature and pressure values to convert the scaled mixing ratio to RH.

Of course, it is necessary to evaluate this correction by using a completely independent source of data. We have used these radiosonde profiles as input to a line-by-line radiative transfer model (LBLRTM; Clough et al. 1992; Clough and Iacono 1995). The LBLRTM is used to calculate the downwelling longwave radiance spectrum at the surface, which then is compared with the longwave radiance spectrum observed by the AERI. We use the spectrum of the residuals between the model and the observations as a measure of the radiosonde accuracy.

Figure 5a demonstrates typical clear-sky AERI spectra for two cases: one where the water vapor content is close to the maximum value typically observed during the summer at SGP, and one where the water vapor content is close to its annual minimum. Figure 5b shows the average residuals calculated by subtracting the spectral radiance obtained by the LBLRTM driven by the unscaled and scaled radiosonde profiles from the corresponding AERI observations. If both the LBLRTM and the AERI are accurate, then the difference in the residuals is due primarily to errors in the radiosonde water vapor profile. The results in Fig. 5b, which were based on a population of 745 clear-sky cases spanning a period from April 1994 to December 1998 show that the residuals based on the MWR-scaled radiosonde data are much lower than those based on the unscaled radiosonde data. The sign of the average residuals for the unscaled data is positive, again indicating a dry bias in the radiosonde RH measurement. However, the MWR-scaled radiosondes also significantly reduce the standard deviation of the longwave residuals between the AERI observations and the model, as compared to the model runs that utilized the unscaled radiosonde profiles (Fig. 5c). The rms difference between the AERI observed and modeled residuals for the unscaled radiosonde calculations is 2.2 mW $(m^2 \text{ ster cm}^{-1})^{-1}$, while the rms difference using MWR-scaled radiosondes is 1.1 mW $(m^2 \text{ ster cm}^{-1})^{-1}$. Given the demonstrated stability of the AERI instrument (due to its calibration sequence), this indicates that the MWR scaling significantly reduces the variability in the radiosonde moisture profiles.

Further evidence supporting the value of MWR-scaling may be seen in the AERI/LBLRTM analysis, when the results are separated diurnally, as in Fig. 6 (Turner et al. 1998). Driving the LBLRTM with the MWR-scaled and unscaled radiosonde profiles at night results in a very slight decrease of the mean spectral residual, indicating that the radiosonde was moistened slightly by the scaling. However, during the day, the MWR-scaling results in a large decrease in the mean residual (corresponding to a much larger moistening of the ra-

