Ground-Based Microwave Radiometer Measurements and Radiosonde Comparisons During the WVIOP2000 Field Experiment

D. Cimini
University of L'Aquila
L'Aquil, Italy

E. R. Westwater
Cooperative Institute for Research in the Environmental Sciences
University of Colorado
National Oceanic and Atmospheric Administration
Environmental Technology Laboratory
Boulder, Colorado

Y. Han
Science System Applications
National Aeronautics Space Administration
Goddard Space Flight Center
Greenbelt, Maryland

S. Keihm
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Introduction

During September to October 2000, a water vapor intensive operational period (WVIOP) was conducted at the Atmospheric Radiation Measurement (ARM) Program's Southern Great Plains (SGP) Cloud and Radiation Testbed site near Lamont, Oklahoma.

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-sky, and also a wide range of integrated water vapor amounts. To supplement the routine observing capabilities during the WVIOP2000, a powerful array of tools 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 microwave radiometers (MWRs).

Among the WVIOP2000 activities, we were particularly interested in evaluating the microwave radiometers absolute 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 absolute accuracy in direct measurements and derived products.

Instruments and Calibration

The set of radiometers under study is composed of three dual-channel units and one three-channel unit, each measuring downwelling Tb in the 20/30 GHz spectral region. Radiometers operating in this frequency band are commonly used to estimate PWV. ARM deployed two identical dual-channel (23.8 and 31.4 GHz) radiometers, located a few meters apart, called hereafter as the Central Facility (CF) and Spare Unit (SU). The National Aeronautics Space Administration (NASA) Jet Propulsion Laboratory (JPL) participated with a three channel radiometer (20.7, 22.2, 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, and this is actually desirable, because it provides a good opportunity to study calibration.

During the experiment, we performed several absolute accuracy tests on ARM MWRs, using a high emissivity absorber cooled in a liquid nitrogen (LN2) bath. The LN2 boiling temperature is considered stable at 77 K, while the absorber emissivity, the box attenuation, and the atmospheric pressure are taken into account to compute the expected brightness temperature of the target (Solheim 2000). We occasionally placed this target for tens of minutes 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 field-of-view. Hence, during these observations the two channels should measure the same brightness temperature, whose value should be close to the expected temperature. Results from this analysis are shown in Figure 1. Although the agreement is usually satisfactory, there are some interesting differences. Usually the agreement between two channels is better than the agreement with the expected value, which is usually warmer than the measurements. This might be related to LN2 vapors leaking in the interface at the bottom of the styrofoam box. Indeed, as suggested in (Solheim 2000), 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 temperature. Nevertheless, the differences between Tbs at the two channels usually remained within 0.7 K.

The MWRs ran simultaneously for 21 days, scanning continuously in the east-west direction. Each radiometer uses a slightly different automatic calibration procedure, using a combination of noise diode sources and internal and external reference targets. Moreover, for each radiometer we apply a tipping curve correction (Han and Westwater 2000) to refine an initial gain parameter value.

