Accuracy of Ground-Based Microwave Radiometer and Balloon-Borne Measurements During the WVIOP2000 Field Experiment

Domenico Cimini, *Member, IEEE*, Ed R. Westwater, *Fellow, IEEE*, Yong Han, and Stephen J. Keihm

***Abstract*—We discuss the performances of a set of four microwave water vapor radiometers operating in the 20–30-GHz band during a field experiment, with an emphasis on calibration and achievable accuracy. The field experiment was conducted at the Department of Energy's Atmospheric Radiation Measurement Program's field site in north central Oklahoma, and was focused on clear-sky water vapor measurements by both radiometers and radiosondes. A comparison between two published radiometric tip curve calibration procedures is presented, and these procedures are applied to measurements from two nearly identical instruments placed a few meters apart. Using the instantaneous tip cal method of the Environmental Technology Laboratory, the brightness temperature measurements for the two identical instruments differed by less than 0.2 K over a 24-h period. Results from reference load cryogenic tests and brightness temperature cross comparisons have shown differences within 0.7 K. In addition, we compare radiometric measurements with calculations of brightness temperature based on the Rosenkranz absorption model and radiosonde observations. During the experiment, both Vaisala-type RS80 and RS90 humidity sensors were used. Our comparisons demonstrate the improvements achieved by the new Vaisala RS90 sensors in atmospheric humidity profiling, which reduce or eliminate the "dry bias" problem.**

***Index Terms*—Microwave radiometry, radiometric accuracy, radiometric calibration, water vapor.**

I. Introduction

THE ACCURATE measurement of water vapor profiles is important to a variety of studies, including weather forecasting, development of radiative transfer algorithms, radio propagation, remote sensor evaluation, and the development of climate models [1]–[6]. Because of the importance to the Department of Energy's Atmospheric Radiation Measurement Program (ARM) [7], [8], a series of water vapor experiments have been conducted by ARM in the tropics [9], in the arctic [10], and in the central United States [11]. All of these experiments, mainly through the inter comparison of Microwave Radiometers (MWR) and Vaisala radiosondes with the RS80 humidity sensors, have indicated the presence of about a 5% dry bias in the measurements of relative humidity. This bias and its uncertainty have limited the development of radiative transfer models of importance to climate studies [11], [12].

During September–October 2000, a Water Vapor Intensive Operational Period (WVIOP) was conducted at the ARM Program's Southern Great Plains (SGP) Cloud and Radiation Testbed (CART) site near Lamont, OK [7]. The main goal of the WVIOP2000 was to characterize the accuracy of several current water vapor measurements under a wide range of values, with an emphasis on instrument comparisons, and to develop and test techniques to improve the accuracy of these observations. This particular site and time of year was chosen by ARM because it offers a high probability of clear skies and a wide range of integrated water vapor amounts. To supplement the routine observing capabilities during the WVIOP2000, a powerful array of instruments for measuring water vapor was deployed at the SGP Central Facility area. A comprehensive suite of airborne *in situ* and ground-based remote sensors was involved, including multiple radiosondes, a ceilometer, and MWRs.

Among the WVIOP2000 activities, we wanted to evaluate the microwave radiometers' calibration accuracy. By comparing brightness temperature (Tb) measurements and precipitable water vapor (PWV) estimates from a set of independent MWRs, we were able to assess the relative uncertainty in both quantities. Moreover, we performed several absolute accuracy tests using a cryogenic target. In addition, by comparing MWR measurements with Tb calculations based on radiosonde observation of the atmospheric thermodynamic state, we were able to demonstrate and quantify the improvements achieved by the new generation of radiosonde sensors in humidity profiling.

In Section II we introduce the instruments we used; in Section III we show the comparison between two alternative tip curve calibrations; in Section IV we compare measurements from the four MWR units; in Section V we discuss the agreement between MWR and observations from two different kinds of radiosondes. In Section VI, a summary of the results is proposed, together with a hint for future developments.

II. Instruments and Calibration

The set of radiometers under study is composed of one three-channel unit and three dual-channel units, each mea-

Manuscript received December 2, 2002; revised May 7, 2003. This work was supported by the Environmental Sciences Division, Department of Energy under the Atmospheric Radiation Measurement Program.

D. Cimini is with CETEMPS, University of L'Aquila, L'Aquila, 67010 Italy (e-mail: Nico.Cimini@aquila.infn.it).

E. R. Westwater is with the Environmental Technology Laboratory, Cooperative Institute for Research in Environmental Sciences, University of Colorado/NOAA, Boulder, CO 80305 USA (e-mail: Ed.R.Westwater@noaa.gov).

Y. Han is with the Office of Research and Application, National Environmental Satellite, Data, and Information Service (NESDIS), National Oceanic and Atmospheric Administration (NOAA), Camp Springs, MD 20746 USA (e-mail: Yong.Han@noaa.gov).

S. J. Keihm is with Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, 91109 USA (e-mail: stephen.j.keihm@jpl.nasa.gov).

Digital Object Identifier 10.1109/TGRS.2003.815673

TABLE I
MAIN CHARACTERISTICS FOR THE SET OF MWR DEPLOYED DURING THE WVIOP2000. WITH TIP CURVE CALIBRATION FACTOR IS INDICATED THE VARIABLE TO BE DETERMINED WITH THE TIPPING CURVE METHOD. UNDER CALIBRATION PROCEDURE ARE LISTED THE TOOLS USED BY EACH RADIOMETER TO ESTABLISH CALIBRATION. FOR FURTHER DETAILS, SEE [15] AND [19] FOR THE CF AND SU, [13] AND [16] FOR THE CSR, AND [14] FOR THE JPL UNITS

NAME	OWNER	CENTRAL FREQUENCY (GHZ)	BANDWIDTH (MHZ)	BEAMWIDTH (DEG)	TIP CURVE CALIBRATION FACTOR	CALIBRATION PROCEDURE
CF	ARM	23.8 – 31.4	400	6 – 4	NOISE DIODE TEMPERATURE	NOISE DIODE, INTERNAL REFERENCE, TIP CURVE
SU	ARM	23.8 – 31.4	400	6 – 4	NOISE DIODE TEMPERATURE	NOISE DIODE, INTERNAL REFERENCE, TIP CURVE
CSR	NOAA/ETL	20.6 – 31.65	500	4 – 4	TRANSMISSION LOSS FACTOR	TWO INTERNAL REFERENCES, ONE EXTERNAL TARGET, TIP CURVE
JPL	NASA/JPL	20.7 – 22.2 – 31.4	400	9 – 6	NOISE DIODE TEMPERATURE	NOISE DIODE, INTERNAL REFERENCE, TIP CURVE

suring downwelling Tb in the 20/30-GHz spectral region. Radiometers operating in this frequency band are commonly used to estimate PWV [13]. ARM deployed two identical dual-channel (23.8 and 31.4 GHz) radiometers, located a few meters apart, hereafter called the central facility (CF) and spare unit (SU). The CF unit has been operational at the ARM SGP CART site for several years. The National Aeronautics and Space Administration (NASA) Jet Propulsion Laboratory (JPL) participated with a three-channel radiometer (20.7, 22.2, and 31.4 GHz), while the National Oceanic and Atmospheric Administration (NOAA) Environmental Technology Laboratory (ETL) operated the dual channel Circulary Scanning Radiometer (CSR) (20.6 and 31.65 GHz). Some of these channels overlap in frequency, which is desirable because it provides a good opportunity to study calibration. The main characteristics of the CSR, JPL and ARM units are listed in Table I, while details are given in [13]–[15], respectively.

The MWRs ran simultaneously for 21 days, continuously scanning in the east–west direction. Each radiometer used a slightly different automated calibration procedure, including a combination of noise diode sources and internal and external reference targets. Moreover, measurements taken during each scan were processed to compute the tip curve calibration correction [16], with which we refine an initial gain parameter value.