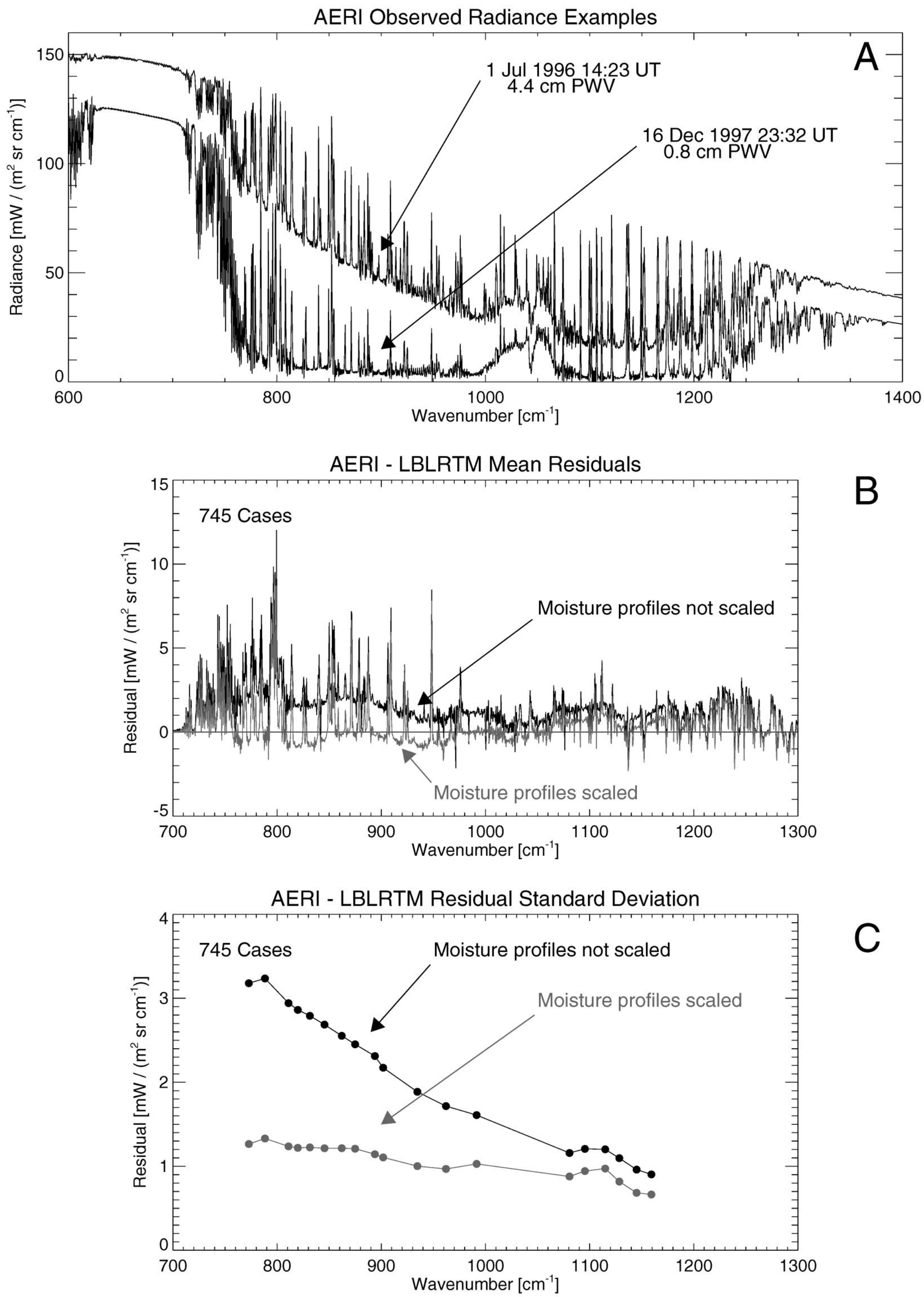

FIG. 5. (a) Typical clear-sky downwelling longwave spectra observed by the AERI for large water vapor burdens (4.4 cm PWV, near the maximum observed at SGP) and small water vapor burdens (0.8 cm PWV, near the minimum observed at SGP). (b) Mean spectral residuals from the AERI − LBLRTM, where the LBLRTM utilized moisture profiles from the original (black) and MWR-scaled (gray) radiosondes. (c) The spectral standard deviation about the mean residual in (b) for the unscaled and MWR-scaled input into the LBLRTM. The data exhibited a trend as a function of PWV, and this trend was removed in the microwindows (i.e., the spectral regions between strong absorption lines) before the std dev was computed. There were 745 clear-sky cases from Apr 1994 to Dec 1998 used in this analysis. Note that the rms difference of the unscaled and MWR-scaled residuals is 2.2 mW $(m^2\ sr\ cm^{-1})^{-1}$ and 1.1 mW $(m^2\ sr\ cm^{-1})^{-1}$, respectively.

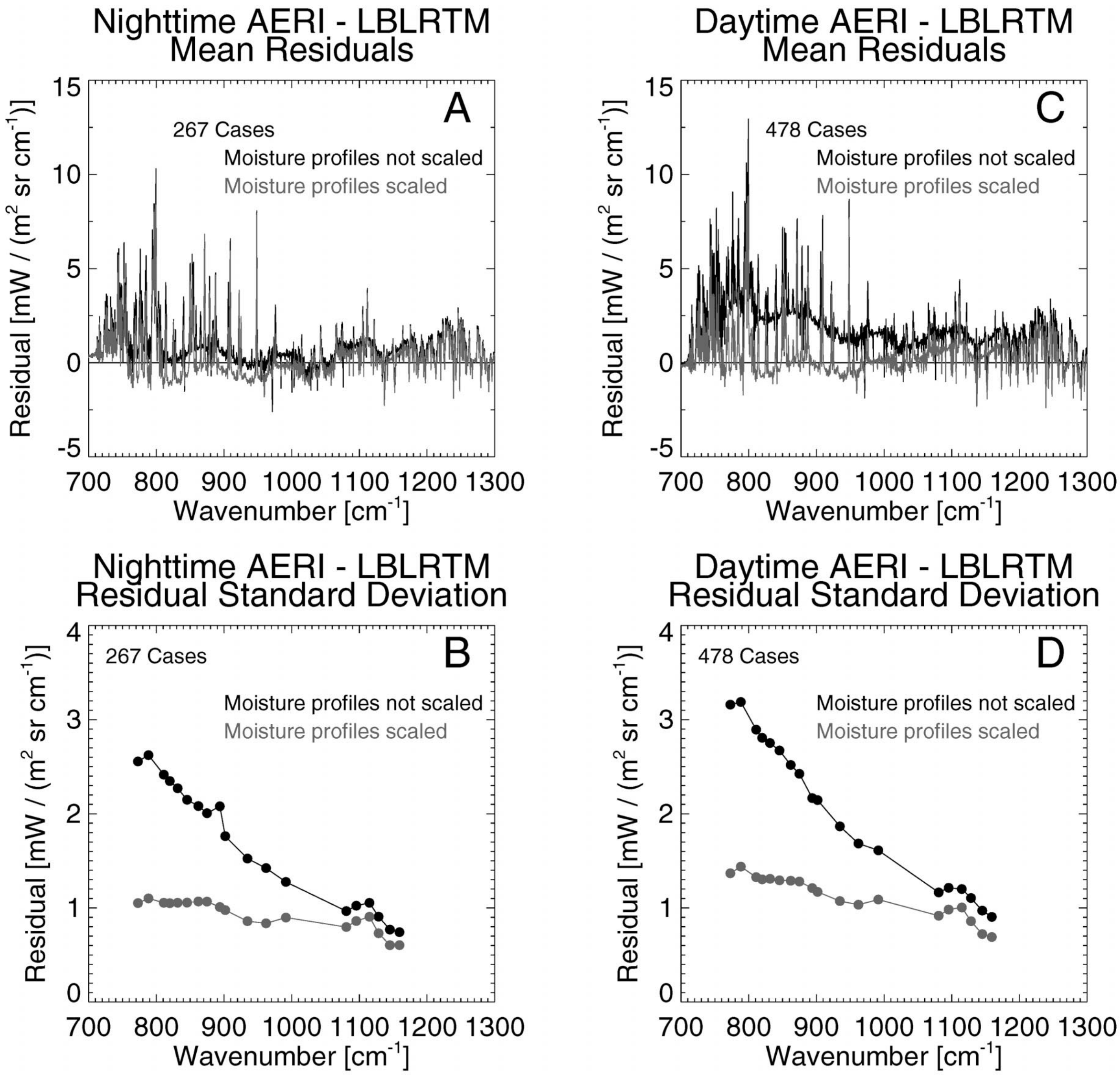

FIG. 6. Same as Figs. 5b and 5c except the results are segregated into (left) nighttime and (right) daytime periods. See text for more details.

diosonde profiles). Both the nighttime and daytime results show that the MWR-scaling significantly reduces the variability in the radiosonde profiles, as indicated by the standard deviation of the residuals about the mean. Also, the MWR-scaled radiosonde results have mean residual and standard deviation spectra that are nearly identical day and night. This indicates that the MWR's sensitivity to PWV does not have a diurnal dependence.

