Radiosonde Humidity Soundings and Microwave Radiometers during Nauru99

EDGEWORTH R. WESTWATER,* B. BOBA STANKOV,+ DOMENICO CIMINI,# YONG HAN,@ JOSEPH A. SHAW,& BARRY M. LESHT,** AND CARLES N. LONG++

*Cooperative Institute for Research in the Environmental Sciences, University of Colorado, and NOAA/Environmental Technology Laboratory, Boulder, Colorado
+NOAA/Environmental Technology Laboratory, Boulder, Colorado
#CETEMPS, University of L'Aquila, L'Aquila, Italy
@NOAA/National Environmental Satellite and Data Information Service, Camp Springs, Maryland
&Montana State University, Bozeman, Montana
**Argonne National Laboratory, Argonne, Illinois
++Pacific Northwest National Laboratory, Richland, Washington

(Manuscript received 12 August 2002, in final form 6 December 2002)

ABSTRACT

During June–July 1999, the NOAA R/V *Ron H. Brown* (*RHB*) sailed from Australia to the Republic of Nauru where the Department of Energy's Atmospheric Radiation Measurement (ARM) Program operates a long-term climate observing station. During July, when the *RHB* was in close proximity to the island of Nauru, detailed comparisons of ship- and island-based instruments were possible. Essentially identical instruments were operated from the ship and the island's Atmospheric Radiation and Cloud Station (ARCS)-2. These instruments included simultaneously launched Vaisala RS80-H radiosondes, the Environmental Technology Laboratory's (ETL) Fourier transform infrared radiometer (FTIR), and ARM's atmospheric emitted radiance interferometer (AERI), as well as cloud radars/ceilometers to identify clear conditions.

The ARM microwave radiometer (MWR) operating on Nauru provided another excellent dataset for the entire Nauru99 experiment. The calibration accuracy was verified by a liquid nitrogen blackbody target experiment and by consistent high quality tipping calibrations throughout the experiment. Comparisons were made for calculated clear-sky brightness temperature (T_b) and for precipitable water vapor (PWV). These results indicate that substantial errors, sometimes of the order of 20% in PWV, occurred with the original radiosondes. When a Vaisala correction algorithm was applied, calculated T_bs were in better agreement with the MWR than were the calculations based on the original data. However, the improvement in T_b comparisons was noticeably different for different radiosonde lots and was not a monotonic function of radiosonde age. Three different absorption algorithms were compared: Liebe and Layton, Liebe et al., and Rosenkranz. Using AERI spectral radiance observations as a comparison standard, scaling of radiosondes by MWR data was compared with both original and corrected soundings.

1. Introduction

The Department of Energy's Atmospheric Radiation Measurement (ARM) Program (Stokes and Schwartz 1994) has established long-term Atmospheric Radiation and Cloud Stations (ARCSs) in the tropical western Pacific (TWP) at Manus Island, Papua New Guinea, and the Republic of Nauru. The concentration of a wide variety of instruments at these sites allows for the continuous measurement of upwelling and downwelling radiative fluxes, as well as important meteorological parameters that affect radiative balance. The sites on Manus Island (ARCS-1) and Nauru (ARCS-2) are intended to characterize climate and radiation processes in the ''heat engine'' of the TWP (Webster and Lucas 1992).

The Nauru99 campaign was organized to see if island measurements are representative of the surrounding ocean (Post and Fairall 2000), and used a variety of in situ and remote sensors deployed on the R/V *Ron H. Brown* (*RHB*) and the Japanese R/V *Mirai.* As discussed in section 2, a variety of instruments, both on the *RHB,* and on ARCS-2, were available to verify instrument performance and site characterization for the surrounding ocean. In our subset of Nauru99 activities, we were particularly interested in verifying, and perhaps improving, both microwave and infrared radiative transfer models (RTMs) for clear air conditions. In such forward model development, one starts with a radiative transfer equation (RTE) model that is based on fundamental physics (Clough et al. 1989; Liebe 1989). Then, after supplying the model with measured profiles of meteo-

Corresponding author address: Dr. Edgeworth R. Westwater, Cooperative Institute for Research in the Environmental Sciences, University of Colorado/NOAA/Environmental Technology Laboratory, 325 Broadway, Boulder, CO 80305-3328.
E-mail: ed.r.westwater@noaa.gov

rological variables such as pressure, temperature, and water vapor density, spectral and/or angular radiance is calculated. The radiance calculations are then compared with observations from a calibrated radiometer (Westwater et al. 1990; Han et al. 1997). If the measurements and calculations agree to within the uncertainty of the radiometer measurements, confidence is obtained that the radiometric data can be used for remote sensing or climate studies. However, if well-calibrated measurements agree poorly with calculations, and if the state of the atmosphere is correctly specified, then adjustments to the RTE model are usually required (Fleming et al. 1991). For example, adjustments to the parameterization of the water vapor continuum have been made on this basis (Han et al. 1997; Westwater et al. 1999; Tobin et al. 1999). However, accurate measurements of temperature and humidity profiles are necessary for the development and verification of RTE models.

Although it is generally realized that spatial and temporal differences limit the accuracy of radiosonde and radiometer comparisons, an implicit assumption is often made that radiosonde temperature and humidity sensors provide an unbiased measure of the true state of the atmosphere. An early example of a possible bias was the reporting procedure of National Weather Service (NWS) humidity soundings that gave a default dewpoint depression of -30°C when the relative humidity was below 20% (Wade 1994). Such a bias was noted earlier in comparisons of microwave humidity soundings with radiosondes (Westwater et al. 1989). Several articles have also addressed the problem of radiosonde biases in humidity. Wade (1994), in addition to recommending that the practice of reporting default dewpoint depressions be modified, noted that the commonly used carbon hygristor has problems below 20% RH. A recent article by Turner et al. (2003) also discusses the dry bias problem and its correction for midlatitude data. And finally, the problem of dry bias of Vaisala humidity soundings in a tropical environment has been carefully discussed by Guichard et al. (2000) and Wang et al. (2002).

Additional experience, both in the tropical western Pacific (Westwater et al. 1999) and during several water vapor intensive operating periods at the ARM Southern Great Plains Cloud and Radiation Testbed (CART) site (Lesht 1995, 1999; Turner et al. 2003), has indicated the need for adjustments to Vaisala Humicap RS80 humidity soundings. The need for such corrections has been identified by measurements of water vapor by microwaves radiometers (MWRs) and with subsequent adjustments by scaling of water vapor profiles (Clough et al. 1996, 1999; Turner et al. 2003). Furthermore, well-calibrated infrared radiance measurements have been taken by the atmospheric emitted radiance interferometer (AERI) operated at many ARM CART sites (Revercomb et al. 1993). The scaled water-vapor profiles, when inserted into the line-by-line (LBL) RTM (Clough et al. 1989, 1992), have produced radiance calculations that are in much better agreement with AERI measurements than are calculations based on the original radiosonde data (Clough et al. 1999; Turner et al. 2003). In addition to producing better agreement on the average than the original radiosonde data, the residuals, as a function of precipitable water vapor (PWV), are also significantly reduced by the scaling. In this paper, we evaluate scaling and correction of radiosonde humidity profiles by a Vaisala correction algorithm and by quantities derived from microwave radiometers.

2. Instruments

During Nauru99 (Post and Fairall 2000), a variety of ground- and ship-based instruments were available to test the quality of remote and in situ soundings in the tropical environment around the island of Nauru (0.521°S, 166.916°E). In this paper, the *RHB* instruments of primary use were Vaisala radiosondes, a cloud ceilometer, a 35-GHz cloud radar, and a Fourier transform infrared radiometer (FTIR) (Han et al. 1997). Similar instruments were operated by ARM at the ARCS-2 site that had, in addition, two types of ceilometers, a dual-frequency MWR, and the counterpart to the Environmental Technology Laboratory's (ETL) FTIR—the AERI. The operation of essentially duplicate instruments in close proximity allowed us to examine the quality and consistency of a variety of data and retrieved products. We describe in this section the salient characteristics of the instruments used in our analysis.

a. Radiosondes

A variety of radiosonde soundings from the ARCS-2 and the *RHB*, many within 15 min of each other, as well as limited number of soundings from the Japanese R/V *Mirai* (Yoneyama 2000), were available. All of the radiosondes were manufactured by Vaisala and contained RS80-H Humicap humidity sensors. In contrast to the earlier Pilot Radiation Observation Experiment (PROBE) (Westwater et al. 1999), the identification numbers of the Vaisala radiosondes were available for subsequent analysis and categorization. As is now well known, the Vaisala RS80 radiosondes suffer from the dry bias problem due to contamination of the humidity element by its protective packaging (Miller et al. 1999; Lesht 1999; Guichard et al. 2000; Wang et al. 2002). Vaisala, together with the National Center for Atmospheric Research (NCAR) and ARM, has developed an algorithm to correct for this bias. The radiosonde identification number, which identifies the age at which the radiosonde was manufactured, is the basis for the corrections used here (Lesht 1999).

b. The ARCS-2 microwave radiometer

The ARM program operates dual-frequency MWRs at all of their CART sites, including the two ARCS sites in the TWP. The ARM MWRs are commercial units that

were purchased from Radiometrics Inc., in Boulder, Colorado. These instruments operate at 23.8 and 31.4 GHz; their description and operational calibration algorithm (tipcal) are given by Liljegren (2000) (see also information available online at http://www.arm.gov/docs/instruments/static/mwr.html). Many of our comparisons with radiosondes rely on the accuracy and consistency of the ARCS-2 MWR whose basic measurement, after calibration, is the brightness temperature (T_b). During this experiment, the radiometer was run in a discrete scanning mode in which the radiometer scans in a plane at air masses 2.5, 2.0, 1.5, and 1.0 on each side of zenith with two measurements at zenith (elevation angles of 23.6°, 30.0°, 41.8°, 90°). During this scan, about 6 s were spent at each angle, with the total sequence requiring 50 s. When the sky conditions were favorable, as determined by symmetry of radiometer scans, the radiometer continued scanning. When clouds were present, angular symmetry was destroyed and the radiometer went into a zenith-observing mode. Because we cannot accurately calculate T_b from radiosonde data during conditions in which liquid-bearing clouds are present, we will focus on clear air conditions only; another reason for choosing clear conditions is that during these conditions, tip calibration data are available every 50 s.

The original ARM calibration algorithm (Liljegren 2000) was used on the data to produce ARM's so-called line of sight (LOS) data. We also applied the ETL calibration method (Han and Westwater 2000) to the same tipcal data and used 10-min samples of data to compare the two methods. The ARM and the ETL procedures have many common features, including corrections for antenna beamwidth, effects of earth curvature and atmospheric refractive index, estimation of mean radiating temperature depending on the season, and a method for accessing the quality of each individual tip curve calibration. Moreover, the two procedures use the same radiometric equation:

$$T_b = T_{\text{Ref}} + \Gamma(V_{\text{Sky}} - V_{\text{Ref}}), \tag{1}$$

where T_b and T_{Ref} are, respectively, the sky equivalent blackbody brightness temperature and the measured temperature of the reference target; V_{Sky} and V_{Ref} are, respectively, the signals measured when the radiometer is looking the sky and the reference target; while Γ is a multiplicative factor. Calling $V_{\text{Ref+nd}}$ the signal measured during reference target observation coupled with injection from a noise diode, T_{nd} the noise diode equivalent brightness temperature, and f_W the polycarbonate foam window loss factor, we find that

$$\Gamma = f_W T_{\text{nd}}/(V_{\text{Ref+nd}} - V_{\text{Ref}}) = f_W G. \tag{2}$$

The most significant difference between the two procedures is the method of deriving the calibration coefficient from the tip curve data. As described in Liljegren (2000), the ARM operational algorithm estimates the value of T_{nd} from each set of measurements taken at eight different elevation angels (ranging from 1 to 2.5 air masses at Nauru). If the tip curve is judged to be valid, then its zenith T_b estimate is used to calculate a corresponding estimate of the noise injection temperature T_{nd}. Once a sufficient number of valid T_{nd}s is accumulated, a robust regression (minimizing the absolute deviation) of T_{nd} on the internal reference T_{Ref} is performed. To determine a linear relationship between T_{nd} and T_{Ref}, this procedure retains a long time history (between 1500 and 3000 samples) of these parameters. Every time a new, quality-tested, couple of T_{nd} and T_{Ref} is measured and stored, the oldest is discharged and the linear regression coefficients are updated. Thus, the radiometer output is calibrated according to the most recent regression coefficients and the instantaneous value of T_{Ref}. Note that, because each tip curve takes about 50 s, the collection of 1500 samples requires more than 24 h. Conversely, the ETL procedure determines the gain correction $f_W T_{\text{nd}}$ by applying the tip curve calibration to each set of observations taken at eight different elevation angles, and recalibrates the same observations according to the new value of $f_W T_{\text{nd}}$. Therefore, the ETL procedure is also called ''instantaneous'' calibration. Note that f_W is only a dimensionless constant factor very close to one ($\sim$1.002).