During the experiment, we occasionally performed absolute accuracy tests on the ARM MWRs, using a high-emissivity absorber cooled in a liquid nitrogen (LN2) bath. The LN2 boiling temperature is 77 K under standard conditions, while the absorber emissivity, the box attenuation, and the atmospheric pressure are taken into account to estimate the brightness temperature of the target. The uncertainty for the equivalent blackbody temperature for this target was estimated to be better than 0.2 K [17]. We placed this target for about 10 min on an aluminum saddle that perfectly fits the radiometers' windows. The target and aluminum saddle system prohibits external microwave radiation from getting into the radiometer antenna's field of view. Hence, during these observations the two channels should measure the same brightness temperature, the value of which should be close to the expected temperature. Some results from this analysis are shown in Fig. 1. Although the agreement is usually satisfactory, there are some apparent differences. The comparison between measurements from the two channels on the same instrument gives better results than the comparison with the estimated value of the target Tb, which was usually warmer than the measurements. This effect might be related to LN2 vapors leaking in the interface at the bottom of the container. Indeed, as suggested in [17], the polystyrene box is slightly porous and allows nitrogen gas to permeate the walls. The net effect, which is difficult to quantify, would result in overestimating the box's insertion loss and thus the expected brightness temperature. This feature puts a question on the actual accuracy for the equivalent blackbody temperature of the target in the present design. Nevertheless, the differences between Tbs at the two channels and the estimated target Tb usually remained within 0.7 K.

III. TIPPING CURVE ALGORITHMS

The tipping curve method relies on the assumption of horizontally-stratified atmosphere, which leads to a linear relationship between atmospheric opacity (τ) and the air mass (α). The air mass is indeed defined as the ratio of τ at elevation angle θ and τ at the zenith $[\alpha(\theta) = \tau(\theta)/\tau(\pi/2)]$. By relating $\tau(\theta)$, estimated by measurements of Tb(θ) [4], and theoretical values of $\alpha(\theta)$ at two or more elevation angles, the tipping curve method can provide an adjustment of a single calibration parameter, as is carefully discussed in [16]. The tipping curve method is very general, as it is applicable to any radiometer working with frequency corresponding to low attenuation (τ roughly between 0.005 and 0.5 nepers), but the algorithm to implement depends on which calibration parameter needs to be tuned and on the radiometer equation specific to the instrument in use. Details for the ETL and ARM MWR units are given in [16].

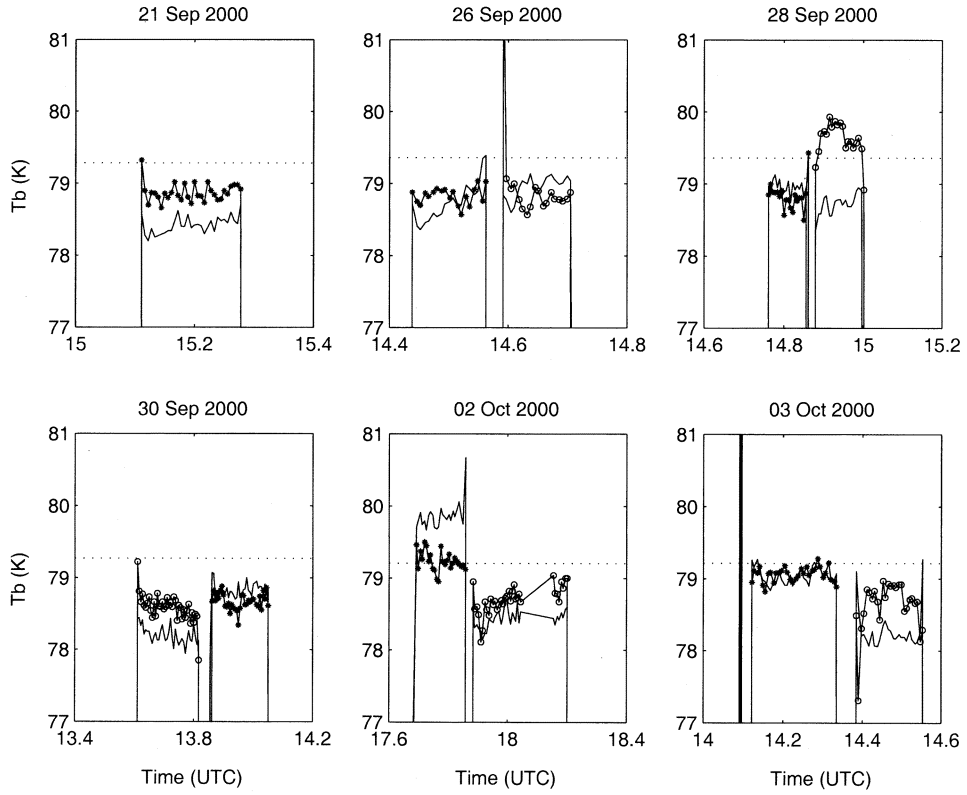


Fig. 1. LN2-based absolute calibration tests performed on ARM CF and SU. The CF 23.8-GHz Tb is shown by a solid line with asterisks, while the SU 23.8-GHz Tb is shown by a solid line with circles. The CF and SU 31.4-GHz Tbs are both shown with plain solid lines, but are distinguishable because are simultaneous with measurements at 23 GHz from the corresponding unit. The expected value (dashed line) has been computed according to [17].

During the WVIOP2000, we used the tipping curve method to adjust the gain factor for all of the four MWR in operation. In the postprocessing stage, we deployed two different procedures: the first was proposed by ETL and it is described in [16], while the second one, which corresponds to the ARM operational calibration, is described in [18] and [19]. We will therefore refer to them as the ETL and ARM tip curve calibration procedures. 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.

To compare the ARM and ETL procedures, we particularly focus on the two ARM units, since they are identical and located just few meters apart. Thus, any difference in brightness temperature should be only related to the effects of instrumental noise and calibration. The radiometric equation for such MWR units is

$$T_{\mathrm{Sky}} = T_{\mathrm{Ref}} + \Gamma \cdot (V_{\mathrm{Sky}} - V_{\mathrm{Ref}}) \tag{1}$$

where T_{Sky} 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 at the sky and when it is looking at the reference target, while Γ is a multiplicative factor. Calling $V_{\mathrm{Ref}+nd}$ the signal measured during reference target observation coupled with injection from the noise diode, T_{nd} the noise diode equivalent brightness temperature, and f_W the polycarbonate foam window loss factor, we have

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

where G is the radiometer gain factor (see [16], [18], and [19]). Equation (2) assumes that the polycarbonate foam window is at the reference target temperature T_{Ref}, since this value is also the temperature inside the radiometer enclosure.

When applied to the ARM MWR units, the ARM and ETL procedures use the tip curve method to adjust the value of T_{nd}, which is proportional to the gain factor G. In fact, the output of the noise diode is constant, since its temperature is controlled (± 0.25 K), but the value of T_{nd} can fluctuate because the antenna and feedhorn are not thermally stabilized [18]. The most significant difference between the two procedures is the method of deriving the calibration coefficient from the tip curve data. As described in [18] and [19], the ARM operational algorithm estimates the value of T_{nd} from each set of measurements taken at ten different elevation angles (corresponding to 1, 1.5, 2, 2.5, and 3 air masses, both sides). In order 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. When new, quality-tested values of T_{nd} and T_{Ref} are measured and stored, the oldest are 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}.

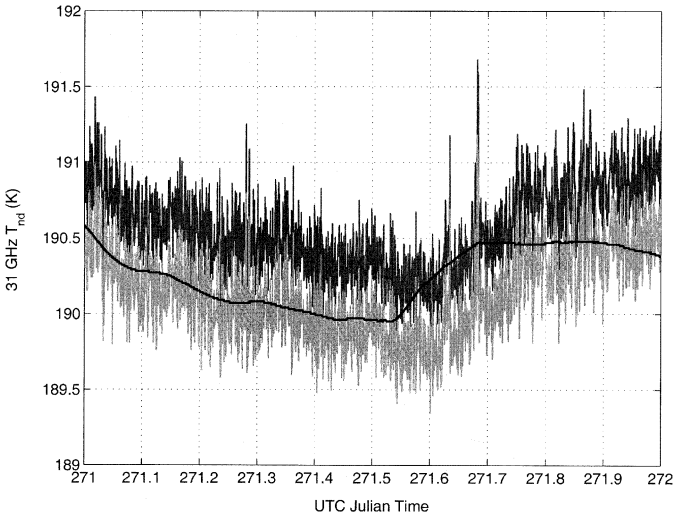


Fig. 2. Time series of CF 31.4-GHz tip curve calibration coefficient during Julian day 271 (September 27). The ARM ($T_{\rm nd}$) and ETL ($f_W \cdot T_{\rm nd}$) instantaneous values are shown in light and dark gray, respectively, while the ARM operational value is shown in black.