5. Application of a physically based correction

Although ARM observations and data from other experiments have long suggested a problem with the RH calibration of Vaisala radiosondes, the cause of the bias was not clear. Engineers from Vaisala, the radiosonde manufacturer, and scientists from NCAR conducted an extensive array of tests in 1998 on the RS80 radiosonde to explore the dry bias question. They determined that the source of the dry bias was contamination of the capacitive RH sensor by outgassing of the packaging materials used on the RS80 radiosondes (Lesht 1999; Guichard et al. 2000; Wang et al. 2002). This contamination, which was found to effect both the A- and H-type humicap sensors, is cumulative with age and thus the dry bias increases as the sonde became older (Wang et al. 2002). Vaisala attempted to address the contamination problem by changing the dessicant with which the radiosonde was packaged in August 1998 (just prior to batch 8344). However, analysis of MWR and radiosonde data after this modification to the packaging was made (the ''new packaging'' period in Fig. 1) indicates that this change had no statistical effect on the dry bias of the radiosonde with respect to the MWR. Vaisala again modified the packaging of the RS80 radiosonde in August of 2000 by covering the sensor with an inert plastic shield. At the time of this writing, ARM has not collected enough data yet to see if this modification has statistically reduced the dry bias in these radiosondes.

Based on their knowledge of the RH sensor's properties and the results of their laboratory experiments, Vaisala and NCAR have developed a correction algorithm that can be applied to the RS80 data to remove this dry bias (Wang et al. 2002). We were supplied with a proprietary version of this correction (henceforth called the *Vaisala correction*) algorithm under the condition that we would not disclose its details. The paper by Wang et al. (2002) provides a complete explanation of several factors that contribute to errors in the RS80 (both A and H versions) relative humidity measurements and presents the same correction algorithm in a form that does not compromise the Vaisala's trade secrets. Thus, Wang et al. (2002) make the corrections available to the community. We have confirmed the equivalence of both versions but all of our results reported here are based on our application of the original proprietary version.

In general terms, the correction has three component factors: one that accounts for small errors in the basic humidity sensor calibration model; one that accounts for the cumulative affect of the contamination of the sensor with age; and one that includes an improved representation of the humidity sensor's temperature response at saturation (Lesht 1999; Wang et al. 2002). All three components depend on RH, but only the contamination correction depends on the radiosonde age. Therefore, the combined correction is a function age, RH, and ambient temperature, though the contamination is the most significant factor for the RS80-H. Because the RH sensor equilibrates with the packaging contamination over time, the bias increases rapidly with age from zero to about 4% RH after 3 months to a maximum of about 8% RH after 1 yr, assuming that the ambient RH is 80% and the ambient temperature is 0°C. The correction due to the model error under the same conditions is about 0.5% RH, which is approximately offset by the low-temperature saturation error.

We applied the Vaisala correction to over two thousand radiosondes launched at the ARM SGP site. As the correction was developed for data with the original packaging, only radiosondes before batch 8344 have been corrected. Figure 7 shows the ratio of the PWV derived from the corrected sounding to that from the original sounding as a function of sonde age (i.e., the number of days between the time the sonde was originally calibrated and when it was actually launched). In all cases, the correction moistens the sounding. The correction is seen to increase as the radiosonde becomes older, with the size of the correction ranging from less than 2% for relatively "young" radiosondes to over 12% for radiosondes that are approaching one year old.

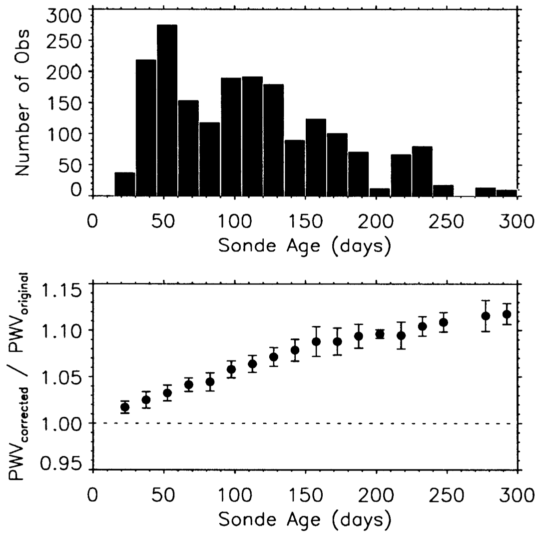

FIG. 7. (top) Histogram showing the age of the radiosondes launched at the SGP CART site from Feb 1996 to Dec 1998. (bottom) The ratio of the PWV from the Vaisala-corrected radiosonde to the uncorrected radiosonde for these radiosondes.

