Analysis of Radiosonde and ground-based remotely sensed PWV data from the 2004 North Slope of Alaska Arctic Winter Radiometric Experiment

V. Mattioli[1, 2], E. R. Westwater[3], D. Cimini[3, 4], J. S. Liljegren[5], B. M. Lesht[5], S. I. Gutman[6], and F. J. Schmidlin[7]

[1] Dipartimento di Ingegneria Elettronica e dell'Informazione, Università di Perugia via G. Duranti 93, 06125 Perugia, Italy

[2] Science and Technology Corporation, 10 Basil Sawyer Drive, Hampton, VA 23666, USA

[3] Cooperative Institute for Research in Environmental Sciences, University of Colorado/NOAA-ESRL Physical Science Division, 325 Broadway, Boulder, CO 80305, USA

[4] Institute of Methodologies for Environmental Analysis, National Research Council, C. da S. Loja - Zona Industriale, I-85050 Tito Scalo (PZ), Italy

[5] DOE/Argonne National Laboratory, Bldg 203, 9700 South Cass Avenue, Argonne, IL 60439, USA

[6] NOAA/Forecast Systems Laboratory, 325 Broadway, Boulder, CO 80305, USA

[7] NASA/Goddard Space Flight Center, Wallops Flight Facility, Wallops Island, VA 23337, USA

Submitted: October 14, 2005

Revised: April 15, 2006

Revised #2: July 3, 2003

Corresponding author: Vinia Mattioli, DIEI, Università di Perugia, via G. Duranti 93, 06125 Perugia, Italy, Ph: 011-39-075-585-3660, Fax: 011-39-075-585-3654, Email: vinia.mattioli@diei.unipg.it

Abstract

During March 9-April 9, 2004, the North Slope of Alaska Arctic Winter Radiometric Experiment was conducted at the Atmospheric Radiation Measurement (ARM) Program's "Great White" field site near Barrow, Alaska. The major goals of the experiment were to compare microwave and millimeter wavelength radiometers and to develop forward models in radiative transfer, all with a focus on cold (temperature 0 to -40ºC) and dry (Precipitable Water Vapor– PWV < 0.5 cm) conditions. To supplement the remote sensors, several radiosonde packages were deployed: Vaisala RS90 launched at the ARM Duplex and at the Great White, and Sippican VIZ-B2 operated by the NWS. In addition, eight dual-radiosonde launches were conducted at the Duplex, with Vaisala RS90 and Sippican GPS Mark II, the latter one modified to include a chilled mirror humidity sensor. Temperature comparisons showed a nighttime bias between the VIZ-B2 and the RS90, which reached 3.5ºC at 30 hPa. Relative humidity comparisons indicated better than 5% average agreement between the RS90 and the chilled mirror. A bias of about 20% for the upper-troposphere was found in the VIZ-B2 and the Mark II measurements relative to both the RS90 and the chilled mirror.

Comparisons in PWV were made between a Microwave Radiometer, a Microwave Profiler, a Global Positioning System receiver, and the radiosonde types. RMS agreement of 0.034 cm was found between the radiometer and the profiler, and better than 0.058 cm between the radiometers and the GPS. RS90 showed a daytime dry bias on PWV of about 0.02 cm.

1. Introduction

Although many years of research and experiments have focused on radiosonde measurements of humidity, many recent experiments have been conducted, primarily because of the importance of humidity to modeling of radiative transfer (Clough et al. 1999; Revercomb et al. 2003; Ferrare et al. 2004). In forward model studies, calculations based on radiosondes are compared to both infrared and microwave radiometer observations (Westwater 1997; Westwater et al. 2003; Liljegren et al. 2005; Mattioli et al. 2005a; Hewison et al. 2006). Thus, the accuracy of radiosonde observations has a direct impact on the evaluation and development of forward models, as well as in the evaluation of radiometers themselves. Remote sensor measurements of Precipitable Water Vapor (PWV) have also played an important role in the evaluation of radiosonde accuracy (Clough et al. 1999; Revercomb et al. 2003; Westwater et al. 2003). Both radiosonde measurements and remotely sensed PWV have also significant applications in climate research (Revercomb et al. 2003), and in the calibration and validation of remote sensing instruments (Westwater 1997; Turner and Goldsmith 1999; Turner et al. 2003). For these studies, inter-comparisons between radiosonde types and from different manufacturers, as well as between various types of remote sensors, are quite useful in evaluating accuracies, and in discovering possible inconsistencies in the measurements.