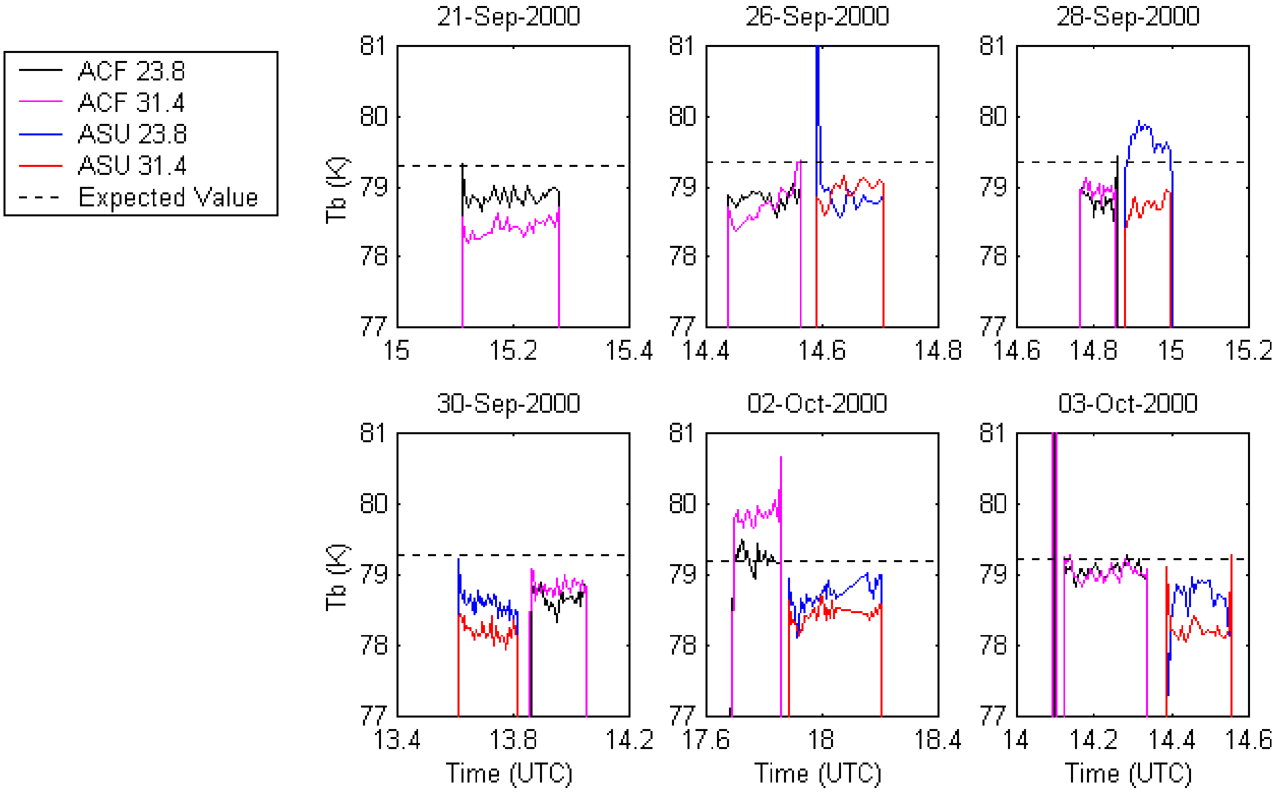

Figure 1. LN2-based absolute calibration tests performed on ARM CF and SU. The expected value has been computed according to Solheim (2000).

Two alternative tipping curve procedures are described in (Liljegren 1999), which corresponds to the ARM operational calibration, and in (Han and Westwater 2000), which was proposed by ETL. 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. 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 1999), the ARM operational algorithm estimates the value of the noise diode temperature (T_{nd}) from each set of measurements taken at ten different elevation angles (ranging from 1 to 3 air masses); in order to determine a linear relationship between T_{nd} and the reference black body target physical temperature (T_{Ref}), this procedure retains a long time history (between 1500 and 3000) 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, since each tip curve takes about 50 seconds, to collect 1500 samples requires more than 24 hours. Conversely, the ETL procedure determines the gain correction $f_W \cdot T_{nd}$ (f_W is only a dimensionless constant factor very close to one) by applying the tip curve calibration to each set of observations taken at 10 different elevation angles, and recalibrates the same observations according to the new value of $f_W \cdot T_{nd}$. Therefore, the ETL procedure is also called “instantaneous” calibration.

It is worth to mention that the current ETL calibration is only applicable during horizontally homogenous sky conditions. This is a limit with respect to the ARM procedure, which was designed for being deployed regardless of the applicability of tip curve calibration. Although considering only clear-sky measurements might seem restrictive, well-calibrated data are useful for the atmospheric absorption models study. Moreover, clear-sky tip curves provide a solid way to analyze the radiometer's performances and effective gain fluctuations. Nevertheless, it is our intention to extend the ETL procedure in order to be applicable during both clear and cloudy conditions. Unfortunately, this was not possible during the WVIOP2000, since during cloudy conditions the ARM radiometers switched to the line-of-sight (LOS) zenith looking mode, for which tip data were not available.