For each tipcal, the ETL method uses angular-dependent mean radiating temperatures $T_{\text{mr}}(\theta)$ and equivalent zenith brightness temperatures (EZTBs) as a measure of calibration quality. To compute EZTB, one takes a T_b measurement at any elevation angle θ, $T_b = T_b(\theta)$, and, using an estimated $T_{\text{mr}}(\theta)$, derives the optical depth $\tau(\theta)$ by

$$\tau(\theta) = \ln\left[\frac{T_{\text{mr}}(\theta) - T_c}{T_{\text{mr}}(\theta) - T_b(\theta)}\right]. \tag{3}$$

In (3), T_c is the cosmic big bang brightness temperature of 2.75 K. After using (3) to derive $\tau(\theta)$, we get $\hat{\tau}(\pi/2) = \tau(\theta)/m(\theta)$, where $m(\theta) = \csc(\theta)$. Finally, we get EZTB $\equiv \hat{T}_b(\pi/2)$ from

$$\hat{T}_b(\pi/2) = T_c e^{-\hat{\tau}(\pi/2)} + T_{\text{mr}}(\pi/2)[1 - e^{-\hat{\tau}(\pi/2)}]. \tag{4}$$

The ETL method of deriving 10-min-averaged data is to average within the prescribed time interval only those data that have passed a EZTB threshold criterion. We mention that the ETL procedure was designed for clear conditions only, and would have to be modified when there are long intervals with clouds. On the other hand, the ARM procedure was designed for being deployed during both clear and cloudy conditions, regardless of the applicability of tip curve calibration.

c. The ETL FTIR and the ARM AERI

1) FTIR

The FTIR developed by ETL (Shaw et al. 1991) is a simple, ground-based Fourier transform spectroradiometer. It is based on a compact and rugged commercially purchased Michelson interferometer, operating

between roughly 500 and 2000 cm^{-1} (5.0–20.0 μm) with 1 cm^{-1} spectral resolution. Downwelling atmospheric emission in the entire bandwidth is viewed simultaneously with a single liquid nitrogen (LN2)-cooled HgCdTe detector. A fast Fourier transform of the measured interferogram then yields the emission power spectrum. The FTIR views two blackbody calibration targets immediately before and after each complete atmospheric emission measurement. Collection of each spectrum takes about 1 s, and we average 100 such spectra to reduce random noise. With this technique, calibrated atmospheric spectra are collected about once every 6–10 min.

2) AERI

The AERI is deployed widely by ARM at many of its CART sites. It measures the absolute infrared spectral radiance of the sky directly above the instrument with a spectral measurement range of 500–3300 cm^{-1} or 20–3 μm. The spectral resolution is 1.0 cm^{-1} with an instrument field of view of 1.3°. A calibrated sky radiance spectrum is produced every 7 min. The AERI data can be used for 1) evaluation of line-by-line radiative transport codes (Tobin et al. 1999), 2) detection and quantification of cloud effects on ground-based measurements of infrared spectral radiance (Deslover et al. 1999), and 3) calculation of vertical atmospheric profiles of temperature and water vapor (Feltz et al. 1998). (Further information on the instrument can be found online at the Web site http://www.arm.gov/docs/instruments/static/aeri.html.) In this paper, we use the two infrared spectroradiometers as tools to examine the accuracy of original and corrected radiosonde data. More precisely, after verifying anecdotically that the two instruments agreed during cloud-free conditions, we used the AERI measurements on the island to compare with calculations based on line-by-line radiative transfer codes and applied to the atmospheric profile data.

d. Millimeter wavelength cloud radar (MMCR)

The MMCR is a zenith-pointing radar that operates at a frequency of 35 GHz. The main purpose of this radar is to determine cloud boundaries (e.g., cloud bottoms and tops), although this radar will also report radar reflectivity (dBZ) to 20 km (Clothiaux et al. 2000). (A more complete description of the MMCR can be found online at http://www.arm.gov/docs/instruments.html#Cloud.)

e. Vaisala ceilometer (Model CT25K)

The Vaisala ceilometer (VCEIL) is a self-contained, ground-based, active, remote sensing device designed to measure cloud-base height at up to three levels. Model CT25K has a maximum vertical range of 25 000 ft. The same model was used by ETL on the R/V *Ron H. Brown* and at ARCS-2 (information on the ARCS-2 unit is found online at http://www.arm.gov/docs/instruments.html#Cloud). In this paper, we use the combined cloud–boundary product of ARM to screen our data from clouds. This product will identify clouds from the surface to 20 km (Clothiaux et al. 2000).

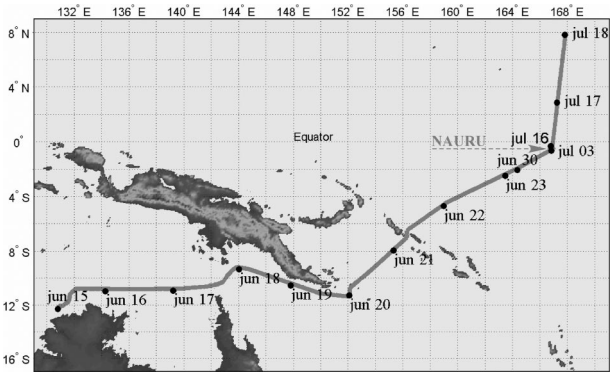

FIG. 1. Ship track of the R/V *Ron H. Brown* from Darwin, Australia, to the Republic of Nauru and on to Kwajalein during the Nauru99 experiment. The ship was in the vicinity of Nauru from 3–15 Jul 1999.

3. Initial observations from the R/V *Ron H. Brown* and ARCS-2

As shown in Fig. 1, the *RHB* departed Darwin, Australia, on 15 June 1999, and arrived at Nauru on 3 July 1999, after spending 7 days positioned near a Tropical Atmosphere–Ocean (TAO) buoy about 40 km away from the island. On arrival, the ship assumed a stationary position approximately 1 km away from the ARCS-2 location. One of the first data comparisons of interest was that of simultaneously launched radiosondes. A substantial difference between ARCS-2 and *RHB* radiosondes was observed on the very first day that the ship was in close proximity to the island. Figure 2 shows a time series of 23.8- and 31.4-GHz brightness temperatures (T_bs) observed by the MWR at the ARCS-2 site, and T_bs calculated from various radiosonde profiles using the latest Rosenkranz (1998) absorption model (ROS). Somewhat surprisingly, the original radiosonde data from the *RHB* agreed much better with the MWR data than those from the collocated ARCS-2 soundings. The open circles and triangles show data that were calculated from ARCS-2 and RHB soundings that were corrected using the procedure desscribed by Lesht (1999).

The results of Fig. 2 showed that, at the very least, there were significant differences between radiosondes that were launched at nearly the same time from the *RHB* and from the ARCS-2. In a blind test a change in the experimental plan was made, and for five soundings radiosonde packages from the two sites were interchanged. This time, the original *RHB* radiosondes, now

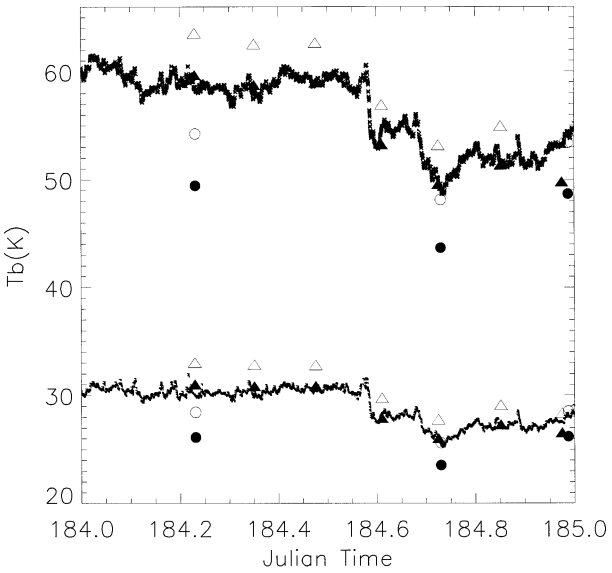

FIG. 2. The 24-h time series of T_b at 23.8 (*) and 31.4 GHz (+) on 3 Jul 1999 calculated from original ARCS-2 radiosondes (age = 315 days) (closed circles), corrected ARCS-2 radiosondes (open circles), original *RHB* radiosondes (age = 172 days) (closed triangles), and corrected *RHB* radiosondes (open triangles). RTE model ROS.

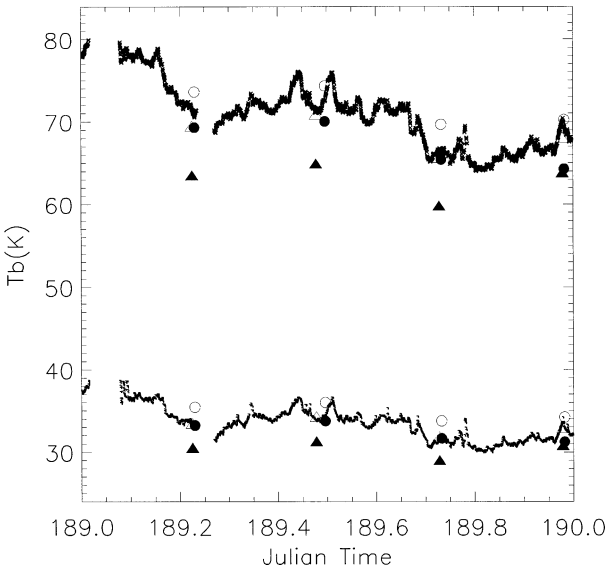

FIG. 3. The 24-h time series of T_b at 23.8 (*) and 31.4 (+) GHz on 8 Jul 1999 calculated from original ARCS-2 radiosondes (age = 177 days) (closed circles), corrected ARCS-2 radiosondes (open circles), original *RHB* radiosondes (age = 407 days) (closed triangles), and corrected RHB radiosondes (open triangles). For the final two radiosondes for the day, the exchange experiment was discontinued and the age of the ARCS-2 and *RHB* radiosondes were now 407 and 177 days. RTE model ROS.

launched from ARCS-2, were in close agreement with the MWR, while the original ARCS-2 radiosondes, now launched from the *RHB*, were in substantial disagreement with the MWR (see Fig. 3). Again, using the procedure of Lesht (1999), the corrected radiosondes were in good agreement with the MWR T_bs. The problem is now known to be associated with contamination of the RS80 humidity element as it ages. Thus, when the radiosondes were exchanged, the results were consistent with a radiosonde problem, not a site problem. In section 5, we present results evaluating the accuracy of the original radiosondes and their correction as a function of radiosonde age over the range of radiosonde ages deployed during the experiment (71–406 days old).