Statistically and when taken as a whole, the Vaisala-corrected radiosondes are in excellent agreement with the MWR (Table 1), as the mean ratio of PWV from the MWR to the corrected radiosonde is 0.992 with a standard deviation of 0.054. However, significant calibration uncertainty still exists in the corrected radiosonde dataset. When the corrected radiosondes are analyzed as a function of batch, significant differences still exist between the mean values of different batches (Fig. 8). For many of the calibration batches, applying the Vaisala correction has resulted in much better agreement with the MWR. However, in some cases, calibration batches that originally agreed well with the MWR, such as batch 6346, are now 6%–7% wetter than the MWR after being corrected. Also, the Vaisala correction has had no impact on the magnitude of the variability of the radiosonde calibration within any particular batch; that is, the variability within any given calibration batch remains the same for both the uncorrected and corrected radiosondes. This point is more easily seen in Fig. 9, which shows the distribution of "scale factors"; that is, the ratio of PWV from the MWR to the PWV from the radiosonde, for both the uncorrected (top panels) and corrected (bottom panel) radiosondes. This figure shows that the correction translates the peak of the distributions by moistening the radiosondes and that the variability is slightly reduced, but that significant (peak-to-peak differences over 25%) variability still exists in the corrected data.

The Vaisala correction also does not affect the diurnal differences in the radiosonde bias. The approximately 3.5% difference between the daytime and nighttime ra-

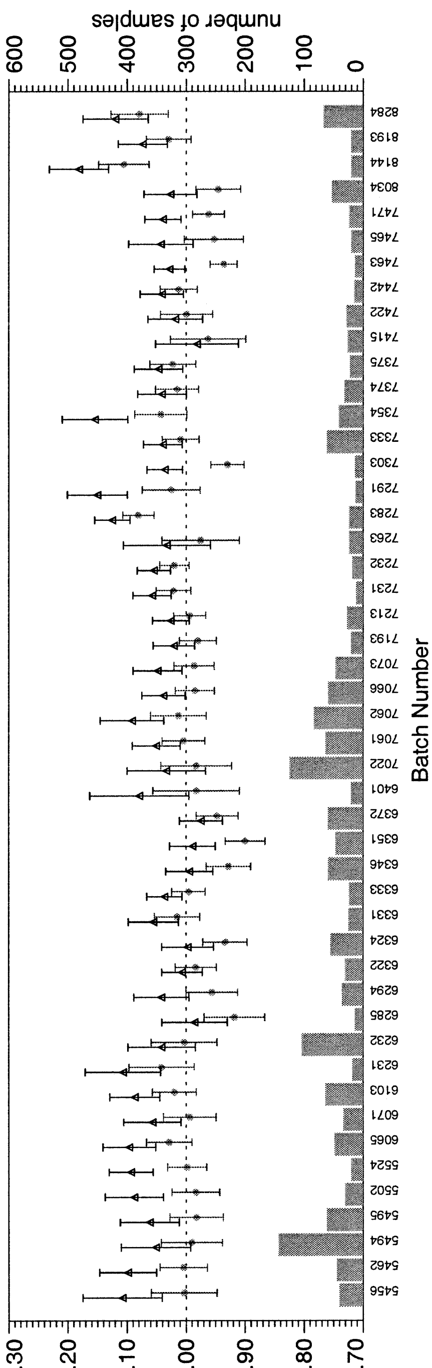

FIG. 8. Comparison of the ratio of PWV from the MWR to the PWV derived from the uncorrected (open triangles) and Vaisala corrected (solid circles) radiosondes from Feb 1996 to Dec 1998. As in Fig. 4, radiosondes that did not achieve an altitude of 10 km were discarded, as well as periods where the MWR reported cloud liquid water above 50 μm. Only data from the "original packaging" period (Fig. 1) are shown here, as the version of the Vaisala correction used in this paper is not appropriate for the "new packaging" period.

diosondes still exists in the corrected radiosonde dataset (Table 1). Figure 10 shows scatterplots of PWV from the MWR and the radiosonde, both corrected and uncorrected, separated into nighttime and daytime. The corrected nighttime radiosondes are in excellent agreement in sensitivity (with a slope of 0.999 $\pm$ 0.003) with the MWR. However, a slight offset (0.5 mm) exists, which introduces a small moist bias into the corrected results.

Unlike scaling the radiosonde's moisture profile to agree with another sensor (like the MWR), the Vaisala correction changes the shape of the profile in the lower troposphere. The impact on the lower troposphere is most easily seen when the planetary boundary layer is well mixed, such as during a warm summer afternoon. In these cases, the water vapor mixing ratio should be nearly constant throughout the well-mixed layer. The uncorrected, MWR-scaled, and Vaisala-corrected radiosonde profiles for two such cases are shown in Fig. 11. Applying the Vaisala correction to the radiosonde profile results in a much better agreement with the MWR-scaled radiosonde profile. Note how the uncorrected and MWR-scaled profiles maintain a nearly constant mixing ratio throughout the depth of the boundary layer (from just above the surface to approximately 2.1 and 2.0 km for 10 September 1996 and 21 July 1998, respectively). The Vaisala correction induces a gradual drying with altitude from the surface to the top of the boundary layer. This linear-like decay could be due to a slight error in the temperature-dependent portion of the Vaisala correction. Unfortunately, the random error in the collocated Raman lidar moisture profile during the middle of the daytime (Turner and Goldsmith 1999) are slightly larger than the differences between these profiles, and thus is unable to offer insight on the validity of either the uncorrected or corrected radiosonde profile.