The WVIOP2000 experiment gave us the opportunity to compare the ARM and ETL calibration procedures. In fact, during the post-processing 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 Figure 2, in which we have a time series of measurements at 31.4 GHz from CF and SU, as calibrated with the ARM original and the ETL procedures. It can be seen that using the ARM procedure leads to differences from 0.5 K to 2 K for about 12 of the 24 hours shown. Conversely, using the ETL procedure, the difference is reduced to 0.2 K for all of the 24 hours. Also, the noise level seems to be sensibly reduced.

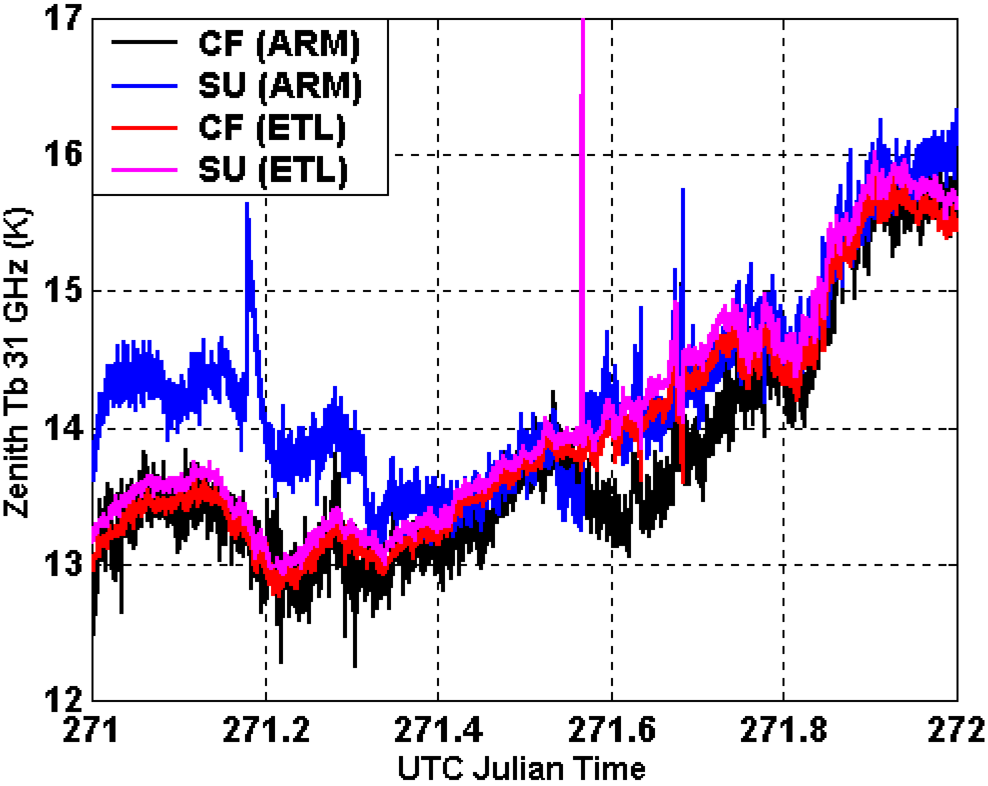

Figure 2. Time series of 31.4 GHz Tb measured by ARM CF and SU, calibrated following the ETL procedure (CF is shown in red, SU in magenta) and the ARM procedure (CF is shown in black, SU in blue).