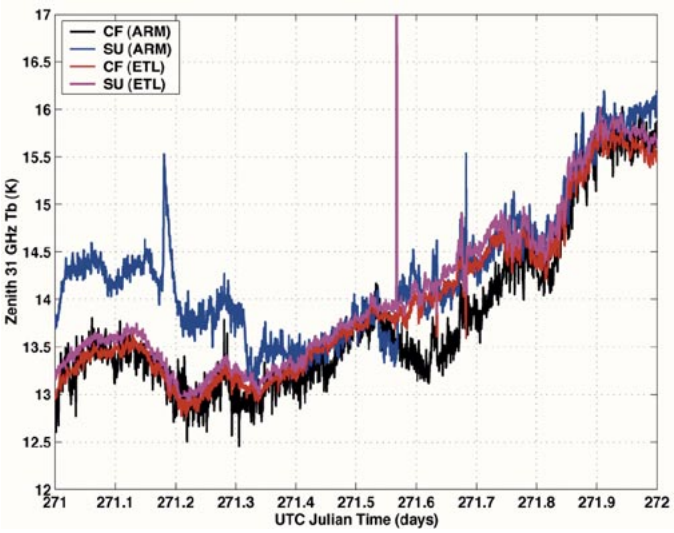


Fig. 3. Time series during Julian day 271 of 31.4-GHz Tb measured by ARM CF and SU, calibrated following the ARM (CF: black; SU: blue) and ETL procedures (CF: red; SU: magenta).

Note that since each tip curve takes about 50 seconds, to generate 1500 samples requires more than 24 h. Conversely, the ETL procedure determines the gain correction $f_W \cdot T_{\rm nd}$ by applying the tip curve algorithm to each set of observations taken at the ten elevation angles, and recalibrates the same observations according to the new value of $f_W \cdot T_{\rm nd}$. 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).

The ETL procedure was designed for clear conditions only, and would have to be extended when there are long intervals with clouds. On the other hand, the ARM procedure was designed for deployment during both clear and cloudy conditions, regardless of the applicability of tip curve calibration.

In Fig. 2, we show a 24-h time series of the CF unit 31-GHz tip curve correction coefficient $(f_W \cdot T_{\rm nd})$ as used in the ETL procedure, together with the $T_{\rm nd}$ values determined by each tip curve and by the linear regression on $T_{\rm Ref}$ in the ARM procedure. The offset between the instantaneous values resulting from ARM and ETL algorithms is due to the f_W factor.

The WVIOP2000 experiment gave us the opportunity to compare the ARM and ETL calibration procedures. In fact, during the postprocessing of data, we applied both procedures to the measurements from the two identical ARM CF and SU radiometers, to determine which procedure gave the best agreement between simultaneous observations. A typical example is shown in Fig. 3, in which we have a time series of Tb measurements at 31.4 GHz from CF and SU, as calibrated with the original ARM and the ETL procedures.

We notice that the Tb difference between CF and SU, as calibrated with the ARM procedure, is considerable, especially within [271 to 271.3] and [271.6 to 271.8] UTC Julian time, sometimes exceeding 2 K (at 271.18 UTC). Conversely, applying the ETL instantaneous calibration, we reduce such differences to some tenths of a degree. Also, the noise level appears to be significantly reduced. By comparing data calibrated with the two procedures, it is clear that the ARM procedure can

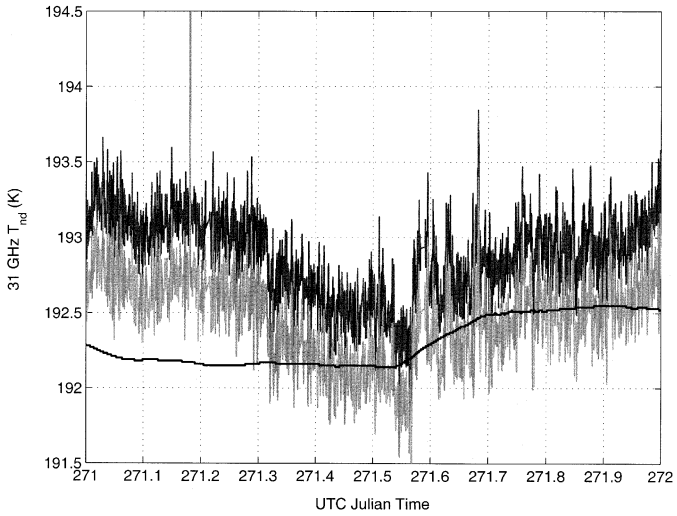


Fig. 4. Same as Fig. 2, but for SU instead of CF.

produce spurious data, e.g., with the SU during ([271 to 271.3] UTC), and then with the CF ([271.6 to 271.8]). Considering the $T_{\rm nd}$ time series during the same day (CF: Fig. 2, SU: Fig. 4), we notice that the departures are related to the ARM operational $T_{\rm nd}$, which cannot follow the instantaneous value, indeed between [271 to 271.3] UTC for SU and between [271.6 to 271.8] for CF.

In summary, using ARM original calibration, simultaneous measurements from identical radiometers differ from 0.5–2 K for about 12 of the 24 h shown. On the other hand, using ETL calibration procedure, the same measurements agree within 0.2 K for the entire 24 h.

For a statistical analysis over the entire experiment (21 days), we divided the time period into 5-min intervals, and for each interval we computed standard deviation and mean value of zenith Tb at both 23.8 and 31.4 GHz.

The signal standard deviation over 5-min intervals, assuming a stable atmosphere, gives an idea of the noise level. Thus, for

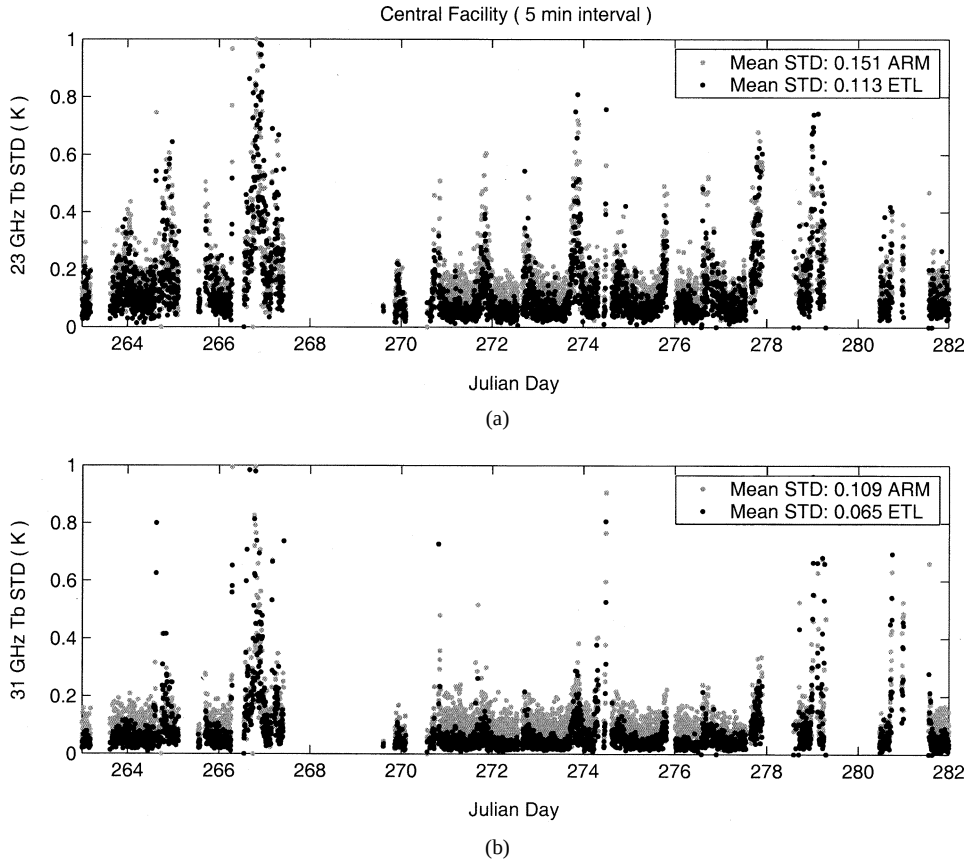


Fig. 5. Standard deviation over 5-min intervals for CF Tb [$\sigma_{\Delta t}^{\mathrm{CF}(f)}(i)$ as defined in (3)]. Results from data calibrated with ARM procedure are shown in gray, while with ETL procedure are shown in black. (a) 23-GHz Tb. (b) 31-GHz Tb.