6. Discussion and summary

The large uncertainty in the absolute calibration of the Vaisala RS80-H radiosonde moisture observations translates into large uncertainties in the applications that use radiosonde data. One of the ARM program's goals is to improve the treatment of atmospheric radiative transfer in climate models, which necessitates the ability to develop and validate radiative transfer models. The comparison of observed (AERI) minus calculated (LBLRTM) radiance residuals demonstrated that the calibration uncertainty of the radiosonde moisture profiles was large and had a diurnal characteristic. The dual-sonde launches demonstrated that the calibration differences between any two radiosondes could be approximated by a height-independent scale factor in the lower troposphere. By scaling the radiosonde moisture profiles to agree with the PWV observed by the MWR, the calibration uncertainty in the radiosonde moisture profile could be markedly decreased and the diurnal characteristic eliminated. This empirical correction has

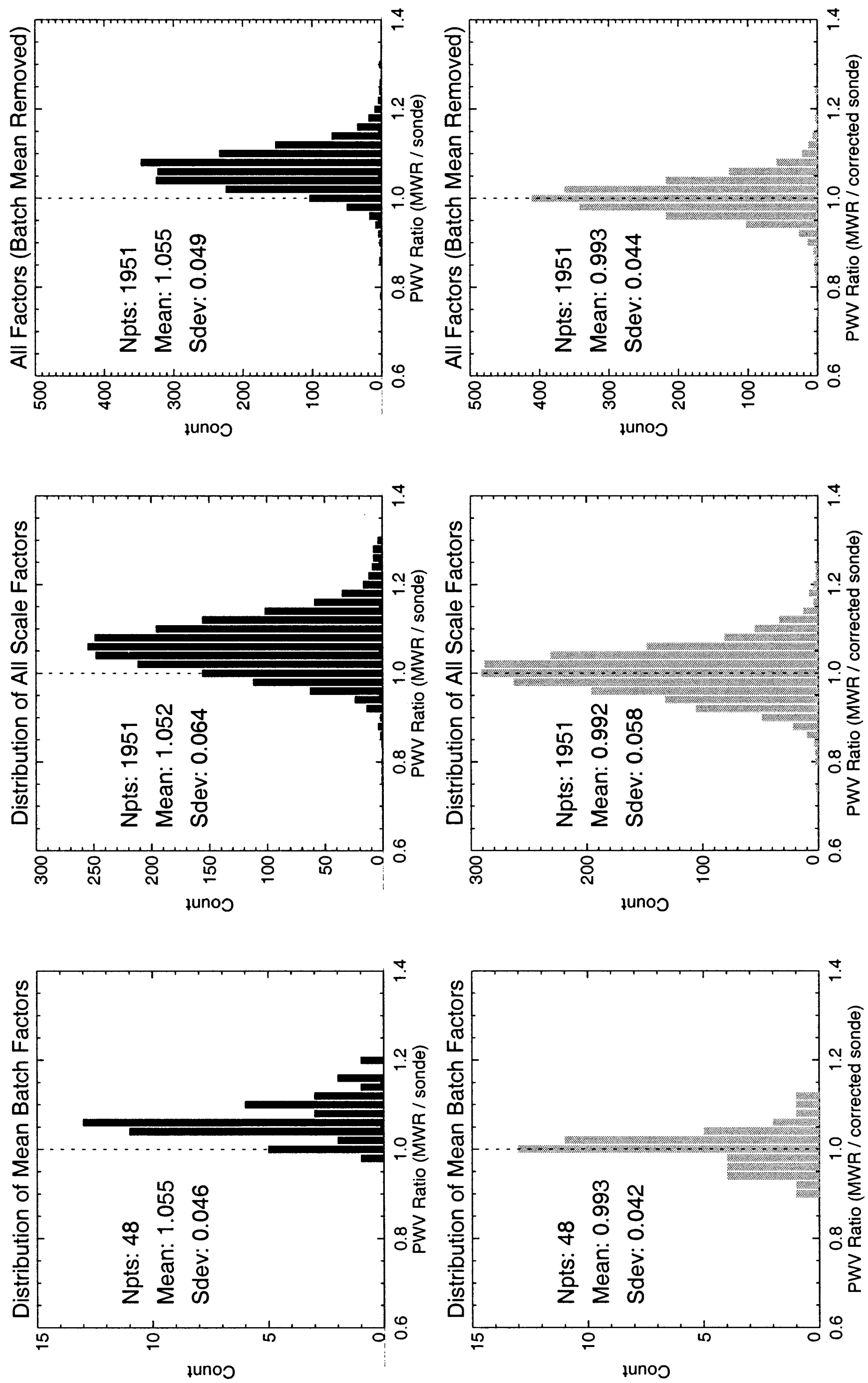

FIG. 9. (left) Histograms of the mean batch "scale factors" (i.e., the ratio of the MWR PWV to the radiosonde PWV) for both the (upper) uncorrected and (lower) corrected radiosondes. (middle) The distribution of all of the scale factors. (right) The distribution of all of the scale factors, where the mean value for the batch the radiosonde is in was removed.