For a statistical analysis over the whole experiment (21 days), we divided the time period into 5-minute intervals and for each interval, we computed standard deviation and mean value of Tb measured at different channels. The signal standard deviation over 5-minute intervals, assuming the atmosphere to remain constant, gives an idea of the noise level; we found that the ETL procedure reduces this quantity by about 30% at 23.8 GHz and 50% at 31.4 GHz with respect to ARM calibration. From the mean values, we computed the standard deviation of the difference between Tbs measured by identical channels, finding that, using ETL instead of the ARM procedure, it decreases from 0.63 K to 0.41 K for the 23.8 GHz channels and from 0.45 K to 0.18 K for the 31.4 GHz channels.

Concluding, during horizontally homogeneous atmospheric 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 days) time scales.

Brightness Temperature Comparisons

To compare Tb measurements at different frequencies, we choose the ARM MWRs as a standard and predict the equivalent Tbs at 23.8 and 31.4 GHz from JPL and CSR measurements at other frequencies. The prediction linear-fit coefficients are derived from a simulated database of clear-sky Tb, computed with a Radiative Transfer Model (RTM) using the Rosenkranz 1998 absorption model (Rosenkranz 1998). We divided the time series in five-minute 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 comparison between the different measurements, we restrict 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.

In order to consider only well-calibrated data, the quality of each tip curve needs to be checked against atmospheric inhomogeneity, which would destroy the linear relationship between slant path optical depth (τ) and equivalent air mass (M). For each angle, we divided τ by M, obtaining an equivalent zenith Tb; we then computed the standard deviation (STD) relative to set of ten angles forming each tip curve. This STD of Equivalent Zenith Tb (EZTb) is a measure of the tip curve quality and atmospheric homogeneity. Thus, we further restrict our sample to measurements in which the tipping curve calibration has passed the quality control STD (EZTb) <0.5 K. This criterion reduces the Tb range to roughly 25 K for 23.8 GHz channels and 15 K for 31.4 channels.

The comparison between measurements from different radiometers obtained using this sample set is described in Table 1 in terms of average difference (BIAS), standard deviation (STD) and root-mean-square (RMS) of the difference. The STD is usually very good, ranging between 0.25 and 0.57 K, although sometimes we see 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 has been 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 interferes in the calibration process (Solheim 2001).

Table 1. Statistical Comparison Between Tb from Different Radiometers (N = 1177)

		23.8 GHz Tb (K)			31.4 GHz Tb (K)		
		BIAS	**STD**	**RMS**	**BIAS**	**STD**	**RMS**
CF	**SU**	0.62	0.17	0.64	-0.15	0.08	0.17
CF	**JPL**	-0.27	0.36	0.45	-0.43	0.23	0.49
CF	**CSR**	-0.63	0.53	0.82	-0.56	0.27	0.62
SU	**JPL**	-0.89	0.32	0.94	-0.28	0.23	0.36
SU	**CSR**	-1.25	0.44	1.32	-0.42	0.25	0.48
CSR	**JPL**	0.36	0.47	0.59	0.13	0.33	0.36

With the exception of the SU 23 GHz measurements, we can say that the radiometers' absolute accuracy is typically 0.5 K or better. This value agrees with theoretical predictions found in (Han and Westwater 2000).

Radiosondes and Radiometers

During WVIOP2000, balloons were launched usually every three hours. For an absolute accuracy study, each balloon carried two independent packages, yielding approximately 300 sets of atmospheric temperature (T), pressure (P), relative humidity (RH) and height profiles.

The radiosondes available were Vaisala RS80 and RS90, 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 co-located 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 are supposed to solve the "dry-bias" issue associated with the RS80 sensors (Lesht 1999). From simultaneous sounding comparisons, we found a mean BIAS of about 0.1 cm in PWV between RS90 and RS80 sensors. The RS90 sensors seem to reduce the RAOB-to-RAOB RMS by more than 60% in PWV (from 0.135 to 0.045 cm), although we have to consider that the sample size is different. In order to compare MWR measurements and RAOBs, we compute down welling Tb from the atmospheric profiles of T, P, RH with a RTM using the Rosenkranz 1998 (Rosenkranz 1998) absorption model. A more extensive comparison using different forward models is currently under study. We averaged measured Tbs for the first half hour of the balloon ascension, in order to have temporal consistency during the passage through the first few kilometers, where most of the water vapor is distributed.