each ARM unit (U) and frequency (f), we computed the standard deviation of Tb over the ith 5-min (Δt) interval, which contains $N(i)$ measurements:

$$\sigma_{\Delta t}^{U(f)}(i) = \mathrm{STD}\left(Tb_{\Delta t(i)}^{U(f)}\right) = \sqrt{\sum_{j=1}^{N(i)}\left(Tb_{\Delta t(i)}^{U(f)}(j) - \overline{Tb_{\Delta t(i)}^{U(f)}}\right)^2 \Big/ (N(i)-1)} \quad (3)$$

where the overbar means the averaged value of the $N(i)$ measurements falling inside the interval $\Delta t(i)$.

In Fig. 5, we plot the time series of $\sigma_{\Delta t}^{\mathrm{CF}(23.8\mathrm{GHz})}$ and $\sigma_{\Delta t}^{\mathrm{CF}(31.4\mathrm{GHz})}$ for the entire experiment, considering both ARM and ETL calibration procedures. As anticipated, we found that on the average the ETL procedure reduces $\sigma_{\Delta t}^{\mathrm{CF}(f)}$ by about 25% at 23.8 GHz and 40% at 31.4 GHz with respect to the ARM original values. The same considerations apply for the SU unit, although we do not show plots for brevity.

From the set of mean values over the 5-min intervals, we computed statistics of the comparison between CF and SU over the entire experiment. In Fig. 6, we plot CF against SU 31-GHz brightness temperature, as calibrated with the ARM and the ETL procedures, screened by a ± 5 K threshold to remove outliers and obstructed antenna episodes. For brevity, we introduce the difference between Tb measured by CF and SU at frequency f and for the ith interval Δt as $\Delta Tb_f(i) = \overline{Tb_{\Delta t(i)}^{\mathrm{CF}(f)}} - \overline{Tb_{\Delta t(i)}^{\mathrm{SU}(f)}}$. Fig. 6 shows that by applying the ETL calibration instead of the original ARM, we decrease the scatter of $\Delta Tb_{31\mathrm{GHz}}$. This is true also for the 23-GHz Tb: with respect to the ARM calibration, the ETL procedure reduces the standard deviation of ΔTb_f from 0.62– 0.42 K for the 23.8-GHz channels and from 0.45–0.19 K for the 31.4-GHz channels. Note that in Fig. 5 we show the standard deviation of CF Tb inside each 5-min bin $(\sigma_{\Delta t}^{\mathrm{CF}(f)})$, while in Fig. 6 we show the standard deviation over the entire experiment of the 5-min averaged Tb difference, i.e.,

$$\sigma_{Exp}^{f} = \mathrm{STD}\left(\Delta Tb_f(i)\right) = \sqrt{\sum_{i=1}^{N}\left(\Delta Tb_f(i) - \overline{\Delta Tb_f}\right)^2 \Big/ (N-1)} \quad (4)$$

where in this case the overbar means the average over the entire set of ΔTb_f (sample size equal to N).

We note that although for this analysis we considered data from the entire experiment, disregarding atmospheric conditions, the ARM operational procedure switches to zenith line-of-sight (LOS) mode when horizontal inhomogeneity is detected. Therefore, during moderate to heavy cloudy conditions, tip data were not available, and so we were not able to extend the ETL procedure for cloudy conditions.

Thus, during clear to light cloudy conditions, the ETL procedure improves the accuracy for Tb measurements with respect to the ARM operational calibration, either for short (5-min) or long (21-day) time scales, as summarized in Table II.

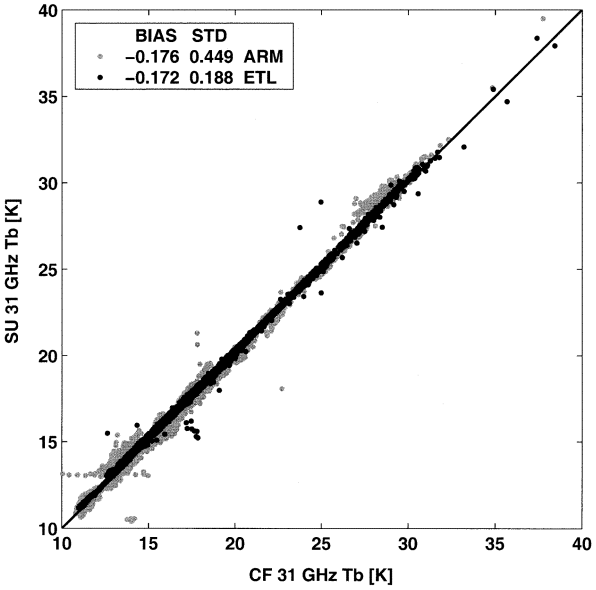


Fig. 6. Scatter plot of CF versus SU 31-GHz Tb. Data calibrated with ARM procedure are shown in gray, while with ETL procedure are shown in black.

TABLE II
STATISTICAL COMPARISON BETWEEN ARM AND ETL CALIBRATION PROCEDURES

	23.8 GHz Tb (K)		31.4 GHz Tb (K)	
	ARM	**ETL**	**ARM**	**ETL**
$\sigma_{\Delta I}^{U(f)}$	0.15	0.11	0.11	0.06
σ_{Exp}^{f}	0.62	0.42	0.45	0.19

IV. BRIGHTNESS TEMPERATURE COMPARISONS

To compare Tb measurements at different frequencies, we chose the ARM MWRs as a standard and predicted the equivalent Tbs at 23.8 and 31.4 GHz from JPL and CSR measurements at other frequencies. The prediction linear-fit coefficients were derived from a simulated database of clear-sky Tbs, computed with a radiative transfer model (RTM) using the Rosenkranz 1998 absorption model [20]. We divided the time series into 5-min bins and averaged each Tb (measured and predicted) inside every bin in order to set up our sample in a common grid. In computing the statistical comparisons between the different measurements, we restricted our sample to those time intervals in which all radiometers were working simultaneously. This screening reduces substantially our sample set, although it is necessary if we want to compare the overall agreement and distribution.

To consider only well-calibrated data, the quality of each tip curve needed to be checked against atmospheric inhomogeneity, which would have destroyed the linear relationship between slant path opacity and equivalent air mass. For each elevation angle θ, we estimated $\tau(\theta)$ from the measured Tb(θ) [4], and forced the $\tau(\theta)$ versus $\alpha(\theta)$ line of best fit to decrease to zero at zero air mass to establish calibration [16]. Then, we converted each $\tau(\theta)$ to an equivalent zenith Tb and computed the standard deviation (STD) relative to each tip curve. Assuming that the MWR's random noise is low, this STD of equivalent zenith Tb (EZTb) is a measure of the tip curve quality and atmospheric homogeneity. In Fig. 7, we show one-day time series of STD(EZTb) for CF and SU, for both the 23- and 31-GHz channels. It is evident that for most of the time the quality of tip curve was excellent, with an STD(EZTb) smaller than 0.3 K. Note that the 23 GHz STD(EZTb) for the SU was higher than for the CF: this was due to an instrumental problem, which will be discussed later in this section. Considering the CF as a standard reference during the entire experiment, more than 50% of the time, we had an STD(EZTb) smaller than 0.5 K, both for the 23- and 31-GHz channels.