4. Data from the ARCS-2 microwave radiometer

As was shown in section 3, there were differences between measurements and calculations of T_b, sometimes as large as 20% from the original radiosonde data and by several degrees from the corrected radiosonde data. It was thus imperative to access and verify the calibration accuracy of the ARM MWR. First, we compared the ARM calibration algorithm (Liljegren 2000) with the ETL calibration method (Han and Westwater 2000) on the same tipcal data and used 10-min samples of data to compare the two methods. As discussed in section 2b, ETL uses EZTB as a measure of tipcal quality. The standard deviations (std dev) of this EZTB over the complete set of angles were frequently better than 0.3 K, indicating a high degree of atmospheric stratification and antenna beam symmetry. An example of a 24-h time series of std dev (EZTB) is shown in Fig. 4. Although there are times when clouds or departures from stratification caused std dev (EZTB) to be larger than 5 K, several periods have std dev (EZTB) less than 0.3 K std dev, as indicated by the solid line in Fig. 4.

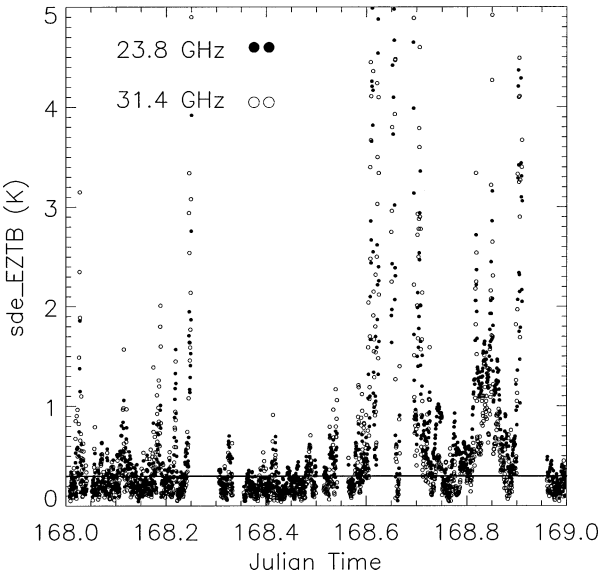

FIG. 4. The 24-h time series of the std dev of equivalent zenith T_b at 23.8 (solid circles) and 31.4 GHz (open circles) for Nauru99. The quality-control threshold of 0.3-K std dev is indicated with the solid line.

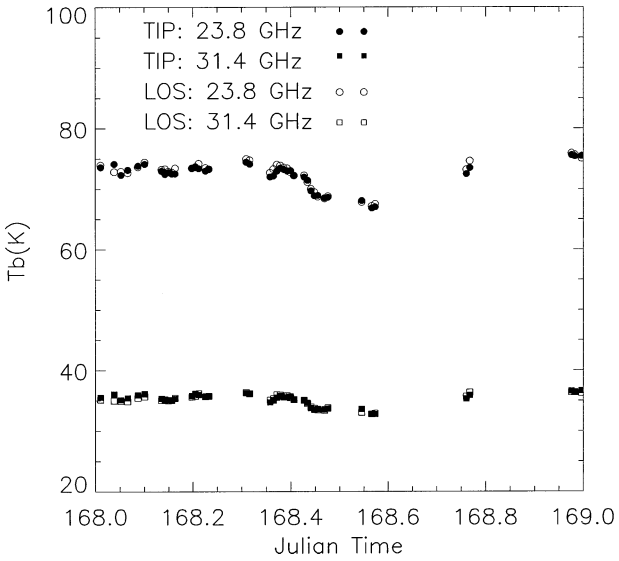

FIG. 5. The 24-h time series of the 10-min-averaged T_b at 23.8 and 31.4 GHz for Nauru99. The ETL TIP data were determined from the tipcal method of Han and Westwater (2000) and the ARM LOS data from the tipcal method of Liljegren (2000). Only data that passed the quality-control threshold of 0.3-K std dev are shown.

When we restrict the MWR data to those passing the 0.3 K of std dev (EZTB) quality-control threshold, both the original ARM-processed data (LOS) and the ETL-processed data are in, at least, qualitative agreement (see Fig. 5). To quantify this agreement, we also present in Table 1 the statistical comparisons of the two methods for the entire dataset. From this table we see that the rms difference between the two methods is less than 0.54 K. In the remainder of this paper, when studying clear periods for the MWR, we used ARM LOS data whose inferred liquid water path was less than 0.005 cm. Within the noise levels of Table 1, the restriction to the 0.005-cm value would yield results close to the results using only data that have passed the 0.3-K std dev threshold with use of the ETL processing.

As further evidence that we were working with a well-calibrated MWR, we also performed an LN2 calibration experiment, in which a blackbody reference target (or load) was filled with LN2 and placed over the MWR. The measured T_bs during this experiment are shown in Fig. 6. For the first 2 min after the load was inserted, the measured T_bs were about 79.6 K, which is close to the expected value of 79.4 $\pm$ 1.9 K. After the first 2 min, moisture condensed on the underside of the Styrofoam container and increasingly spurious observations were obtained. However, the few minutes of good measurements indicated that the MWR was accurate to within ± 1 K. This single target calibration measurement, together with the continuous high quality of tipcals, indicated that the MWR could be used as a comparison standard for the experiment.

TABLE 1. Comparison of ARM vs ETL tip calibration results for 10-min-averaging period. Threshold of ETL quality control parameter std dev EZTB = 0.3 K. Sample size is 782. Nine-point median filter applied to 50-s data to eliminate outliers.

	Avg (K) 23.8 GHz	Std dev (K) 23.8 GHz	Avg (K) 31.4 GHz	Std dev (K) 31.4 GHz
ARM LOS	65.12	8.13	32.12	3.08
ETL Tip	64.82	8.06	32.00	3.06
Tip − LOS	−0.29	0.45	−0.12	0.24

5. The Vaisala humidity correction algorithm vs MWR data

Comparisons of Vaisala radiosondes with remote sensors at the Department of Energy's Southern Great Plains site in central Oklahoma, as well as various types of intercomparisons between the radiosondes themselves, have indicated significant problems with the relative humidity measurements (Lesht 1999; Lesht and Liljegren 1996; Turner et al. 2003). Similar problems were observed during the Tropical Ocean and Global Atmosphere (TOGA) Coupled Ocean–Atmosphere Response Experiment (COARE) (Miller et al. 1999; Wang et al. 2002; see also online at http://www.atd.ucar.edu/dir_off/tc_corr/radiosonde/index.html). This problem has been shown to arise from contamination of the humidity element during storage. The manufacturers of Vaisala radiosondes, together with NCAR, developed a proprietary algorithm to correct for the dry bias problem (Miller et al. 1999; Lesht 1999). Subsequently, Wang et al. (2002) developed several non proprietary algorithms to correct for this bias. In our study of the Nauru MWR data, we used a version of the original Vaisala

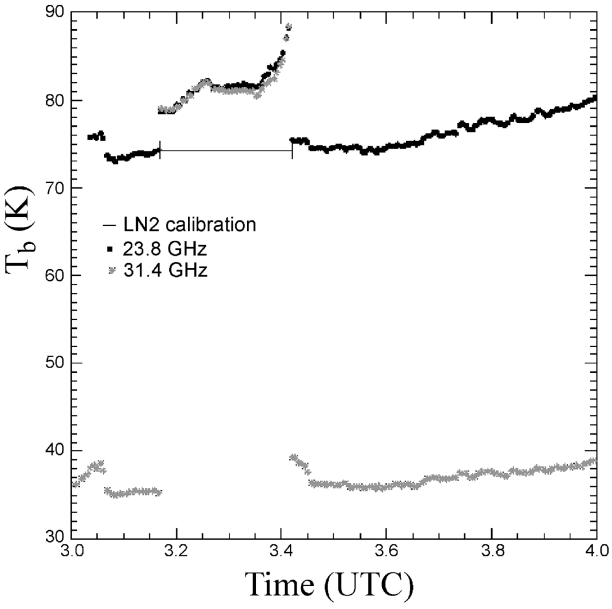

FIG. 6. Results of the LN2 calibration experiment. A blackbody reference target (or load) was placed in the bottom of a Styrofoam container, which was filled with LN2 and placed over the MWR.

algorithm that makes the correction based only on the age of the radiosonde. Examples based on the application of this algorithm are shown in Fig. 7, which shows radiosondes that are about ½ yr old (Figs. 7a and 7b) and more than 1 yr old (Figs. 7c and 7d). Note how the ½-yr-old sounding (age = 178 days) shown in Figs. 7a and 7b is changed by about 6% by the correction, while the 406-day-old sounding, shown in Figs. 7c and 7d, is changed by about 10%. Figure 7 shows both ARCS-2 and *RHB* radiosondes after correction has been applied. The differences still remain but they can be explained by Fig. 8, which shows the ascent of the two radiosondes and their trace at the ground. The ARCS-2 radiosonde ascended faster in the area of westerly winds near the ground and reached the easterly wind regime above ground sooner than the *RHB* radiosonde. Because of that occurrence, the *RHB* radiosonde was carried 10 km farther away from the island.

Examples of T_b calculations using this correlation algorithm were shown in Figs. 2 and 3. As is evident for these two figures, the 23.8-GHz channel is more sensitive to water vapor than the 31.4-GHz channel. The correction algorithm removes, and perhaps slightly overcorrects, some of the dry bias with respect to the MWR data. Because we were worried about diurnal effects, we divided our data samples into day and night subsets. Then for each subset we compared T_bs measured by the ARM MWR with T_b calculations based on the Rosenkranz (1998) absorption model for both the original and corrected radiosondes. Figures 9 and 10 show scatterplots giving the day versus night comparisons of corrected and uncorrected T_b calculations for the period of 15 June–15 July 1999. Our analysis showed a modest diurnal effect, with the original radiosondes being drier during the day (about 1.3 K at 23.8 GHz and 0.6 K at 31.4 GHz). Although the tipcal EZTB data did not show any apparent diurnal tendencies, such effects on the MWR cannot be ruled out.

Because the performance of the algorithm as a function of radiosonde age was an important issue, we also plotted the differences between measurements and calculations as a function of radiosonde age. The surprising results, shown for ARCS-2 in Figs. 11a and 11b, and for the *RHB* in Figs. 11c and 11d, and in Tables 2 and 3, indicated that although the algorithm generally improved the differences, the improvement did not always occur for all radiosonde lots. In fact, for the radiosonde lots corresponding to age around 367–370 days, 19 radiosondes in all, the correction worsened the 23.8-GHz results by about 5 K. Except during the radiosonde exchange experiment, the radiosonde launched from the *RHB* were much more homogeneous in age than those from ARCS-2 and generally were newer (the *RHB* radiosondes were clustered around 180 days old). In the last four rows of Tables 2 and 3, we show the results from radiosondes that range from 174 to 182 days and those that range from 301 to 409 days. The original biases at 23.8 GHz increase by about 3 K with the older set, and the Vaisala correction does an excellent job of reducing these biases. In summary, the correction algorithm generally improves the fit with the MWR, but it certainly is not perfect.