Figure 3 presents an example of this analysis, showing the scatter plot between Tbs measured by the JPL radiometer and that computed from simultaneous RAOB profiles. We distinguish the RAOB packages in RS80 (blue circle) and RS90 (red triangles), and we include the main statistical information, as BIAS, STD and sample size. We also added a least-squares linear fit for each sub-sample (blue for RS80 and red for RS90), whose slope (P1) and offset (P0) coefficients are also shown. Although the sample size for RS90 is much smaller compared to RS80, it appears that the new packages reduce sensibly the dry bias. The linear fit analysis suggests the same considerations, showing a slope closer to one for RS90 than for RS80 sensors. We found that these considerations are true for all of the four radiometers under

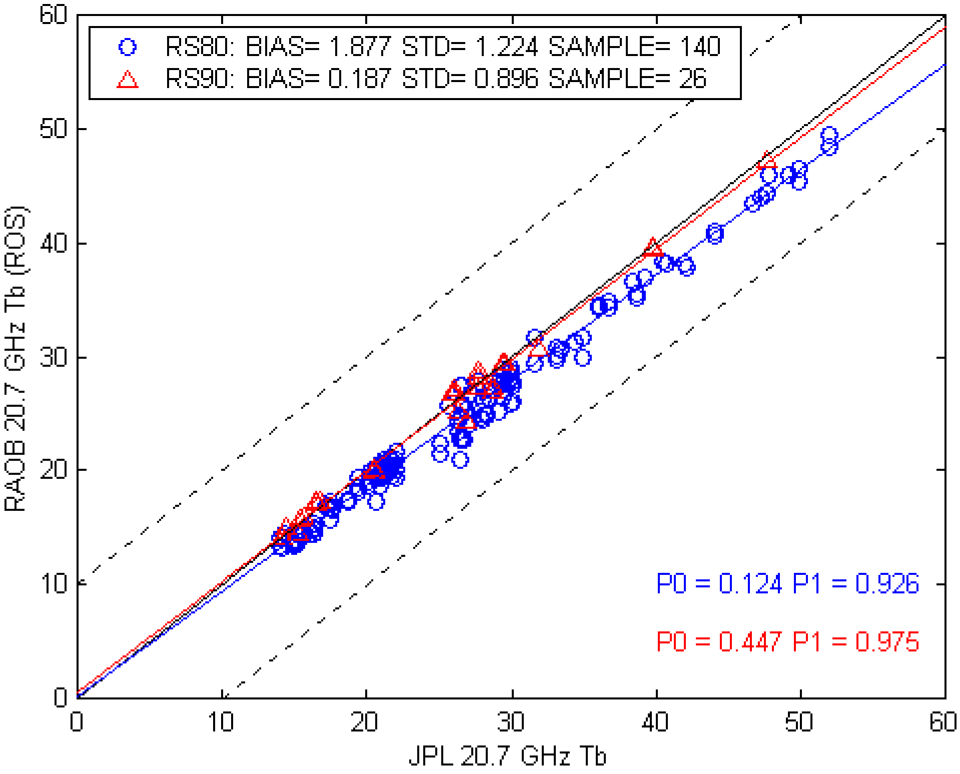

Figure 3. 20.7 GHz Tb measured by JPL vs. computed from RAOBs. Simulations from RAOB measurements are shown in blue circles if the sensors were RS80, while in red triangles if the sensors were RS90. Sample size, BIAS, STD, and linear fit coefficients (P0: offset, P1: slope) are also shown.

study: the BIAS between 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 will be available by other investigators (Lesht 2001).