Thus, we further restricted our analysis to measurements in which the tipping curve calibration had passed a quality control, considering only those cases in which the CF STD(EZTb) was found smaller than 0.5 K. This criterion reduced the Tb range to approximately 25 K for 23.8-GHz channels and 15 K for 31.4-GHz channels. The sample is reduced to 1177 5-min intervals, resulting in about 100 h of observations.

The comparison between measurements from different radiometers obtained using this sample set is shown in Table III in terms of average difference (BIAS), standard deviation (STD), and root-mean-square (RMS) of the difference. The STD was usually very good, ranging between 0.25 and 0.57 K, although sometimes we saw a large BIAS, which could reach 1.2 K. The largest values of BIAS were found in the 23.8-GHz ARM SU measurements. This problem was already noticed during the experiment. After the WVIOP2000, the SU was subjected to an in-house diagnosis by the manufacturer: it was found that the observed BIAS was caused by a misplacement of the sidelobes collar, which slightly interfered with the calibration process [30]. Based on the manufacturer's findings, we were able to apply a correction on SU calibration coefficients. We determined a single correction factor by comparing the SU and the CF gain coefficients: to ensure independence of the SU and CF data during the experiment, we used observations during eight days immediately before and after the WVIOP2000. The improvements obtained with such corrections are shown in Table IV: the BIAS between all of four radiometers is now limited within 0.5 K.

Assuming that the standard uncertainty is equivalent but uncorrelated among the four MWR units, this quantity can be estimated simply from RMS/$\sqrt{2}$. If we do not consider the SU 23-GHz measurements or adopt the gain correction for them, from Tables III and IV we can conclude that the standard uncertainty for the MWR units deployed during the WVIOP2000 ranged from 0.1–0.6 K for the 23.8-GHz channels, and from 0.1–0.4 K for the 31.4-GHz channels. Thus, the WVIOP2000 showed that the uncertainty for MWR calibrated with tip curve is better than 0.6 K. This value agrees with theoretical predictions found in [16].

V. RADIOSONDES AND RADIOMETERS

During WVIOP2000, balloons were launched every 3 h. To study the relative uncertainty, each balloon carried two independent packages, yielding a total of approximately 300 sets of atmospheric temperature (T), pressure (P), and relative humidity (RH) profiles.

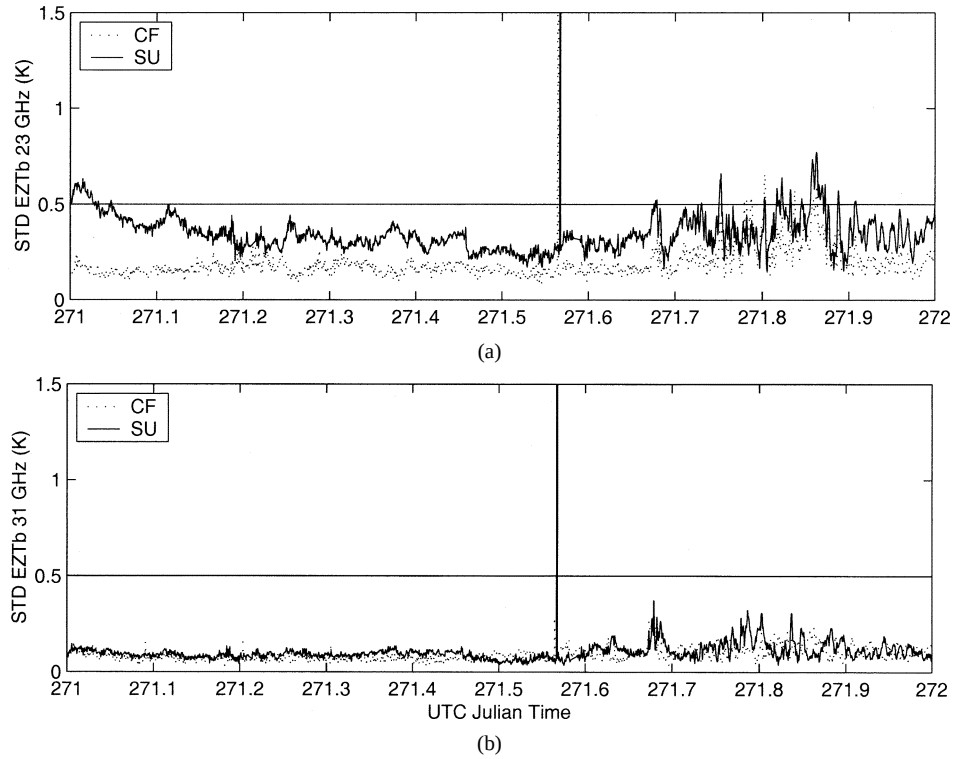


Fig. 7. Time series during Julian day 271 of standard deviation of equivalent zenith Tb (EZTb) for the (dashed) CF and (solid) SU MWRs. (a) 23-GHz channels. (b) 31-GHz channels.

TABLE III
STATISTICAL COMPARISON BETWEEN TB FROM THE SET OF MWR DEPLOYED DURING THE WVIOP2000

	23.8 GHz Tb (K)			31.4 GHz Tb (K)		
	BIAS	STD	RMS	BIAS	STD	RMS
CF SU	0.65	0.21	0.68	-0.14	0.09	0.17
CF JPL	-0.22	0.38	0.44	-0.44	0.24	0.50
CF CSR	-0.51	0.70	0.86	-0.55	0.29	0.62
SU JPL	-0.87	0.33	0.93	-0.29	0.24	0.38
SU CSR	-1.16	0.56	1.29	-0.40	0.26	0.48
CSR JPL	0.29	0.60	0.66	0.11	0.35	0.37

TABLE IV
STATISTICAL COMPARISON BETWEEN TB FROM THE SET OF MWR DEPLOYED DURING THE WVIOP2000, CONSIDERING SU POSTCALIBRATION

	23.8 GHz Tb (K)			31.4 GHz Tb (K)		
	BIAS	STD	RMS	BIAS	STD	RMS
CF SU	-0.01	0.09	0.09	-0.14	0.09	0.17
SU JPL	-0.21	0.38	0.44	-0.29	0.24	0.38
SU CSR	-0.50	0.71	0.87	-0.40	0.26	0.48

The radiosondes available were Vaisala RS80 and RS90 [21], which do not carry any sensor capable of measuring atmospheric liquid content. Thus, a comparison between radiometric measurements and radiosonde observations (RAOBs) was restricted to clear-sky cases only, as determined by cloud detection from a collocated ceilometer. To our knowledge WVIOP2000 was the first ARM field experiment in which RS90 sensors were deployed. RS90s represent the new generation of Vaisala radiosondes, which were designed to solve the "dry-bias" issue associated with the RS80 sensors [9], [11], [22]. From simultaneous sounding comparisons, we found the RS80 measurements of PWV dryer by about 0.1 cm in average with respect to the RS90 [23]. The RS90 sensors reduce the RAOB-to-RAOB RMS by more than 60% in PWV (from 0.135– 0.045 cm), although we have to consider that the RMS is influenced by the sample size, and, during the WVIOP2000, the number of RS90 was four times smaller than the number of RS80 [23].

To compare MWR observations and RAOBs, we compute downwelling Tb from the measured atmospheric profiles of T, P, and RH with an RTM using the Rosenkranz 1998 absorption model [20]. A more extensive comparison using different forward models is currently under study. We averaged MWR Tbs for the first half hour of the balloon ascension, in order to have temporal consistency during the passage through the first few kilometers in altitude, where most of the water vapor is distributed.