6. Effects of microwave absorption model

In the previous sections, we have shown how the calculations of T_b that were based on the Rosenkranz (1998) absorption model, when using the corrected radiosonde data, were in better agreement with the MWR T_b data than were calculations based on the original radiosonde data. However, these results could be affected by the choice of absorption model, and although we believe that ROS is state of the art, other models should at least be considered to show the magnitude of the effect. There are many issues in the determination of parameters that enter into water vapor absorption modeling, and a clear discussion of several of these issues is given by Rosenkranz (1998). Relevant to the discussion here is the choice of parameters to calculate the pressure-broadened line width, which, in the case of water vapor, arises from the collisions of H_2O with other H_2O molecules (self-broadening), or from collisions of H_2O molecules with those of dry air (foreign broadening). In fact, Rosenkranz based his model on using Liebe and Layton's (1987) model (L87) values for the foreign-broadened component, and those from Liebe et al.'s (1993) model (L93) for the self-broadened component. We therefore examined the Rosenkranz model and the other two as well. These models can be directly applied to radiosonde data to calculate T_b at each of the two MWR frequencies. As a point of reference, we mention that the range of PWV varied from 2.5 to 6.1 cm, as measured by the original radiosondes, and from 2.8 to 6.4 cm, as measured by the corrected radiosondes. Scatterplots of the comparisons with the MWR data are shown in Figs. 12 and 13, for 23.8- and 31.4-GHz data, respectively. As is seen from these figures, L87 and ROS are in close agreement at 23.8 GHz, but L93 differs considerably from the other two by 3 K. If we assume that the Vaisala correction algorithm is correct, these comparisons indicate serious problems with L93. At 31.4 GHz, where the absorption is more dependent on the water vapor continuum [the long wavelength wings of self- and foreigned-broadened lines are greater than 22.235 GHz; see Ma and Tipping (1990)] than on the parameters of the 22.235-GHz absorption line, there is about 1-K difference in the bias between L87 and ROS. The corrected radiosonde data show that at 31.4 GHz the L87 results agree slightly better with the MWR than do the ROS results. However, the uncertainties in both original and corrected radiosondes, for the ages of radiosondes encountered here (generally greater than ½ year), preclude making any definite conclusions of L87 versus ROS. We also investigated the data to see if differences between measured and calculated T_b depended on PWV, either for

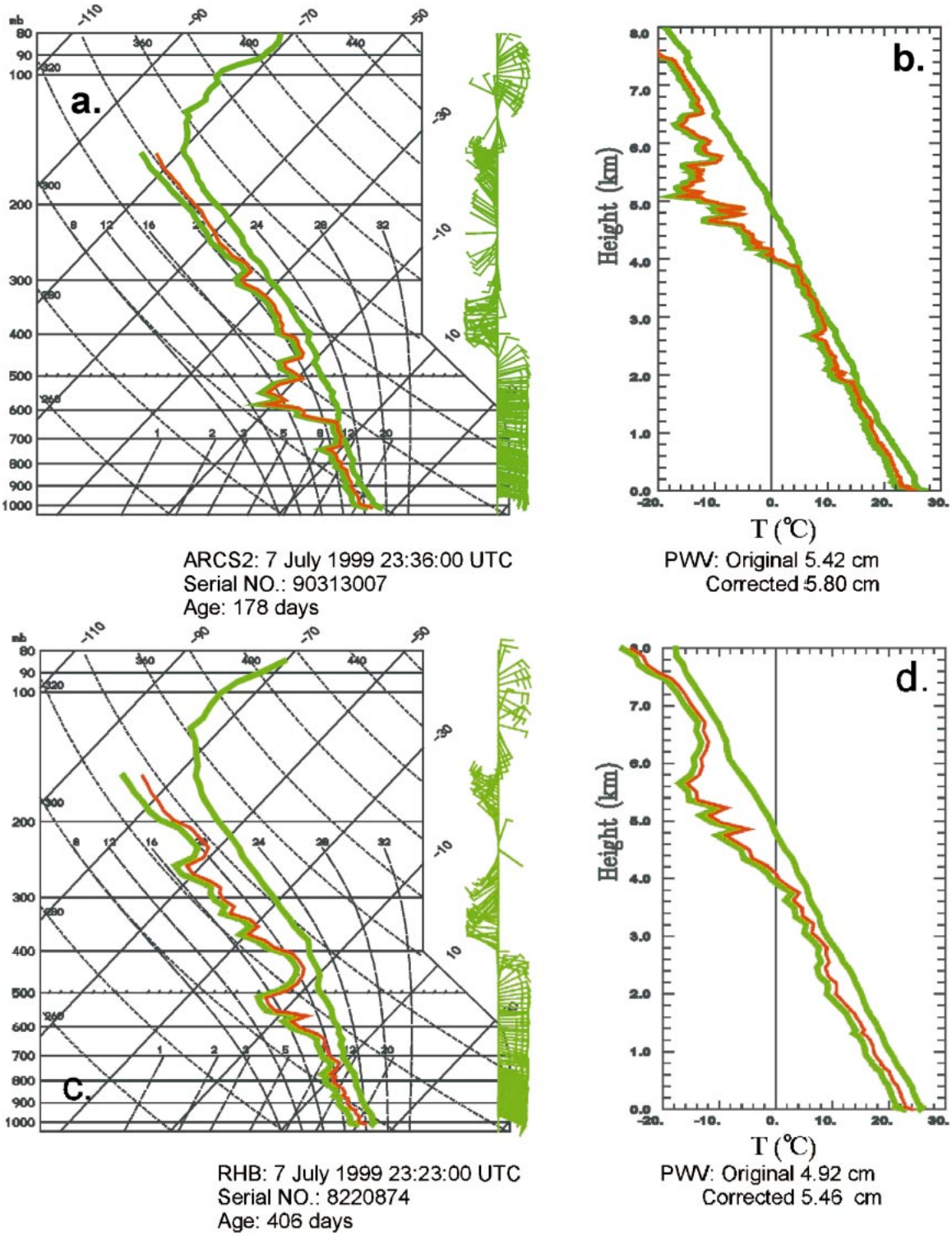

FIG. 7. ARCS-2 and *RHB* original (green) and corrected (red) radiosondes: (a) and (c) are skew *T*–log*p* profiles, (b) and (d) are temperature and dewpoint temperature vs height.

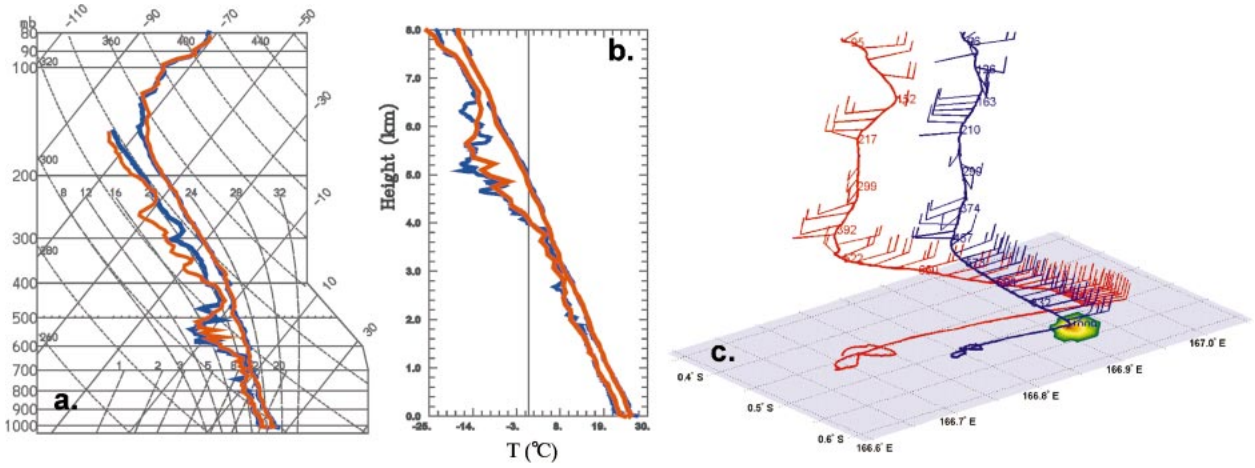

FIG. 8. ARCS-2 (blue) and *RHB* (red) corrected radiosondes: (a) skew T–log$p$ profile, (b) temperature and dewpoint temperature vs height, and (c) radiosonde trajectory with wind barbs and the trajectory trace at the ground.

the original or the corrected radiosondes. We calculated the correlation coefficient between measured minus calculated T_b versus PWV for all three models and for both original and corrected radiosondes. Neither for the absorption models nor for the original versus corrected radiosondes were there appreciable correlations: the correlation coefficients ranged from −0.10 to 0.10 over the set.

We also performed calculations for the set of radiosondes released from ARCS-2, the set released from the *Ron H. Brown*, and the composite containing both sets; the results are shown in Tables 4 and 5. We see immediately that at 23.8 GHz, there is close agreement between L87 and ROS, with only about 0.1 K of difference separating the two, both in the average difference and the standard deviation for both the original and corrected datasets. If we assume that the Vaisala correction algorithm is correct on the average, then ROS is slightly closer to the MWR than is L87. However, because of the large scatter in the results, std dev ≈ 3.5–4.0 K, L87 and ROS are statistically indistinguishable within the 99% confidence limits shown in Table 4. At 31.8 GHz, however, there is about a 1.0-K difference in the average difference, and, again, if we assume the Vaisala correction algorithm is correct on the average, then L87 agrees better with the MWR with 99% confidence limits.

We also display our comparisons of MWR data versus model calculations by histogram analysis; for brevity, we show only comparisons for the composite ARCS-2 plus *RHB* datasets. In Fig. 14, we show the comparisons for 23.8 GHz. We note that (a) the probability distributions and the median values for L87 and ROS are within about 0.1 K; (b) for all three models, the median values for the original and corrected radiosondes are quite different, about 5 K; (c) L93 is inconsistent with the other two models; and (d) for either corrected or uncorrected datasets, there is a large range (peak-to-peak differences of 20–25 K). Certainly, the large variance reduces the statistical confidence with which one could make definitive statements about the differences between L87 and ROS, but it is encouraging that the median values differ by only 0.1 K. For the 31.4-GHz comparisons, we present Fig. 15, which is consistent with Fig. 14, except for one major difference: the me-

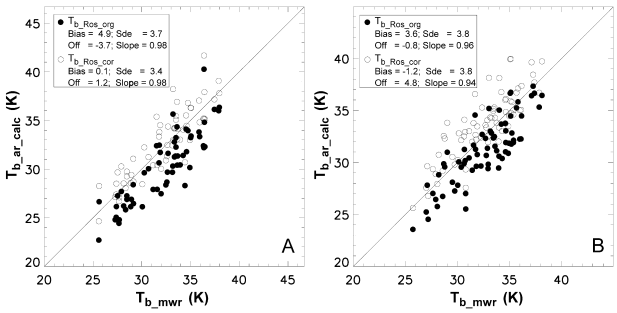

FIG. 9. Comparison of T_b at 23.8 GHz calculated from original (closed circles) and corrected (open circles) radiosondes vs ARM MWR data: (a) day, $N = 61$; (b) night, $N = 72$ (RTE model ROS). Observations at ARCS-2. In this figure, ''Av'' and ''Sde'' refer to the avg and std dev of (measured − calculated) brightness temperature, respectively.

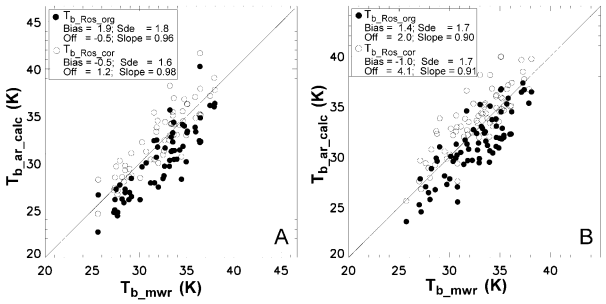

FIG. 10. Comparison of T_b at 31.4 GHz calculated from original (closed circles) and corrected (open circles) radiosondes vs ARM MWR data. Same notation and sample sizes as in Fig. 9.

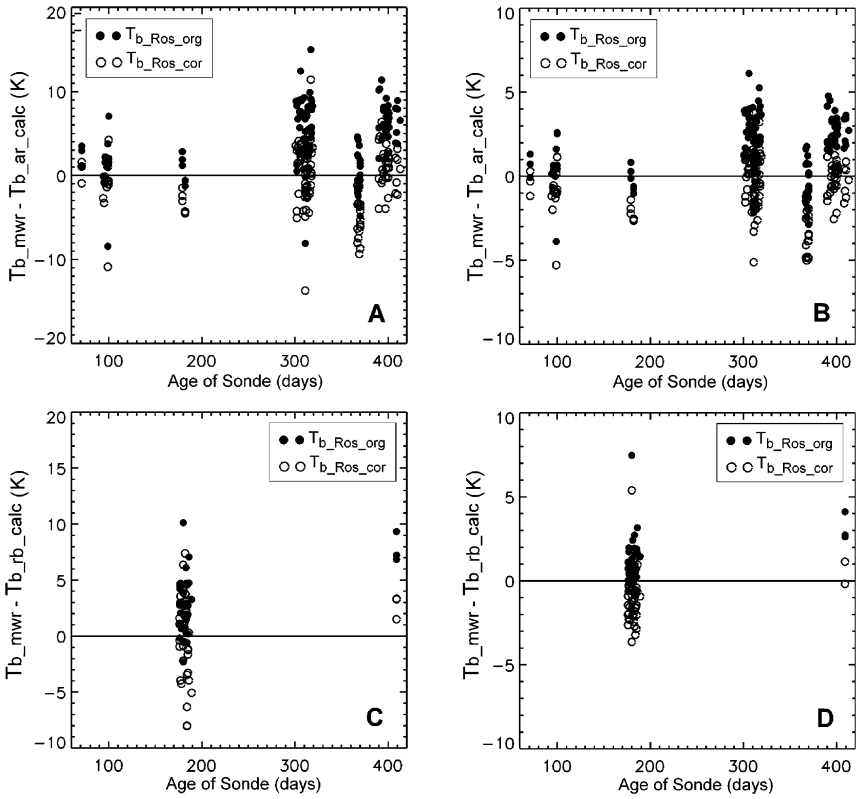

FIG. 11. Comparison of T_b differences as a function of radiosonde age (RTE model ROS). Observations at ARCS-2: (a) 23.8 GHz and (b) 31.4 GHz. Observations on *RHB*: (c) 23.8 GHz and (d) 31.4 GHz. Original (closed circles) and corrected (open circles) observations. Sample sizes for the age categories are given in Tables 2 and 3.