Summary 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 measurements from the two identical ARM MWRs can be reduced using the calibration procedure proposed by ETL (Han and Westwater 2000) instead of the ARM operational procedure (Liljegren 1999). The improvements at any single time obtained by using the ETL procedure are significant, of the order of 1 K. The total improvement in STD is approximately 30% for the 23 GHz channel, and about 50% for the 31 GHz channel. Using ETL tip curve calibration procedure, the ARM radiometers agreed with a STD of 0.4 K at 23.8 GHz and 0.2 K at 31.4 GHz channel over a very large range of water vapor and atmospheric conditions. Moreover, as seen before during another experiment (Han 2000), the ETL procedure reduces the signal scatter within 5-minute intervals with respect to the ARM operational calibration. As a possible 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 weights the latest tip curve result with the recent past (24 hours or more) and leaves the only instantaneous parameter being the T_{Ref}. 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). 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.

The LN2-based absolute accuracy tests and Tb cross-comparisons show that the theoretical prediction of an absolute accuracy of 0.5 K for microwave radiometers (Han and Westwater 2000) is realistic. Excluding the uncertainties related to the choice of the atmospheric absorption model, such an absolute accuracy for direct measurements would provide an accuracy in PWV estimates of about 0.04 cm.

We also analyzed the performances of the Vaisala new generation RS90 radiosondes, which, to our knowledge, were deployed for the first time in a 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-radiometer comparisons show only a small residual (BIAS = 0.05 cm) of the typical "dry bias" signature that affects RS80 sensors (Rosenkranz 1998).

As a result of our analysis during the WVIOP2000, it is clear that the tip curve is a powerful method for the calibration and the study of gain fluctuations for those channels in which the attenuation is low enough to allow using it. Moreover, measurements at different elevation angles provide a way to monitor the atmospheric homogeneity. Even in presence of stratiform clouds, when the horizontal homogeneity is not completely lost, the tip curve method may still provide an acceptable calibration. Even 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 tip curve method to the remaining data. Therefore, it is our intention to extend the ETL procedure in order to be used during both clear and cloudy conditions.

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

Since the WVIOP2000 experiment provided a large set of simultaneous and independent measurements of down-welling Tb and atmospheric thermodynamic profiles, our ongoing and near-future research focus on microwave absorption models comparisons, in order to determine which one fits the empirical data best.

Acknowledgments

The work presented in this paper was sponsored by the Environmental Science Division of the Department of Energy as a part of their Atmospheric Radiation Program. Part of this work has been carried out when D. Cimini was a Ph.D. student at University of L'Aquila, Italy, under European Social Funding (FSE) Program. The authors also thank Fredrick Solheim, Henry E. Revercomb, and Shepard A. Clough for their useful and insightful comments.

Corresponding Author

D. Cimini, nico.cimini@aquila.infn.it

References

Han, Y., and E. R. Westwater, 2000: Analysis and improvement of tipping calibration for ground-based microwave radiometers. *Trans. Geosci. Rem. Sens*., **38** (3), 1260-1276.

Han Y., E. R. Westwater, P. E. Racette, M. Klein, A. Gasiewski, W. Manning, and B. M. Lesht, 2000: Radiometric Observations of Water Vapor During the 1999 Arctic Winter Experiment. Proc. Tenth ARM Science Team Meeting, 2000.

Lesht, B. M., 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 ARM Science Team Meeting.

Lesht, B. M., 2001: Personal communication.

Liljegren, J. C., 1999: Automatic Self-Calibration of ARM Microwave Radiometers, Proc. of Ninth ARM Science Team Meeting, 1999.

Rosenkranz, P. W., 1998: Water vapor microwave continuum absorption: a comparison of measurements and models. *Radio Sciences*, **33** (4), 919-928.

Solheim, F., 2000: An inexpensive, accurate microwave target, "microCal", First Intern. Microwave Radiometer Calibration Workshop, 2000.

Solheim, F., 2001: Personal communication.