In Fig. 8, we show the results of this analysis, showing the scatter plot of Tbs as measured by MWR [Fig. 8(a): CF; 8(b): SU; 8(c): JPL; 8(d): CSR] versus simulations computed from simultaneous RAOB profiles. We distinguish the RAOB packages in RS80 and RS90, and we add the main statistical information, as BIAS, STD, and sample size. We also computed the least squares linear fit for each subsample, the slope (P1) and offset (P0) coefficients of which are also shown. Although the sample size for RS90 is much smaller compared to RS80, it appears that the new packages reduce substantially the dry bias. The linear fit analysis suggests the same considerations,

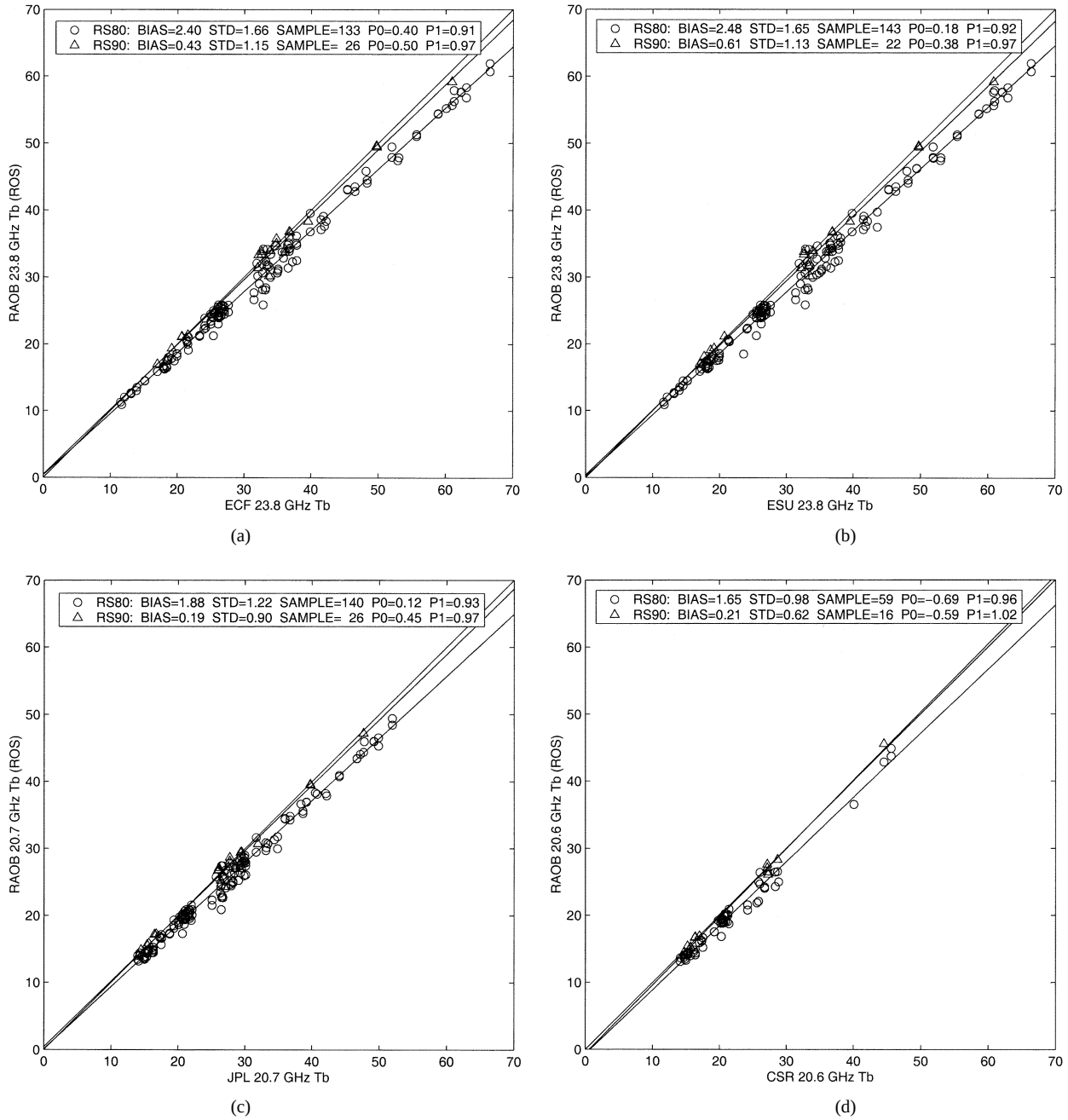


Fig. 8. Scatter plot of Tb measured by MWR versus Tb computed from RAOB. (a) ARM CF calibrated with ETL procedure (ECF). (b) ARM SU calibrated with ETL procedure (ESU). (c) JPL. (d) CSR. RAOBs are divided into RS80 (circles) and RS90 (triangles). The slope (P1) and offset (P0) coefficients from a least squares linear fit are also shown for each subsample.

showing a slope closer to one for RS90 than for RS80 sensors. These results are true for all of the four radiometers under study: the BIAS between MWR-measured and RAOB-computed Tb is at least three times smaller when using atmospheric profiles measured by RS90 instead of RS80 sensors. An extensive analysis on a more significant sample is in progress by other investigators [31].

VI. Summary, Discussion, and Future Plans

We have shown the performances of a set of MWRs working in the 20/30-GHz band during the WVIOP2000 field experiment. We demonstrated that the difference between simultaneous clear-sky measurements from two identical models of ARM MWRs can be reduced using the calibration procedure proposed by the ETL [16] instead of the ARM operational procedure [18], [19]. The improvements at any one time obtained by using the ETL procedure are significant, of the order of 1 K. The total improvement in STD of the difference is approximately 30% for the 23-GHz channel, and more than 50% for the 31-GHz channel. Using the ETL tip curve calibration procedure, the ARM radiometers agreed with an STD of 0.4 K at 23.8 GHz and 0.2 K at 31.4-GHz channels over a very large range of water vapor and atmospheric conditions.

Moreover, as seen previously during another experiment [24], the ETL procedure reduces the signal scatter within 5-min intervals with respect to the ARM operational calibration. A possible explanation of the noise reduction by the ETL procedure was given by Liljegren [32], who argued that the brightness temperature derived by applying the tip curve represents a kind of spatial average over the tip angles, which does not fully capture the actual zenith variations that occur even during clear-sky conditions. Thus, the ETL "instantaneous" calibration results are smoother than the ARM calibration results, which represent the noise diode calibration applied to the zenith measurement only. According to this, the ETL procedure reduces the scatter because it smoothes the atmospheric horizontal variability.

As an alternative explanation, we believe the ETL procedure reduces the instrumental noise level because it actually compensates for short time-scale gain fluctuations. The ARM operational procedure weighs the latest tip curve result with the recent past (24 h or more) and leaves T_{Ref} (internal black body target temperature) the only instantaneous parameter. To get the instantaneous calibration coefficient, the ARM procedure relies on the correlation between T_{nd} and T_{Ref}, which is actually poor (correlation coefficients between 0.2 and 0.55).

If we assume that the radiometer's gain factor changes, not just because of large temperature changes (such as seasonal or during the warm up), but also because of small effects (such as internal gradients caused by wind advection, cloud shade, or fan switching) then the radiometer output has to be considered as a nonstationary random process.

Being "instantaneous," the ETL calibration procedure accounts for all the gain fluctuations that are in the time scale of the order of a minute, which would be missed by any kind of long-term calibration factor. Missing these effects could be interpreted as additional atmospheric noise.

As already mentioned, the current ETL calibration is only applicable during horizontally homogenous sky conditions, whereas the ARM procedure is more generally applicable. Although it might seem restrictive to consider only clear-sky measurements, well-calibrated data are useful for the study of atmospheric absorption models. Moreover, clear-sky tip curves provide a substantial method to analyze the radiometer's performances and effective gain fluctuations.

Nevertheless, it is our intention to extend the ETL procedure to be applicable during both clear and cloudy conditions. Unfortunately, this was not possible during the WVIOP2000, since during moderate to heavy cloudy conditions the ARM radiometers switched to the LOS (zenith looking) mode, for which tip data were not available.