TABLE 2. 23.8-GHz MWR_MEAS − MWR_CALC. Statistics of the differences between measured and calculated T_b as a function of radiosonde age for both the original (ORIG) and corrected (CORR) radiosondes. The absorption model used in the calculations was ROS.

	Age of radiosonde	Count	MWR_CALC (K)	Std dev (K)
AR − ORIG	71–71	3	2.5	1.27
AR − CORR	71–71	**3**	**0.6**	**1.34**
AR − ORIG	94–100	15	1.4	3.21
AR − CORR	94–100	**15**	**−1.3**	**3.13**
RB − ORIG	174–179	36	2.6	2.36
RB − CORR	174–179	**36**	**−1.5**	**2.35**
AR − ORIG	179–182	6	0.6	1.64
AR − CORR	179–182	**6**	**−3.4**	**1.25**
AR − ORIG	301–319	58	5.5	3.61
AR − CORR	301–319	**58**	**0.5**	**3.72**
AR − ORIG	367–370	19	0.4	2.80
AR − CORR	367–370	**19**	**−5.0**	**2.65**
AR − ORIG	390–413	35	6.4	2.22
AR − CORR	390–413	**35**	**0.9**	**2.30**
RB − ORIG	409–409	3	7.8	1.33
RB − CORR	409–409	**3**	**1.9**	**1.29**
ALL − ORIG	174–182	42	−2.3	2.37
ALL − CORR	174–182	**42**	**0.68**	**3.44**
ALL − ORIG	301–409	115	−5.01	3.70
ALL − CORR	301–409	**115**	**0.26**	**3.77**

TABLE 3. 31.4 GHz MWR_MEAS − MWR_CALC. Statistics for the differences between measured and calc T_b as a function of radiosonde age for both the original (ORIG) and corrected (CORR) radiosondes. The absorption model used in the calculations was ROS.

	Age of radiosonde	Count	MWR_CALC (K)	Std dev (K)
AR − ORIG	71	3	0.66	0.69
AR − CORR	71	**3**	**−0.40**	**0.74**
AR − ORIG	94–100	15	0.42	1.48
AR − CORR	94–100	**15**	**−0.95**	**1.47**
RB − ORIG	174–179	36	0.92	1.54
RB − CORR	174–179	**36**	**−1.19**	**1.57**
AR − ORIG	179–182	6	−0.22	0.69
AR − CORR	179–182	**6**	**−2.21**	**0.48**
AR − ORIG	301–319	58	2.29	1.54
AR − CORR	301–319	**58**	**−0.28**	**1.61**
AR − ORIG	367–370	19	−0.32	1.43
AR − CORR	367–370	**19**	**−3.15**	**1.39**
AR − ORIG	390–413	35	2.63	0.98
AR − CORR	390–413	**35**	**−0.10**	**1.04**
RB − ORIG	409–409	3	3.15	0.83
RB − CORR	409–409	**3**	**0.26**	**0.76**
ALL − ORIG	174–182	42	−0.76	1.50
ALL − CORR	174–182	**42**	**1.34**	**1.50**
ALL − ORIG	301–409	115	−1.98	1.71
ALL − CORR	301–409	**115**	**−0.69**	**1.69**

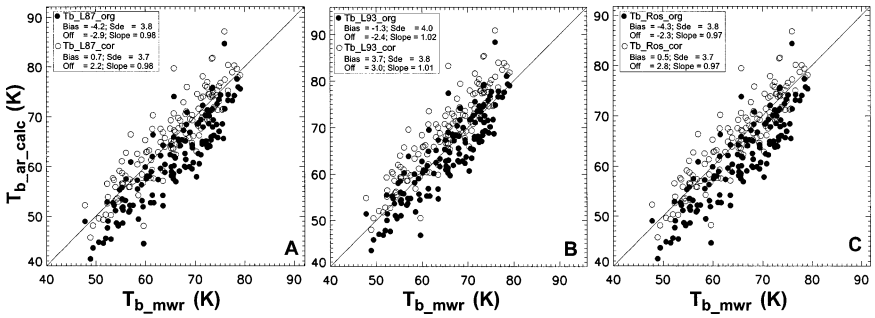

FIG. 12. Comparisons of 23.8-GHz T_b calculated from various absorption models for original and corrected Vaisala radiosondes launched at ARCS-2. The three absorption models (a) L87, (b) L93, and (c) ROS, are defined in the text; original (closed circles) and corrected (open circles). $N = 133$. In this figure, ''Av'' and ''Sde'' refer to the avg and std dev of (measured − calculated) brightness temperature.

dian values for L87 and ROS differ in T_b by about 1.0 K and, if the corrected radiosonde data have zero bias, then L87 is better. As is discussed above, there is a statistically significant difference in the L87 and ROS at 31.4 GHz.

It is frequently convenient to use PWV derived from the MWR (Westwater 1993), rather than T_b, to compare with radiosondes. Our algorithms use the derived-variable optical depth τ. We developed dual- and single-frequency algorithms to derive PWV from τ as follows:

a) a dual-frequency algorithm over a clear + cloudy ensemble of profiles,
b) a dual-frequency algorithm over an ensemble of clear profiles,
c) a single-frequency algorithm using 23.8-GHz τ over an ensemble of clear profiles, and
d) a single-frequency algorithm using 31.4-GHz τ over an ensemble of clear profiles is used.

Each of these algorithms yields PWV as

$$\text{PWV} = a_0 + a_{23.8}\tau_{23.8} + a_{31.4}\tau_{31.4}, \qquad (5)$$

where either physical or purely statistical algorithms (Westwater 1993) can be used to determine the constants *a*. Here, each of the algorithms (a)–(d) was developed for the three absorption models using linear statistical inversion. Our a priori ensemble of profiles was developed from several ocean stations of data taken during TOGA COARE (Webster and Lucas 1992), and we assumed a radiometric noise level of 0.3-K std dev. Examples of retrievals of PWV are shown in Fig. 16. We note that, for a given absorption model, as much as 0.5-cm difference was observed between the various algorithms. We also note that the smallest dispersion between retrievals was for ROS, presumably because it was more self-consistent than the others. In general, one would not use the single channel at 31.4 for the PWV retrieval because of its smaller sensitivity to PWV and its nonideal humidity weighting function (Westwater 1993). For ROS, the PWV retrieval was relatively insensitive to the choice of algorithms (a), (b), or (c).

7. Comparisons with R/V *Mirai* radiosondes

As a part of the measurement intercomparisons, on 3 July the *RHB* and the *Mirai* were in close proximity, and simultaneous radiosondes were launched 4 times at 3-h intervals. As in the case of the ARCS-2 and the *RHB* radiosondes, these units used RS80-H Humicap humidity sensors. However, in contrast to the humidity correction that we used for the ARCS-2 and *RHB* radiosondes, the *Mirai* radiosondes were processed by the

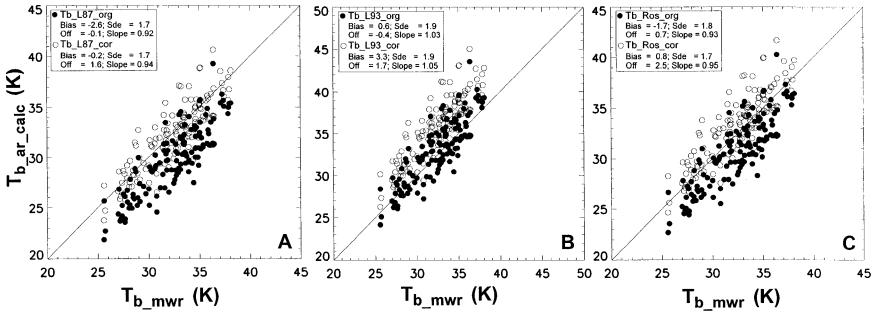

FIG. 13. Comparisons of 31.4-GHz T_b calculated from various absorption models for original and corrected Vaisala radiosondes launched at ARCS-2. Same notation and sample size as in Fig. 12.

TABLE 4. Comparison of measured and calculated brightness temperatures (K) at 23.8 GHz for original and corrected Vaisala radiosondes and for three absorption models; 99% confidence intervals (Ci) for the averages are also given.

		L87		L93		ROS	
		Original	Corrected	Original	Corrected	Original	Corrected
ARCS	Avg	−4.20	0.65	−1.33	3.72	−4.29	0.54
(N = 136)	Std dev	3.84	3.69	3.98	3.82	3.81	3.67
	Ci	±0.85	±0.82	±0.88	±0.84	±0.84	±0.81
RHB	Avg	−2.90	1.33	0.15	4.55	−3.02	1.20
(N = 39)	Std dev	2.69	2.46	2.77	2.56	2.68	2.45
	Ci	±1.11	±1.02	±1.14	±1.06	±1.10	±1.01
Comp	Avg	−3.91	0.80	−1.00	3.90	−4.01	0.69
(N = 175)	Std dev	3.65	3.46	3.79	3.59	3.63	3.44
	Ci	±0.71	±0.67	±0.74	±0.70	±0.71	±0.67

NCAR algorithm (Wang et al. 2002) and, in addition to age, also used independent surface meteorological observations in the process. In Table 6, we show the differences in PWV between the *RHB* (both original and corrected) and the *Mirai*. Note that there is no consistent trend between the original or corrected *RHB* data and the much newer *Mirai* soundings. Because of the age, it is believed that the *Mirai* is the most accurate of the three sets.

For 3 days in June, the *Mirai* was located immediately adjacent to Nauru (Yoneyama 2000). The *Mirai* radiosondes were again RS80-H radiosondes and they were newer (around 70 days) than those used on Nauru (300 and 388 days). In Table 7, we show the complete dataset of seven collocated and simultaneous soundings, including the age of the radiosondes and the ARM LOS MWR data. In this case we used the ARM LOS data to pick up two soundings that were rejected by the ETL method because of clouds, although the cloud liquid water amount was less than 0.01 cm. We note the excellent agreement (maximum difference of 0.3 cm) with the *Mirai* radiosondes. Except for two cases, the age-only-based corrections of the ARCS-2 radiosondes are within about 0.2 cm. However, at 1800 UTC 17 June, 1800 UTC 18 June, 0000 UTC 19 June, the original data are closer to the MWR than are the corrected ones. In Fig. 17, we compare the measured and calculated T_bs and PWV. For these comparisons, we show only the calculations using ROS. It is quite evident that the agreement between the MWR data and those from the *Mirai* radiosondes is excellent, and that deficiencies in both the original and corrected Vaisala ARCS-2 data are shown clearly. More quantitatively, the maximum difference in PWV with the *Mirai* is 0.3 cm, while the same measure for the original and corrected is 0.81 and 0.55 cm, respectively.