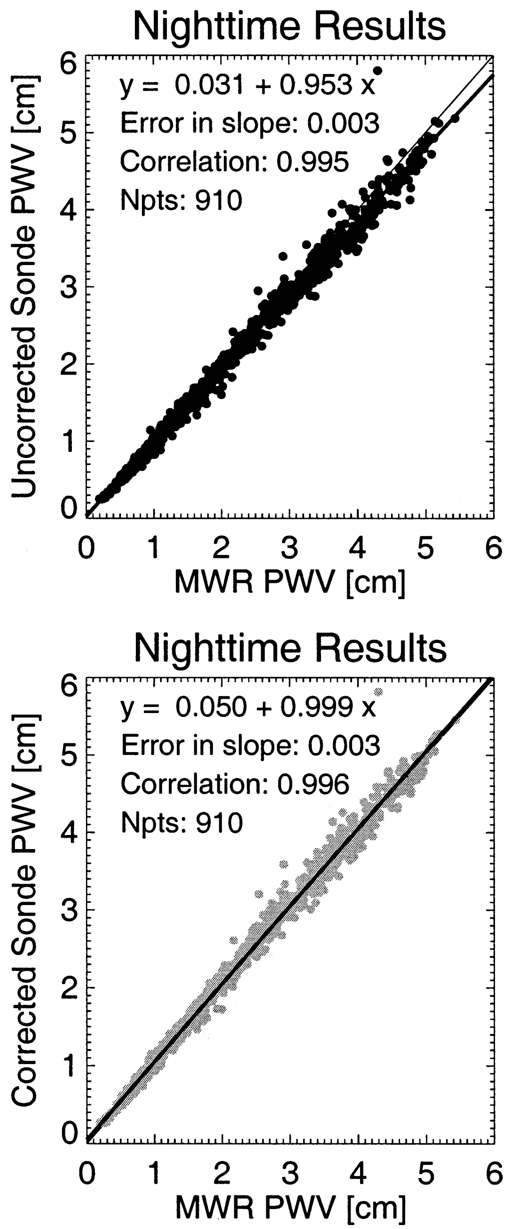

FIG. 10. Scatterplots of PWV from the MWR and the (upper) uncorrected and (lower) corrected radiosondes for the (left) nighttime and (right) daytime. The data used are the same data as used in Figs. 8 and 9.

been automated by the ARM program to create a hybrid sounding product, and all radiosonde moisture data collected by the program are scaled in this manner and provided as an independent data stream to ARM researchers. The appropriateness of this scaling procedure is still being investigated, especially for the upper-tropospheric region.

A physically based correction, which was developed by the radiosonde manufacturer and NCAR from laboratory data to account for the contamination dry bias along with other errors/biases in Vaisala RS80 humidity data, has the desired affect of moistening the radiosonde data. In a mean sense, the correction brings the corrected radiosondes into nearly perfect agreement with the microwave radiometer (in terms of total precipitable water vapor), and corrected nighttime radiosonde data show almost the exact same sensitivity to PWV as does the MWR. Unfortunately, the uncertainty in the calibration of the corrected radiosonde moisture profiles has not been significantly reduced, either within or between the calibration batches. This correction algorithm does not improve diurnal differences seen in the radiosonde calibration. Furthermore, the correction algorithm does not appear to maintain a constant mixing ratio profile in layers that are well mixed. For these reasons, the ARM program

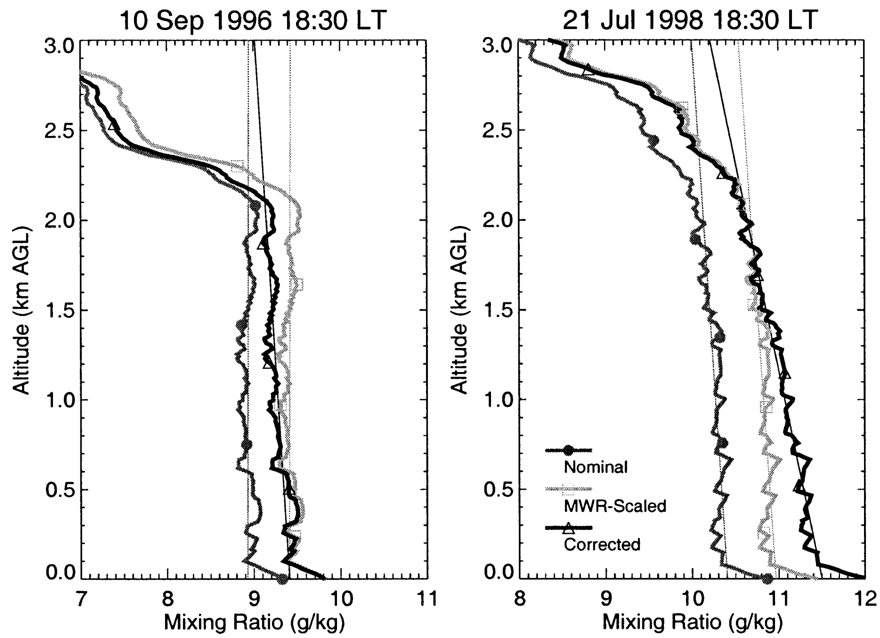

FIG. 11. Examples of water vapor mixing ratio profiles from the uncorrected (nominal), Vaisala corrected, and MWR-scaled radiosondes during two late afternoon cases where the boundary layer was well mixed. The thin solid lines are linear fits of the well-mixed boundary layer (from 200 m to 2.1 km for the 10 Sep 1996 case and to 2.0 km for the 21 Jul 1998 case) to help indicate the trend in the data for the three profiles. Note that the Vaisala-corrected profile dries with altitude in the well-mixed layer.

is continuing to study this correction before applying it to its radiosonde data operationally.

This paper has primarily discussed the radiosonde with respect to the microwave radiometer arguing that the MWR provides a stable reference that can be used to evaluate and scale the radiosonde moisture profile in the lower troposphere. However, the retrieved value may not represent ''truth,'' as calibration issues or inadequacies in the model used to retrieve PWV from the observed microwave brightness temperatures may introduce biases. Also, other water vapor observations, such as the PWV retrieved from GPS observations, could possibly serve as the stable reference that can be used to reduce the calibration uncertainty in radiosonde data. To address this, a set of water vapor experiments at the ARM SGP site was conducted in September 2000 to evaluate the absolute calibration of various water vapor instruments (Revercomb 2000). Analysis of this data is ongoing (some of this is highlighted in Revercomb et al. 2003), and will provide a mechanism to extend this work from RS80-H to the new RS90 sensor, as dual-sonde launches utilizing both were conducted during recent field experiments at the ARM SGP site. These experiments were also conducted in part to help characterize the RS80-H in the upper troposphere (Revercomb 2000; Tobin et al. 2002).