The LN2-based tests and Tb cross comparisons show that the theoretical prediction found in [16] of a calibration accuracy of 0.5 K for microwave radiometers calibrated with tip curve methods is realistic. Excluding the uncertainties related to the choice of the atmospheric absorption model, such calibration accuracy for direct measurements would provide an uncertainty in PWV estimates of about 0.035 cm during clear skies. However, this estimate of PWV uncertainty depends slightly on the absolute value of PWV, and we expect it to range from 0.02–0.04 cm from very dry to very humid conditions, respectively.

In terms of absolute accuracy, the remaining absorption model uncertainties are approximately $\pm 5\%$, which dominate the vapor measurement error budget, especially for high-humidity conditions [25]. Differences in 20–30-GHz Tb from recent models are of the order of 0.5–1.0 K in the arctic [26]–[28], while up to 3.5 K in the tropics [9]. Thus, presently, absorption model uncertainties are the main limiting factor to the retrieval of PWV from MWR [29].

We also analyzed the performances of the Vaisala new generation RS90 radiosondes, which, to our knowledge, were deployed for the first time in an ARM field experiment during the WVIOP2000. RAOB-RAOB and RAOB-MWR comparisons showed that the RS90 sensor represents a notable improvement with respect to the RS80 sensor. The RS90 RAOB-MWR comparisons show only a small residual (BIAS $= 0.05$ cm) of the typical "dry bias" signature that affects the RS80 sensors.

As a result of our analysis during the WVIOP2000, it is clear that the tip curve is a powerful method for the calibration and for the study of gain fluctuations for those channels in which the attenuation is low enough to allow its use. Moreover, measurements at different elevation angles provide a means to monitor the atmospheric homogeneity. Even in the presence of stratiform clouds, when the horizontal homogeneity is not completely lost, the tip curve method may still provide an acceptable calibration. In addition, when the horizontal homogeneity is destroyed by isolated clouds, it might be possible to detect and remove the affected measurements from the set of scanned angles and still apply the tip curve method to the remaining data.

Therefore, in order to have more accurate ground-based measurements in the 20–30-GHz band, we strongly recommend generating tipping data as often as the instrument permits, possibly on 1-min temporal scales, regardless of atmospheric conditions.

Since the WVIOP2000 experiment provided a large set of simultaneous and independent measurements of downwelling Tb and atmospheric thermodynamic profiles, our ongoing and near-future research will focus on absorption models comparisons [29], in order to determine which among the most used best fits the empirical data.

ACKNOWLEDGMENT

The contributions of D. Cimini formed a portion of his Ph.D. thesis at the University of L'Aquila. The authors would like to thank A. Yevgrafov for assistance in deploying the CSR unit. The authors would also like to thank F. Solheim, H. E. Revercomb, S. A. Clough, and P. E. Racette for useful and insightful comments on the manuscript.

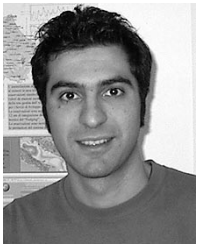
REFERENCES

[1] X. Q. Zou and Y. H. Kuo, "Incorporating SSM/I-derivated precipitable water and rainfall rate into a numerical model: A case study for the ERICA IOP-4 cyclone," *Mon. Weather Rev.*, vol. 128, pp. 87–108, 2000.

[2] R. Ware, F. Solheim, and F. Vandenberghe, "Advanced assimilation of thermodynamic and wind soundings," in *Proc. COST720 Meeting*, L'Aquila, Italy, June 19–21, 2002, [Online]. Available: http://www.cost720.rl.ac.uk/.

[3] S. A. Clough, "The water vapor continuum and its role in remote sensing," in *Optical Remote Sensing of the Atmosphere*, ser. Optical Society of America Technical Digest. Washington, DC: Opt. Soc. Amer., 1995, vol. 2, pp. 76–78.

[4] M. A. Janssen, *Atmospheric Remote Sensing by Microwave Radiometry*. New York: Wiley, 1993.
[5] Y. Han, J. A. Shaw, J. H. Churnside, P. D. Brown, and S. A. Clough, "Infrared spectral radiance measurements in the tropical Pacific atmosphere," *J. Geophys. Res.*, vol. 102, no. D4, pp. 4353–4356, 1997.
[6] F. Guichard, D. Parsons, and E. Miller, "Thermodynamical and radiative impact of the correction of sounding humidity bias in the tropics," *J. Climate*, vol. 13, no. 20, pp. 3611–3624, 2000.
[7] G. M. Stokes and S. E. Schwartz, "The Atmospheric Radiation Measurement (ARM) program: Programmatic background and design of the cloud and radiation testbed," *Bull. Amer. Meteorol. Soc.*, vol. 75, pp. 1201–1221, 1994.
[8] T. Ackerman and J. Stokes, "The atmospheric radiation measurement program," *Phys. Today*, vol. 56, no. 1, pp. 38–44, 2003.
[9] E. R. Westwater, B. B. Stankov, D. Cimini, Y. Han, J. A. Shaw, B. M. Lesht, and C. N. Long, "Radiosonde humidity soundings and microwave radiometers during Nauru99," *J. Atmos. Ocean. Technol.*, vol. 20, no. 7, pp. 953–971, July 2003.
[10] P. E. Racette, E. Westwater, Y. Han, W. Manning, A. Gasiewski, and D. Jones, "Millimeter-wave measurements of low amounts of precipitable water vapor," in *Proc. IGARSS*, July 24–28, 2000, pp. 1154–1156.
[11] H. E. Revercomb, D. D. Turner, D. C. Tobin, R. O. Knuteson, W. F. Feltz, J. Bannard, J. Bosenberg, S. Clough, D. Cook, R. Ferrare, J. Goldsmith, S. Gutman, R. Halthorne, B. Lesht, J. Liljegren, H. Linne, J. Michalsky, V. Morris, W. Porch, S. Richardson, B. Schmid, M. Splitt, T. Van Hove, E. Westwater, and D. Whiteman, "The ARM program's water vapor intensive observation periods: Overview, initial accomplishments, and future challenges," *Bull. Amer. Meteorol. Soc.*, vol. 84, no. 2, pp. 217–236, 2003.
[12] S. A. Clough, P. D. Brown, J. C. Liljegren, T. R. Shippert, D. D. Turner, R. O. Knuteson, H. E. Revercomb, and W. L. Smith, "Implications for atmospheric state specification from the AERI/LBLRTM quality measurement experiment and the MWR/LBLRTM quality measurement experiment," in *Proc. of 6th ARM Science Team Meeting*, San Antonio, TX, Mar. 4–7, 1996, [Online]. Available: http://www.arm.gov/docs/documents/technical/conf_9603/cloug_96.pdf.
[13] D. C. Hogg, F. O. Guiraud, J. B. Snider, M. T. Decker, and E. R. Westwater, "A steerable dual-channel microwave radiometer for measurement of water vapor and liquid in the troposphere," *J. Appl. Meteorol.*, vol. 22, no. 5, pp. 789–806, May 1983.
[14] M. A. Janssen, "A new instrument for the determination of radio path delay due to atmospheric water vapor," *IEEE Trans. Geosci. Remote Sensing*, vol. 4, no. GE-23, pp. 485–490, 1985.
[15] R. Ware and F. Solheim, *WVR-1100. Total Integrated Water Vapor and Liquid Water Radiometer*. Boulder, CO: Radiometrics Corp..
[16] Y. Han and E. R. Westwater, "Analysis and improvement of tipping calibration for ground-based microwave radiometers," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 1260–1276, May 2000.
[17] F. Solheim, "An inexpensive, accurate microwave target," in *microCal, Proc. 1st Int. Microwave Radiometer Calibration Workshop*, College Park, MD, 2000.
[18] J. C. Liljegren, "Automatic self-calibration of ARM microwave radiometers," in *Proc. 9th ARM Science Team Meeting*, San Antonio, TX, March 22–26, 1999, [Online]. Available: http://www.arm.gov/docs/documents/technical/conf_9903/liljegren(1)-99.pdf.
[19] J. C. Liljegren, "Automatic self-calibration of ARM microwave radiometers," in *Microwave Radiometry and Remote Sensing of the Earth's Surface and Atmosphere*, P. Pampaloni and S. Paloscia, Eds: VSP, 2000, pp. 433–441.
[20] P. W. Rosenkranz, "Water vapor microwave continuum absorption: A comparison of measurements and models," *Radio Sci.*, vol. 33, no. 4, pp. 919–928, 1998.
[21] O. Y. Vaisala, "RS90 radiosondes information release,", Vaisala, Helsinki, Finland, Tech. Rep. R634en-2.0. [Online]. Available: http://www.vaisala.com/, 2000.
[22] B. M. Lesht, "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," in *Proc. 9th ARM Science Team Meeting*, San Antonio, TX, Mar. 22–26, 1999, [Online]. Available: http://www.arm.gov/docs/documents/technical/conf_9903/lesht-99.pdf.
[23] D. Cimini, E. R. Westwater, and B. Lesht, "Evaluation of the improvements in humidity sounding by balloon-borne sensors," in *Proc. COST720 Meeting*, L'Aquila, Italy, June 19–21, 2002, [Online]. Available: http://www.cost720.rl.ac.uk/.
[24] Y. Han, E. R. Westwater, P. E. Racette, M. Klein, A. Gasiewski, W. Manning, and B. M. Lesht, "Radiometric observations of water vapor during the 1999 arctic winter experiment," in *Proc. 10th ARM Science Team Meeting*, San Antonio, TX, Mar. 13–17, 2000, [Online]. Available: http://www.arm.gov/docs/documents/technical/conf_0003/han-y.pdf.
[25] S. J. Keihm, Y. Bar-Sever, and J. C. Liljegren, "WVR-GPS Comparison measurements and calibration of the 20-32 GHz tropospheric water vapor absorption model," *IEEE Trans. Geosci. Remote Sensing*, vol. 40, pp. 1199–1210, June 2002.
[26] E. R. Westwater, Y. Han, M. D. Shupe, and S. Y. Matrosov, "Analysis of integrated cloud liquid and precipitable water vapor retrievals from microwave radiometers during SHEBA," *J. Geophys. Res.*, vol. 16, no. 23, pp. 32 019–32 030, 2001.
[27] E. R. Westwater, P. E. Racette, and D. Cimini, "The arctic winter millimeter-wave radiometric experiment: Summary, conclusion and recommendations," in *Proc. 11th ARM Science Team Meeting*, Mar. 19–23, 2001, [Online]. Available: http://www.arm.gov/docs/documents/technical/conf_0103/westwater-er.pdf.
[28] P. E. Racette, E. R. Westwater, Y. Han, A. Gasiewski, M. Klein, D. Cimini, W. Manning, E. Kim, J. Wang, and P. Kiedron, "Measuring low amounts of precipitable water vapor using millimeter-wave radiometry," *J. Atmos. Oceanic Technol.*, 2003, submitted for publication.
[29] D. Cimini, E. R. Westwater, S. J. Keihm, Y. Han, F. S. Marzano, and P. Ciotti, "Empirical evaluation of four microwave radiative forward models based on ground-based radiometer data near 20 and 30 GHz," in *Proc. IGARSS*, Toulouse, France, July 21–25, 2003.
[30] F. Solheim, private communication, 2002.
[31] B. M. Lesht, private communication, 2002.
[32] J. C. Liljegren, private communication, 2000.