8. Comparisons with AERI measurements

We had another independently calibrated measurement with which to compare corrected and uncorrected radiosondes—the AERI data from ARCS-2. For a portion of the measurement period (3–15 July), the ETL FTIR (Han et al. 1997) was operated on the *RHB*. We made several intercomparisons of the data between the two instruments during times when both the ceilometers and cloud radars on ARCS-2 and the *RHB* indicated clear conditions. A typical cloud product that is available from the ARM archives is shown in Fig. 18 (Clothiaux et al. 2000). Detailed examination of the 24-h ARCS-2 cloud product identified promising cloud-free periods that were examined independently on the *RHB*. Comparisons of measured infrared spectral radiance during the mutually cloud-free periods indicated that both completely independent infrared spectrometers were well calibrated. We show an AERI spectrum in Fig. 19; here, we also indicate the portion of the window region where we make detailed comparisons of mea-

TABLE 5. Comparison of measured and calculated brightness temperatures (K) at 31.4 GHz for original and corrected Vaisala radiosondes and for three absorption models; 99% confidence intervals (Ci) for the averages are also given.

		L87		L93		ROS	
		Original	Corrected	Original	Corrected	Original	Corrected
ARCS	Avg	−2.61	−0.22	0.62	3.28	−1.66	0.80
(N = 136)	Std dev	1.73	1.70	1.91	1.88	1.76	1.73
	Ci	±0.38	±0.37	±0.42	±0.42	±0.39	±0.38
RHB	Avg	−2.08	0.04	1.36	3.71	−1.10	1.08
(N = 39)	Std dev	1.59	1.54	1.71	1.69	1.61	1.57
	Ci	±0.65	±0.64	±0.70	±0.70	±0.66	±0.65
Comp	Avg	−2.49	−0.16	0.79	3.37	−1.53	0.86
(N = 175)	Std dev	1.71	1.66	1.89	1.84	1.74	1.69
	Ci	±0.33	±0.32	±0.37	±0.36	±0.34	±0.33

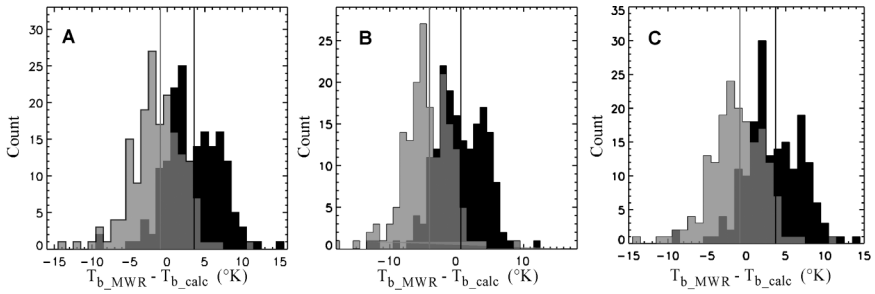

FIG. 14. Histograms of measured minus calculated T_b at 23.8 GHz for original (black) and corrected (light gray) radiosonde data. The overlap regions are shown (dark gray) with the minimum count of the histogram cell (dark gray). Median values for each histogram are shown by the solid vertical lines for (a) L87, (b) L93, and (c) ROS.

surements and calculations. Because comparisons at strongly absorbing spectral regions could be affected by local surface conditions, only the window regions were considered.

Up to now, our examination of radiosonde quality has been based on comparing original and corrected radiosondes with MWR data. For the independently calibrated AERI measurements, we can add another quantity for comparison—radiosonde profiles that have been scaled by the MWR-derived PWV (Clough et al. 1996; Turner et al. 2003). In this method, the absolute humidity at altitude h, $\rho(h)$, is scaled according to

$$\rho_{\text{Scal}} = \rho_{\text{Org}}\left(\frac{\text{PWV}_{\text{MWR}}}{\text{PWV}_{\text{Org}}}\right), \tag{6}$$

where ρ_{Scal} and ρ_{Org} are the corrected and original radiosonde humidities, respectively, and PWV_{MWR} and PWV_{Org} are the derived and original PWV, respectively. In the cases when ρ_{Scal} exceeded saturation, it was replaced by the saturation value. We also checked those cases that exceeded saturation and integrating supersaturated profiles (usually, at most three or four saturated points) gave PWVs that differed by less than 0.05 cm from the truncated values.

Our infrared spectral calculations use the LBL RTM code (Clough et al. 1989, 1992, 1996) that uses the CKD_2.2 continuum model. Before we show comparisons, a few caveats are necessary when making comparisons using the MWR to scale the radiosonde data. The CKD_2.2 continuum model was developed from FTIR and radiosonde observations that were taken during PROBE (Westwater et al. 1999). Because of the radiosonde problems discussed by these authors, only a few (three or four) simultaneous FTIR and radiosonde observations were available during clear conditions and for which the radiosonde agreed with the T_b measured by the ETL MWR to within 0.5 K. The microwave absorption model used by these authors was L87 and thus there is at least some concern about the dependence between the CKD_2.2 continuum model and L87. However, independent comparisons between FTIR measurements and calculations based on radiosondes (Han et al. 1997) indicated an improvement of some 5 K in continuum brightness temperature using the CKD_2.2 model. Thus, although certainly not perfect, we believe that FTIR comparisons are useful in comparing various radiosonde correction algorithms.

Figure 20 shows a comparison between measurements and calculations based on original (315 days old), corrected, and scaled radiosondes, as well as a 172-day-old radiosonde from the *RHB.* From calculations of the AERI over the entire wavelength band, the spectrally integrated radiance from the MWR-scaled radiosonde (using ROS) differed from the measurements by less

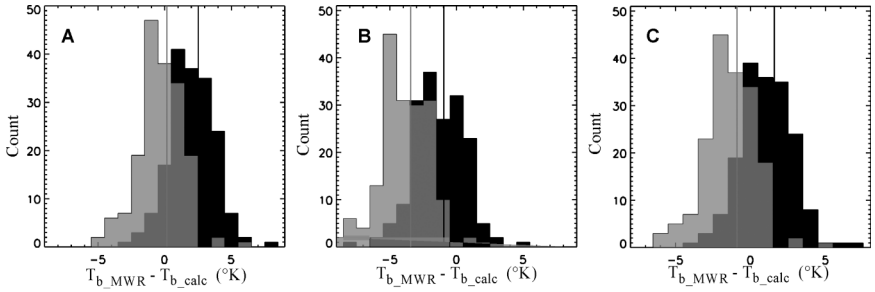

FIG. 15. Histograms of measured minus calculated T_b at 31.4 GHz for original (black) and corrected (light gray) radiosonde data. The overlap regions are shown (dark gray) with the minimum count of the histogram cell (dark gray). Median values for each histogram are shown by the solid vertical lines for (a) L87, (b) L93, and (c) ROS.

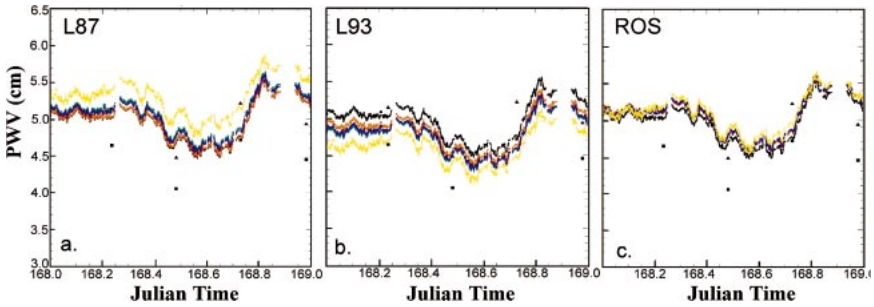

FIG. 16. The 24-h time series of the PWV (10-min averages) at Nauru, showing the original ARM LOS data and the retrievals of methods (a)–(d) described in the text [see Eq. (5)]. Retrieval algorithms: L87, L93, and ROS. Only data with cloud liquid path less than 0.01 cm are shown. Original ARM retrievals that were based on L87 (black), dual-frequency algorithm over a clear + cloudy ensemble of profiles (red), dual-frequency algorithm over an ensemble of clear profiles (green), single-frequency algorithm using 23.8-GHz T_b (blue), single-frequency algorithm using 31.4-GHz T_b (yellow), original radiosonde data (black squares), and radiosondes corrected by the Vaisala algorithm (black triangles).

than 1.5% and the ARCS-2 corrected radiosonde differed by about 6%, whereas the original ARCS-2 differed from AERI by 13%. Another example is shown in Fig. 21, where this time the corrected ARCS-2 radiosonde and the MWR-scaled radiosonde agreed with the measured AERI data to within 3% while the original radiosonde differed by about 9%.

We also computed statistics between AERI (see Fig. 19) spectrally integrated radiance measurements and simulations from original and corrected radiosonde data. We chose a 10-min time interval around the 7-min AERI measurement to check the ceilometer data for clouds and used the AERI data only if none of the points from 15-s data indicated clouds. Figure 22 shows comparisons between original, Vaisala-corrected, and MWR-scaled (using ROS) radiosondes. It is apparent that the original ARCS-2 radiosondes have a substantial dry bias relative to the AERI measurements and that the corrected data agree better with the AERI. Also, there is a slight difference in bias between the Vaisala-corrected and MWR-scaled radiosondes, and slightly less scatter with the MWR-scaled data. In Fig. 23, we compare AERI measurements with all three of the absorption models. It is evident that L87 (the ARM operational model) and ROS give similar results, and both are closer to the AERI measurements than L93.

We note in Figs. 22 and 23, that four sets of radiosonde–MWR-derived points around the AERI measurements of about 3.9×10^4 (mW m^{-2} sr) seem to be inconsistent with the trend exhibited by other MWR-measured points. We went back and examined in detail each of these points and discovered that three of four are on 23 June and the other is on 22 June. The calculations of EZTB during the times in question were around 0.3–0.4-K rms, indicating that the MWR data were accurate. In addition, three of the four radiosondes were from the same lot (365 days old), and thus the application of the Vaisala correction algorithm is perhaps suspect; however, the MWR scaling should have corrected these biases. Because the ETL FTIR on the *RHB* had not yet arrived, a cross check on the AERI calibration was not possible. Because there was no compelling reason to reject these outliers, we left them in the dataset.

9. Discussion and conclusions

The data obtained during Nauru99 provided a rare opportunity to compare basic observations necessary for developing climate and radiative transfer models. Essentially identical instruments were operated from the R/V *RHB* and on ARCS-2. These instruments included simultaneously launched Vaisala radiosondes, the ARM MWR, the ETL FTIR and the ARM AERI, and cloud radars/ceilometers to identify clear conditions. The latter data were used only to screen clouds for FTIR and AERI comparisons.

The ARM MWR operating at the ARCS-2 provided

TABLE 6. *Mirai* vs *RHB* PWV (cm) from radiosonde as a function of radiosonde age for both the original (ORIG) and corrected (CORR) *RHB* radiosondes for Jul 1999.

Time day : UTC	Age ARCS-2	Age *Mirai*	PWV Calc ARCS-2 ORIG	PWV Calc ARCS-2 CORR	PWV Calc *Mirai*
2 : 2100	173	87	4.19	4.52	4.33
3 : 0000	174	88	3.85	4.17	3.83
3 : 0300	174	88	3.45	3.76	3.60
3 : 0600	174	88	3.96	4.27	3.81

TABLE 7. *Mirai* vs ARCS-2 PWV (cm) from radiosonde and MWR measurements as a function of radiosonde age for both the original (ORIG) and corrected (CORR) ARCS-2 radiosondes during Jun 1999. The MWR T_b data were from the operational (LOS) data archive and screened for cloud liquid less than 0.01 cm. The PWV retrieval algorithm used ROS (clear and cloudy).