Acknowledgments. This work was conducted as part of the ARM program sponsored by the U.S. Department of Energy, Office of Energy Research, Office of Health and Environmental Research, Environmental Sciences Division. Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by Battelle under Contract DE-AC06-76RLO-18030. The University of Chicago operates Argonne National Laboratory for the U.S. Department of Energy under Contract W-31-109-ENG-38.

APPENDIX

Decomposing the Radiosonde Serial Number to Determine the Calibration Date

Vaisala has recorded the calibration date in the serial number of each radiosonde package. For the RS80-H radiosondes, there are two formats of the serial number. For packages manufactured prior to October 1995, the serial number is DDMMYTTPP, where:

DD = day of the month (1–31)
MM = month (1–12) + facility identifier (00, 20, 40, or 80)
Y = last digit of the year
TT = calibration tray identifier
PP = position in calibration tray

More recent radiosonde serial numbers are coded as YWWDTTTNN, where:

Y = last digit of the year

WW	=	week number (1–52)
D	=	day of the week (1–7) Monday = 1
TTT	=	calibration tray identifer
NN	=	position in calibration tray

REFERENCES

Antikainen, V., and A. Paukkunen, 1994: Studies on improving humidity measurements in radiosondes. *Proc. WMO Tech. Conf. on Instruments and Methods of Observation,* Geneva, Switzerland, WMO, WMO Tech. Doc. 588, 137–141.

Barr, A. G., and A. K. Betts, 1997: Radiosonde boundary layer budgets above a boreal forest. *J. Geophys. Res.,* **102,** 29 205–29 212.

Clough, S. A., and M. J. Iacono, 1995: Line-by-line calculations of atmospheric fluxes and cooling rates: Application to carbon dioxide, ozone, methane, nitrous oxide, and the halocarbons. *J. Geophys. Res.,* **100,** 16 519–16 535.

——, ——, and J. L. Moncet, 1992: Line-by-line calculations of atmospheric fluxes and cooling rates: Application to water vapor. *J. Geophys. Res.,* **97,** 15 761–15 785.

——, P. D. Brown, J. C. Liljegren, T. R. Shippert, D. D. Turner, R. O. Knuteson, H. E. Revercomb, and W. L. Smith, 1996: Implications for atmospheric state specification from the AERI/LBLRTM Quality Measurement Experiment and the MWR/LBLRTM Quality Measurement Experiment. *Proc. Sixth Atmospheric Radiation Measurement (ARM) Science Team Meeting,* San Antonio, TX, U.S. Dept. of Energy. [Available online at http://www.arm.gov/pub/docs/documents/technical/conf_9603/cloug_96.pdf.]

——, ——, D. D. Turner, T. R. Shippert, J. C. Liljegren, D. C. Tobin, H. E. Revercomb, and R. O. Knuteson, 1999: Effect on the calculated spectral surface radiances due to MWR scaling of sonde water vapor profiles. *Proc. Ninth Atmospheric Radiation Measurement (ARM) Science Team Meeting,* San Antonio, TX, U.S. Dept. of Energy. [Available online at http://www.arm.gov/pub/docs/documents/technical/conf_9903/clough-99.pdf.]

Cole, H., and E. R. Miller, 1995: A correction for low-level radiosonde temperature and relative humidity measurements. Preprints, *Ninth Symp. on Meteorological Observation and Instrumentation,* Charlotte, NC, Amer. Meteor. Soc., 32–36.

Elliott, W. P., and D. J. Gaffen, 1991: On the utility of radiosonde humidity archives for climate studies. *Bull. Amer. Meteor. Soc.,* **72,** 1507–1520.

Fleming, R., 1998: A note on temperature and relative humidity corrections for humidity sensors. *J. Atmos. Oceanic Technol.,* **15,** 1511–1515.

Guichard, F., D. Parsons, and E. Miller, 2000: Thermodynamic and radiative impact of the correction of sounding humidity bias in the Tropics. *J. Climate,* **13,** 3615–3624.

Han, Y., J. B. Snider, E. R. Westwater, S. H. Melfi, and R. A. Ferrare, 1994: Observations of water vapor by ground-based microwave radiometers and Raman lidar. *J. Geophys. Res.,* **99,** 18 695–18 702.

Lesht, B. M., 1995: An evaluation of ARM radiosonde operational performance. Preprints, *Ninth Symp. on Meteorological Observations and Instrumentation,* Charlotte, NC, Amer. Meteor. Soc., 6–10.

——, 1998: Uncertainty in radiosonde measurements of temperature and relative humidity estimated from dual-sonde soundings made during the September 1996 ARM water vapor IOP. Preprints, *10th Symp. on Meteorological Observations and Instrumentation,* Phoenix, AZ, Amer. Meteor. Soc., 80–85.

——, 1999: Reanalysis of radiosonde data from the 1996 and 1997 water vapor intensive observation periods: Application of the Vaisala RS-80H contamination correction algorithm to dual-sonde soundings. *Proc. Ninth Atmospheric Radiation Measurements (ARM) Science Team Meeting,* San Antonio, TX, U.S. Dept. of Energy. [Available online at http://www.arm.gov/pub/docs/documents/technical/conf_9903/lesht-99.pdf.]