Domenico Cimini (M'03) was born in Teramo, Italy, in 1973. He received the M.S. and Ph.D. in degrees from the University of L'Aquila, L'Aquila, Italy, in 1998 and 2002, respectively, both in physics.

Between 1999 and 2001, he spent 18 months with the Environmental Technology Laboratory (ETL), National Oceanic and Atmospheric Administration (NOAA), Boulder, CO, working on radiometer calibration techniques, remote sensing of the ocean surface, and microwave radiative transfer model comparisons. During this period, he was a Guest Scientist with the Water Vapor Intensive Operational Period (WVIOP2000). Currently, he is with the Remote Sensing Division of the Center of Excellence CETEMPS, University of L'Aquila, working on ground- and satellite-based passive microwave and infrared radiometry.

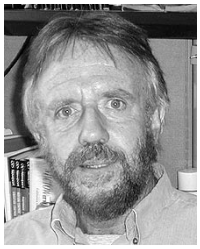
Ed R. Westwater (SM'91–F'01) was born in Denver, CO, in 1937. He received the B.A. degree in physics and mathematics from Western State College of Colorado, Gunnison, in 1959, and the M.S. and Ph.D. degrees from the University of Colorado, Boulder, in 1962 and 1970, respectively, both in physics

He is currently a Research Associate with the Cooperative Institute for Research in the Environmental Sciences (CIRES), University of Colorado/National Oceanic and Atmospheric Administration (NOAA), Boulder, CO, and is associated with the Microwave System Development Division of the Environmental Technology Laboratory (ETL), NOAA, Boulder, CO. He was employed by the U.S. Department of Commerce, Washington, DC, from 1960 to 1995, and joined CIRES in 1995. His research has been concerned with microwave absorption in the atmosphere, remote sensing of the ocean surface, microwave and infrared radiative transfer, ground- and satellite-based remote sensing by passive radiometry, and in the application of mathematical inversion techniques to problems in remote sensing. He was an Associate Editor of *Radio Science* (1998–2002).

Dr. Westwater is a member of the American Meteorology Society, the American Geophysical Union, the Mathematical Association of America, and was Chairman of URSI Commission F (2000 to 2002). He presented the American Meteorological Society Remote Sensing Lecture in 1997 and received the 15th V. Vaisala Award from the World Meteorological Society in 2001. In 2003, he received the Distinguished Achievement Award from the IEEE Geoscience and Remote Sensing Society.

Yong Han was born in Shanghai, China. He received the B.S. and M.S. degrees in 1982 and 1985, respectively, from the Nanjing Institute of Meteorology, Najing, China, both in atmospheric science, and the Ph.D. degree in meteorology from the Pennsylvania State University, University Park, in 1992.

He is currently a Physical Scientist with the National Oceanic and Atmospheric Administration (NOAA)/National Environmental Satellite, Data and Information Service, Boulder, CO. He was a Postdoctoral Research Associate with the National Research Council from 1992 to 1994. He was with the University of Colorado/Cooperative Institute for Research in the Environmental Sciences/NOAA Environmental Technology Laboratory from 1994 to 2001, and at the Science Systems and Applications, Inc./NASA Goddard Space Flight Center, Greenbelt, MD, in 2002. His research includes radiometer calibrations, radiative transfer modeling, mathematical retrieval methods, and remote sensing of atmosphere and the ocean surface from both ground- and satellite-based sensors.



Stephen J. Keihm received the B.S. degree in physics from Fordham University, New York, in 1968, the B.S. degree in mechanical engineering from Columbia University, New York, in 1969, and the M.S. degree in astronautical science from Stanford University, Stanford, CA, in 1970.

From 1970 to 1978, he was a Research Assistant, then as a Staff Associate, with the Lamont-Doherty Geological Observatory, Columbia University. While at Lamont, he served as Co-Investigator for the Apollo 15 and 17 lunar heat flow experiments, and Principal Investigator in studies of the thermal and electrical properties of the lunar regolith. In 1978, he joined the Planetary Science Institute, Pasadena, CA, where he conducted theoretical studies for the interpretation of data from the remote sensing of planetary surfaces. Since 1982, he has been with the Jet Propulsion Laboratory (JPL), California Institute of Technology, Pasadena. At JPL, he developed lunar microwave calibration models for the Cosmic Background Explorer (COBE) experiment and worked extensively in the areas of algorithm development and data interpretation for the earth-based, aircraft, and satellite microwave measurements of the atmosphere and sea surface. Currently, he is Instrument Scientist for the TOPEX and Jason-1 microwave radiometers, as well as for the advanced ground-based water vapor radiometer utilized for Cassini Radio Science tropospheric calibration.