Time day : UTC	Age ARCS-2	Age *Mirai*	PWV CALC ARCS-2 ORIG	PWV CALC ARCS-2 CORR	PWV CALC *Mirai*	PWV MWR (LOS)
17 : 1200	299	71	4.05	4.49	4.63	4.76
17 : 1800	388	71	4.71	5.23	4.68	4.83
18 : 0000	388	72	4.45	4.95	5.02	5.31
18 : 0600	389	72	4.04	4.51	4.60	4.85
18 : 1200	300	72	4.33	4.78	4.86	4.98
18 : 1800	300	72	4.95	5.42	4.76	4.97
19 : 0000	300	73	4.60	5.08	4.27	4.59

an excellent dataset for the entire Nauru99 experiment. The calibration accuracy was verified by an LN2 blackbody target experiment and by consistent high quality tipcals throughout the experiment. The data thus provide an excellent baseline for evaluation of the quality and consistency of Vaisala radiosondes that were launched from ARCS-2 and *RHB*, and a smaller dataset (seven radiosondes) launched from the *Mirai*. Comparisons were made for calculated clear-sky brightness temperature and for precipitable water vapor. Our results indicate that substantial differences, sometimes of the order of 20% in PWV, occurred with the uncorrected radiosondes. When the Vaisala correction algorithm was applied to the radiosondes, better agreement with the MWR was obtained. However, the improvement was noticeably different for different radiosonde lots and was not a monotonic function of radiosonde age. However, because of the demonstrated quality of the MWR data, the large scatter in the differences between measurements and calculations comes mainly from radiosonde error in the older radiosondes.

We also performed our brightness temperature calculations with three absorption algorithms—Liebe and Layton (1987), Liebe et al. (1993), and Rosenkranz (1998). When using the Rosenkranz (1998) absorption model applied to newer radiosondes, both from the *RHB* and *Mirai*, agreement with the ARM MWR was substantially better than with the original radiosondes. For example, at the vapor-sensitive 23.8-GHz channel, the correction algorithm reduced the bias from −2.3 to 0.68 K for the 180-day-old radiosondes and from −5.01 to 0.26 K for the 301–409-day-old radiosondes. L87 was within about 1-K agreement with ROS, but each differed by 3 K with L93. For the 23.8-GHz vapor-sensitive channel, L87 and ROS differed by only 0.1 K in bias in T_b comparisons. At 31.4 GHz, the bias difference was about 1.0 K between the two; however, the uncertainties in the accuracy of either the original or corrected radiosondes do not allow one to make definite statements about the better of the two models. This is in contrast to observations in a very cold and dry arctic environment, when ROS was clearly superior to the other two at 31.4 GHz (Westwater et al. 2001). For each of the absorption models, we also performed both single- and dual-frequency retrievals of PWV based on Eq. (5). However, the PWV retrievals that used only the 31.4-GHz channel demonstrate that ROS was more self-consistent than the other two.

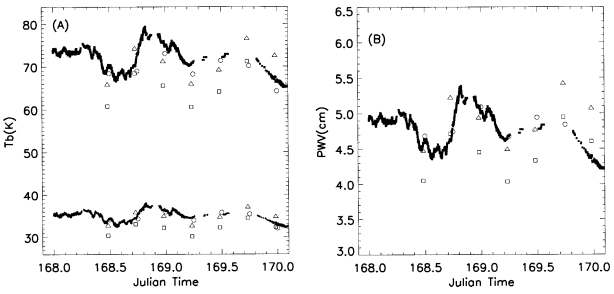

FIG. 17. The (a) 48-h time series of the T_b at 23.8 and 31.4 GHz (10-min-averaged data) showing the MWR data, with cloud liquid path less than 0.005 cm, using the ROS forward model; original ARCS-2 radiosondes (squares), corrected ARCS-2 radiosondes (triangles), and *Mirai* radiosondes (circles) in Nauru99; and (b) 48-h time series of the PWV retrieved from the MWR data. Same notation as in (a).

We had another independently calibrated measurement with which to compare corrected and uncorrected radiosondes—the AERI data from ARCS-2. To compare with AERI data, we calculated radiance over a transmission window from 750 to 950 cm^{-1} from a variety of input atmospheric profiles: original and corrected radiosondes from both ARCS-2 and the *RHB*, as well as the original profiles scaled by PWV derived from the ARM MWR, using algorithms derived from L87, L93, and ROS. Our infrared spectral calculations use the LBL RTM code (Clough et al. 1989, 1992, 1996) that uses the CKD_2.2 continuum model. The CKD_2.2 continuum model was developed from FTIR and radiosonde observations that were taken during PROBE (Westwater et al. 1999). The microwave absorption model used by these authors was L87 and thus there is at least some concern about the dependence between the CKD_2.2 continuum model and L87. However, independent comparisons between FTIR measurements and calculations based on radiosondes (Han et al. 1997) indicated an improvement of some 5 K in continuum brightness tem-

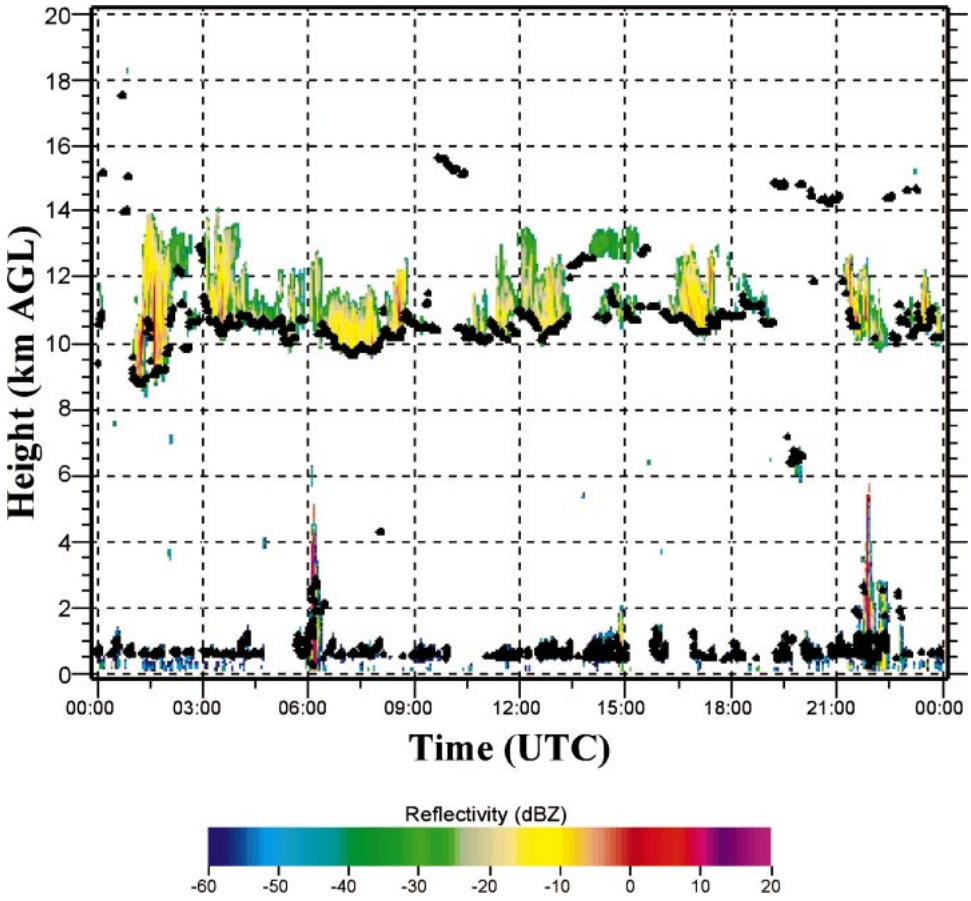

FIG. 18. Time series of cloud boundaries as determined from lidar and cloud radar on 17 Jun 1999 (Clothiaux et al. 2000). The black dots are determined from the micropulse lidar while the colored reflectivities are from the MMCR. The black dots above 14 km are particles seen by the lidar but not by the MMCR. This is an example of an operational product produced by ARM (information available online at http://www.arm.gov/).

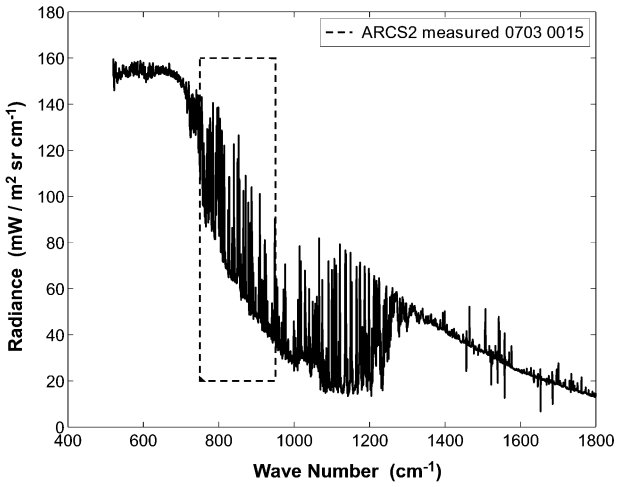

FIG. 19. Infrared radiance spectrum measured by the AERI on ARCS-2 during Nauru99 in clear conditions. The portion enclosed in dashed lines indicates the portion of the transparent atmospheric window that we analyze in subsequent figures.

perature using the CKD_2.2 model. Thus, although certainly not perfect, we believe that FTIR comparisons are useful in comparing various radiosonde correction algorithms. These comparisons showed that, using rms/mean as our measure of quality, that the poorest agreement with AERI came from the uncorrected radiosondes (7.4%), the next two poorest were the corrected radiosondes and the L93-scaled radiosondes (both 5.1%), and the two best were ROS- and L87-scaled radiosondes (4.5% and 4.4%, respectively).

Because all of the ARM CART sites have operational MWR instruments, scaling (Clough et al. 1996, 1999; Turner et al. 2003) and quality control (Westwater et al. 1989, 1999) of radiosonde data by either T_b (clear skies only) or PWV measurements (cloud liquid less than 1 or 2 mm) is promising. The results reported by Turner et al. (2003), obtained at the ARM Central Facility in northern Oklahoma and using L87 to process the MWR data, were consistent with ours and show similar promise. Because of the wide deployment of GPS instruments (Wolfe and Gutman 2000), with similar accuracy to the

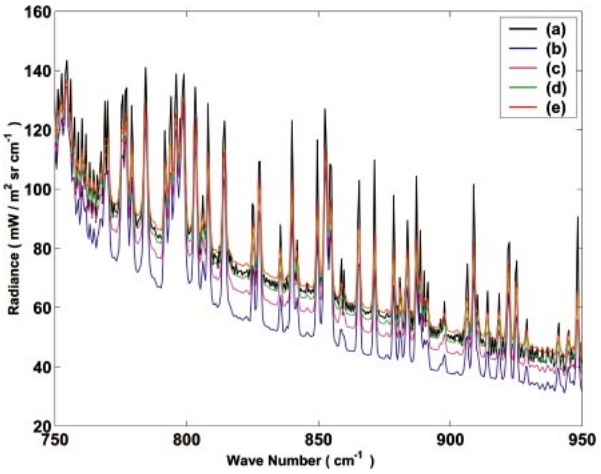

FIG. 20. Measured spectral radiance compared with calc radiance using a variety of original, corrected, and scaled radiosondes on 3 Jul at 0532: (a) measured by AERI, (b) calculated from original ARCS-2 radiosonde (age = 315 days, PWV = 3.14 cm, difference from AERI = 13%), (c) calculated from the corrected ARCS-2 radiosonde (PWV = 3.52 cm; difference from AERI = 5.8%), (d) MWR-scaled (ROS) PWV measurements (PWV = 3.73 cm, difference from AERI = 1.5%), and (e) calculated from the original *RHB* radiosonde (age = 172 days, PWV = 3.98 cm, difference from AERI = 2.0%).