——, and J. C. Liljegren, 1997: Comparison of precipitable water vapor measurements obtained by microwave radiometers and radiosondes at the Southern Great Plains Cloud and Radiation Testbed site. *Proc. Sixth Atmospheric Radiation Measurements (ARM) Science Team Meeting,* San Antonio, TX, U.S. Dept. of Energy, CONF-9603149, 59–62. [Available online at http://www.arm.gov/docs/documents/technical/conf_9603/lesh_96.pdf.]

Liebe, H. J., and D. H. Layton, 1987: Millimeter-wave properties of the atmosphere: Laboratory studies and propagation modeling. NTIA Rep. 87–24, National Telecommunications and Information Administration, Boulder, CO, 74 pp.

Liljegren, J. C., 2000: Automatic self-calibration of the ARM microwave radiometers. *Microwave Radiometry and Remote Sensing of the Environment,* P. Pampaloni, Ed., VSP Press, 433–441. [Available online at http://www.arm.gov/docs/instruments/publications/mwr_calibration.pdf.]

——, and B. M. Lesht, 1996: Measurements of integrated water vapor and cloud liquid water from microwave radiometers at the DOE ARM cloud and radiation testbed in the U.S. Southern Great Plains. *Proc. Int. Geoscience and Remote Sensing Symp. (IGARSS),* Lincoln, NE, IEEE, 1675–1677.

Miloshevich, L. M., H. Voemel, A. Paukkunen, A. J. Heymsfield, and S. J. Oltmans, 2001: Characterization and correction of relative humidity measurements from Vaisala RS80-A radiosondes at cold temperatures. *J. Atmos. Oceanic Technol.,* **18,** 135–156.

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

Revercomb, H. E., cited 2000: Science plan for the 3rd ARM Water Vapor IOP and the ARM-FIRE Water Vapor Experiment (AFWEX) at the ARM SGP CART site, Lamont, OK. [Available online at http://www.arm.gov/docs/iops/2000/sgp2000afwex/afwex.html.]

——, F. A. Best, R. G. Dedecker, T. P. Dirkx, R. A. Herbsleb, R. O. Knuteson, J. F. Short, and W. L. Smith, 1993: Atmospheric Emitted Radiance Interferometer (AERI) for ARM. Preprints, *Fourth Symp. on Global Climate Change Studies,* Anaheim, CA, Amer. Meteor. Soc., 46–49.

——, and Coauthors, 2003: The Atmospheric Radiation Measurement (ARM) Program's water vapor intensive observation periods: Overview, initial accomplishments, and future challenges. *Bull. Amer. Meteor. Soc.,* in press.

Richardson, S. J., M. E. Splitt, and B. M. Lesht, 2000: Enhancement of ARM surface meteorological observations during the Fall 1996 Water Vapor Intensive Observation Period. *J. Atmos. Oceanic Technol.,* **17,** 312–322.

Soden, B. J., and J. R. Lanzante, 1996: An assessment of satellite and radiosonde climatologies of upper-tropospheric water vapor. *J. Climate,* **9,** 1235–1250.

Stokes, G. M., and S. E. Schwartz, 1994: The Atmospheric Radiation Measurement (ARM) program: Programmatic background and design of the cloud and Radiation Test bed. *Bull. Amer. Meteor. Soc.,* **75,** 1201–1221.

Tobin, D. C., H. E. Revercomb, and D. D. Turner, 2002: Overview of the ARM-FIRE Water Vapor Experiment (AFWEX). *Proc. 12th Atmospheric Radiation Measurement (ARM) Science Team Meeting,* St. Petersburg, FL, U.S. Dept. of Energy.

Turner, D. D., and J. E. M. Goldsmith, 1999: Twenty-four-hour Raman lidar water vapor measurements during the Atmospheric Radiation Measurement program's 1996 and 1997 water vapor intensive operating periods. *J. Atmos. Oceanic Technol.,* **16,** 1062–1076.

——, T. R. Shippert, P. D. Brown, S. A. Clough, R. O. Knuteson, H. E. Revercomb, and W. L. Smith, 1998: Long-term analysis of observed and line-by-line calculations of longwave surface spectral radiance and the effect of scaling the water vapor profiles. *Proc. Eighth Atmospheric Radiation Measurement (ARM) Science Team Meeting*, Tucson, AZ, U.S. Dept. of Energy. [Available online at http://www.arm.gov/pub/docs/documents/technical/conf_9803/turner-98.pdf.]

Vaisala, 1996: RS80 Radiosondes information release. Engineering Release R625en 1996–06, Vaisala, OY, Helsinki, Finland, 12 pp.

Wang, J., H. Cole, D. J. Carlson, E. R. Miller, K. Beierle, A. Paukkunen, and T. K. Laine, 2002: Corrections of humidity measurement errors from the Vaisala RS80 radiosonde—Application to TOGA_COARE data. *J. Atmos. Oceanic Technol.*, **19,** 981–1002.

Weckworth, T. M., V. Wulfmeyer, R. M. Wakimoto, R. M. Hardesty, J. W. Wilson, and R. M. Banta, 1999: NCAR/NOAA lower tropospheric water vapor workshop. *Bull. Amer. Meteor. Soc.*, **80,** 2339–2357.

Zipser, E. J., and R. H. Johnson, 1998: Systematic errors in radiosonde humidities: A global problem? Preprints, *10th Symp. on Measurements, Observations, and Instrumentation,* Phoenix, AZ, Amer. Meteor. Soc., 72–73.