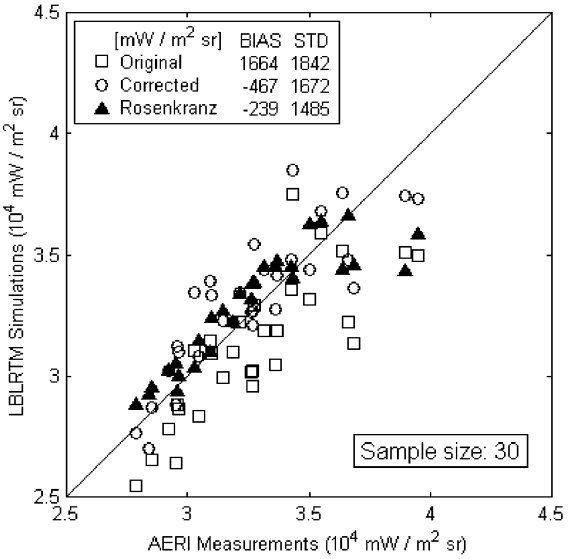

FIG. 22. Scatterplot of calculated and AERI-measured integrated spectral radiance using ARCS-2 original radiosondes, corrected ARCS-2 radiosondes, and ARCS-2 radiosondes that were scaled by PWV derived from the ARCS-2 MWR and the ROS forward model in the retrieval. In this figure, ''BIAS'' and ''STD'' refer to the avg and std dev of (measured − calculated) quantities.

MWR in PWV retrieval, similar scaling and quality control is possible. The temporal continuity of either MWR or GPS measurements is thus a powerful tool for identifying and correcting erroneous humidity soundings.

In addition to the needs for screening and scaling of poor-quality radiosondes, there is also a need for high-quality soundings of water vapor from which accurate forward models for microwave radiometers can be developed. If the residual uncertainties in microwave forward models, for example, ROS and L87, can be reduced, then the quality of both past and present data can be substantially increased.

Acknowledgments. The work presented in this paper was sponsored by the Environmental Sciences Division

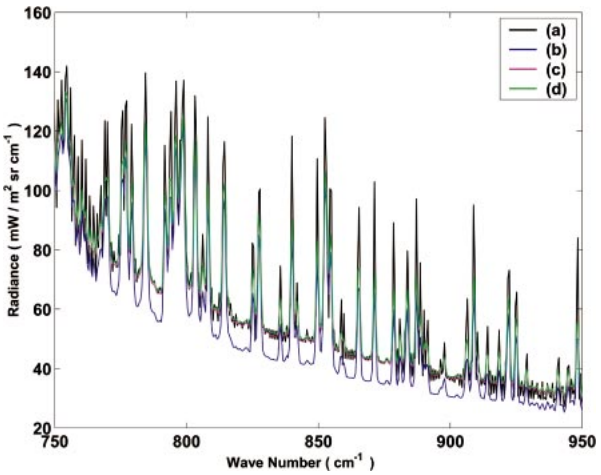

FIG. 21. Measured and calculated spectral radiance using a variety of original, corrected, and scaled radiosondes on 3 Jul at 1731 UTC: (a) measured by AERI, (b) calculated from original ARCS-2 radiosonde (age = 315 days, PWV = 2.69 cm, diff from AERI = 8.7%), (c) calculated from the corrected ARCS-2 radiosonde (PWV = 3.04 cm, difference from AERI = 0.9%), and (d) MWR-scaled (ROS) PWV (PWV = 3.16 cm, difference from AERI = 3.2%).

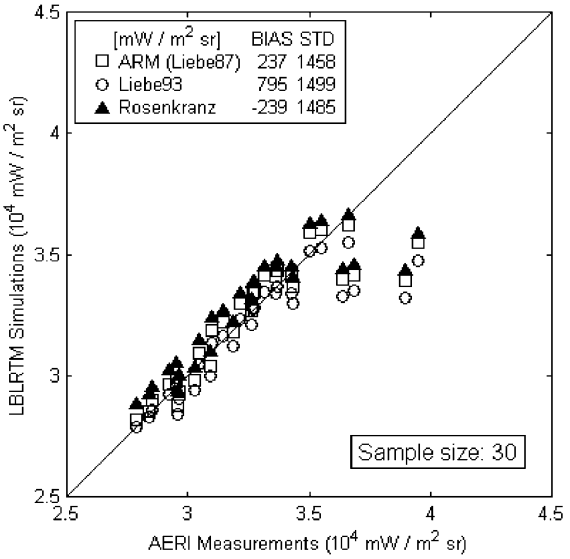

FIG. 23. Scatterplot of calculated and AERI-measured integrated spectral radiance using ARCS-2 original radiosondes, scaled by retrievals using the indicated forward models. The ARM-scaled retrievals use L87. Same notation as in Fig. 22.

of the Department of Energy as a part of their Atmospheric Radiation Measurement Program. The authors also thank K. Yoneyama for providing us with the R/V *Mirai* radiosonde data. Finally, the authors acknowledge the insightful comments of two anonymous reviewers and to Dave Turner for providing us with a preprint of his paper.

REFERENCES

Clothiaux, E. E., T. P. Ackerman, G. G. Mace, K. P. Moran, R. T. Marchand, M. A. Miller, and B. E. Martner, 2000: Objective determination of cloud heights and radar reflectivities using a combination of active remote sensors at the ARM CART sites. *J. Appl. Meteor.,* **39,** 645–665.

Clough, S. A., F. X. Kneizys, and R. W. Davies, 1989: Line shape and the water vapor continuum. *Atmos. Res.,* **23,** 229–241.

——, M. J. Iacono, and J. L. Moncet, 1992: Line-by-line calculation 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, Atmospheric Radiation Measurement Program, 45–49. [Available online at http://www.arm.gov/docs/documents/technical/conf_9603/index.html.]

——, ——, 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, Atmospheric Radiation Measurement Program. [Available online at http://www.arm.gov/docs/documents/technical/conf_9903/index.html.]

Deslover, D. H., W. L. Smith, P. K. Piironen, and E. W. Eloranta, 1999: A methodology for measuring cirrus cloud visible-to-infrared spectral optical depth ratios. *J. Atmos. Oceanic Technol.,* **16,** 251–262.

Feltz, W. F., W. L. Smith, R. O. Knuteson, H. E. Revercomb, H. M. Woolf, and H. B. Howell, 1998: Meteorological applications of temperature and water vapor retrievals from the ground-based Atmospheric Emitted Radiance Interferometer (AERI). *J. Appl. Meteor.,* **37,** 857–875.

Fleming, H. E., N. C. Grody, and E. J. Kratz, 1991: The forward problem and corrections for the SSM/T satellite microwave temperature sounder. *IEEE Trans. Geosci. Remote Sens.,* **29,** 571–583.

Guichard, F., D. Parsons, and E. Miller, 2000: Thermodynamical and radiative impact of the correction of sounding humidity bias in the Tropics. *J. Climate,* **13,** 3611–3624.

Han, Y., and E. R. Westwater, 2000: Analysis and improvement of tipping calibration for ground-based microwave radiometers. *IEEE Trans. Geosci. Remote Sens.,* **38,** 1260–1277.

——, J. A. Shaw, J. H. Churnside, P. D. Brown, and S. A. Clough, 1997: Infrared spectral radiance measurements in the tropical Pacific Ocean. *J. Geophys. Res.,* **102,** 4353–4356.

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.

——, 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 Measurement (ARM) Science Team Meeting,* San Antonio, TX, Atmospheric Research Measurement Program. [Available online at http://www.arm.gov/docs/documents/technical/conf_9903/index.html.]

——, and J. C. Liljegren, 1996: 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 Measurement (ARM) Science Team Meeting,* San Antonio, TX, Atmospheric Radiation Measurement Program, 165–168. [Available online at http://www.arm.gov/docs/documents/technical/conf_9603/index.html.]

Liebe, H. J., 1989: MPM, an atmospheric millimeter wave propagation model. *Int. J. Infrared Millimeter Waves,* **10,** 631–650.

——, and D. H. Layton, 1987: Millimeter wave properties of the atmosphere: Laboratory Studies and propagation modeling. National Telecommunications and Information Administration (NTIA) Rep. 87–24, 74 pp.

——, G. A. Hufford, and M. G. Cotton, 1993: Propagation modeling of moist air and suspended water/ice particles at frequencies below 1000 GHz. *Atmospheric Propagation Effects through Natural and Man-Made Obscurants for Visible through MM-Wave Radiation,* AGARD-CP-542, 3.1–3.10.

Liljegren, J. C., 2000: Automatic self-calibration of ARM microwave radiometers. *Microwave Radiometry and Remote Sensing of the Earth's Surface and Atmosphere,* P. Pampaloni and S. Paloscia, Eds., VSP Press, 433–441.

Ma, Q., and R. H. Tipping, 1990: Water vapor continuum in the millimeter spectral region. *J. Chem. Phys.,* **93,** 6127–6139.

Miller, E. R., J. Wang, and H. L. Cole, 1999: Correction for dry bias in Väisälä radiosonde RH data. *Proc. Ninth Annual Atmospheric Radiation Measurement (ARM) Science Team Meeting,* San Antonio, TX, Atmospheric Radiation Measurement Program. [Available online at http://www.arm.gov/docs/documents/technical/conf_9903/index.html.]

Post, M. J., and C. F. Fairall, 2000: Early results from the Nauru99 campaign on the NOAA ship RONALD H. BROWN. *Proc. Int. Geoscience and Remote Sensing Symp.,* Honolulu, HI, GRSS, 1151–1153.

Revercomb, H. E., 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,* Anaheim, CA, Amer. Meteor. Soc., 46–49.

Rosenkranz, P. W., 1998: Water vapor microwave continuum absorption: A comparison of measurements and models. *Radio Sci.,* **33,** 919–928.

Shaw, J. A., J. H. Churnside, and E. R. Westwater, 1991: An infrared spectrometer for ground-based profiling of atmospheric temperature and humidity. *Proc. SPIE Int. Symp. on Optical and Applied Science and Engineering,* San Diego, CA, Society of Photo-Optical Instrumentation Engineers, 681–686.

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

Tobin, D. C., and Coauthors, 1999: Downwelling spectral radiance observations at the SHEBA ice station: Water vapor continuum measurements from 17 to 26 microns. *J. Geophys. Res.,* **104** (D2), 2081–2092.

Turner, D. D., B. Lesht, S. A. Clough, J. C. Liljegren, H. E. Revercomb, and D. C. Tobin, 2003: Dry bias and variability in Vaisala RS80-H radiosondes: The ARM experience. *J. Atmos. Oceanic Technol.,* **20,** 117–132.

Wade, C. G., 1994: An evaluation of problems affecting the measurement of low relative humidity on the U.S. radiosonde. *J. Atmos. Oceanic Technol.,* **11,** 687–700.

Wang, J., H. L. 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.

Webster, P. J., and R. Lucas, 1992: TOGA COARE: The Coupled

Ocean Atmosphere Response Experiment. *Bull. Amer. Meteor. Soc.*, **73,** 1377–1416.

Westwater, E. R., 1993: Ground-based microwave remote sensing of meteorological variables. *Atmospheric Remote Sensing by Microwave Radiometry*, M. Janssen, Ed., John Wiley and Sons, Inc., 145–213.

——, M. J. Falls, and I. A. Popa Fotino, 1989: Ground-based microwave radiometric observations of precipitable water vapor: A comparison with ground-truth from two radiosonde observing systems. *J. Atmos. Oceanic Technol.*, **6,** 724–730.

——, J. B. Snider, and M. J. Falls, 1990: Ground-based radiometric observations of atmospheric emission and attenuation at 20.6, 31.65, and 90.0 GHz: A comparison of measurements and theory. *IEEE Trans. Antennas Propag.*, **38,** 1569–1580.

——, and Coauthors, 1999: Ground-based remote sensor observations during PROBE in the tropical western Pacific. *Bull. Amer. Meteor. Soc.*, **80,** 257–270.

——, Y. Han, M. D. Shupe, and S. Y. Matrosov, 2001: Analysis of integrated cloud liquid and precipitable water vapor retrievals from microwave radiometers during SHEBA. *J. Geophys. Res.*, **106,** 32 019–32 030.

Wolfe, D. E., and S. I. Gutman, 2000: Developing an operational, surface-based, GPS, water vapor observing system for NOAA: Network design and results. *J. Atmos. Oceanic Technol.*, **17,** 426–440.

Yoneyama, K., 2000: Activities of the R/V *Mirai* Nauru99 Cruise and its early results. *Proc. Int. Geoscience and Remote Sensing Symp.*, Honolulu, HI, GRSS, 1148–1150.