Many comparisons of radiosondes and remote sensors have been conducted in the mid-latitudes (Revercomb et al. 2003; Wang et al. 2003), and in the tropics (Westwater et al. 2003; Wang et al. 2002). However, there is a dearth of radiosonde and remote sensor comparisons for Arctic locations. This is especially important in climate modeling for PWV < 3 mm when infrared radiance in normally opaque regions become partially

transparent and structure in the frequency spectrum becomes apparent. As a first step in evaluating a variety of radiometers in the Arctic in conditions of low PWV, in March 1999, an Intensive Operating Period (IOP) was conducted at the U.S. Department of Energy's Atmospheric Radiation Measurement Program's (ARM) "Great White" field site near Barrow, Alaska (Racette et al. 2005). Due to a limited number of radiosondes, many questions were left unanswered about the accuracy of radiometric remote sensors. In particular, at that time, ARM radiosondes were launched only once-per-day and at asynoptic times, making comparisons with the synoptic launches of the U.S. National Oceanic and Atmospheric Administration (NOAA)/ National Weather Service (NWS) difficult. In addition, Vaisala RS80 radiosondes were also launched by ARM, and these radiosondes are known to have a dry bias, at least at middle and tropical latitudes (Revercomb et al. 2003; Turner et al. 2003; Westwater et al. 2003; Wang et al. 2002). To understand better these and other issues, the 2004 North Slope of Alaska Arctic Winter Radiometric Experiment - Water Vapor Intensive Operational Period (WVIOP04), was conducted at the Great White from March 9 to April 9, 2004. The basic goals of the experiment were to examine the relative sensitivity of millimeter wavelength radiometers to conventional microwave radiometers, to demonstrate a new NOAA instrument and associated calibration techniques, and to compare microwave and millimeter forward models for radiative transfer. A description of the experiment is given by Westwater et al. (2004), and initial results are given in (Mattioli et al. 2005b; Cimini et al. 2005; Westwater et al. 2005). In this experiment, several radiosonde observations (RAOB) by different types of sensors were taken and several remote sensing instruments were operated. This paper presents the results of the comparisons of radiosonde measurements

of temperature and relative humidity profiles, as well as the comparison of measurements of PWV by radiosondes, a dual-channel microwave radiometer (MWR), a microwave radiometer profiler (MWRP), and a Global Positioning System (GPS) receiver. Our results represent the first systematic comparisons of the above remote sensors and radiosonde systems for cold (0 to – 40 ºC in surface temperature) and dry (PWV from 0.08 to 1.5 cm) conditions.

2. Radiosonde launch strategy

In the 2004 IOP, three different humidity sensors were deployed from three separate locations near Barrow. ARM Operational Balloon Borne Sounding System (BBSS) radiosondes were launched daily at 2300 UTC (2 pm AST) at the Great White (GW). In addition, at the ARM Duplex (DPLX) in Barrow, 2.2 km to the west of the GW, BBSS radiosondes were launched four-times daily (0500, 1100, 1700, and 2300 UTC). Data from synoptic radiosondes from the NWS (1100 and 2300 UTC) were also archived. The NWS site is in Barrow, 4.3 km to the southwest of the GW. Finally, during clear conditions, eight dual-radiosonde launches (See Section 3c) were conducted at the ARM Duplex. The location and coordinates of the three RAOB sites are shown on the map in Figure 1. This collection of almost simultaneous and nearly co-located RAOBs allowed us to compare various aspects of temperature and humidity measurements.

3. Radiosonde types

a. VAISALA RS90

From the beginning of the experiment, radiosondes of the Vaisala RS90-A type were launched at the ARM DPLX in Barrow and at the ARM GW site. For convenience, these

radiosondes will be referred to as DPLX-RS90 and GW-RS90, respectively. The RS90-A is a "PTU-only" system, i.e., the primary measurements are pressure (P), temperature (T), and relative humidity (RH). Altitude and dew point temperature are derived quantities in the data. The sensor for the temperature measurement is the Vaisala F-Thermocap, which consists of a capacitive wire. The sensor for the relative humidity is the Vaisala Heated H-Humicap, a thin film capacitor with a heated twin-sensor design: two humidity sensors work in phase so that while one sensor is measuring, the other is heated to prevent ice formation [www.vaisala.com]. Samples were taken every 2 seconds. Details of the sensors accuracies are given in Paukkunen et al. (2001).

b. Sippican VIZ-B2

During the experiment, the synoptic radiosondes of the Sippican VIZ-B2 type were launched in Barrow by the NWS. These radiosondes were implemented at NWS upper-air sites in 1997 and are currently operated at 28 NWS upper-air sites. At present, a new Radiosonde Replacement System (RSS) is under deployment (www.ua.nws.noaa.gov/RRS.htm) and expected to continue until 2009, to replace the radiosondes in use with GPS-based radiosondes. The initial GPS radiosonde is the Sippican MkIIA.

The VIZ-B2 radiosonde measure P, T, and RH every 1.2 seconds and reported every 6 seconds for archiving. Altitude and dew point temperature are derived quantities in the data. Wind speed and direction are also computed, by using the recorded elevation and azimuth information from the radiosonde tracking system. Here, these soundings will be referred to as NWS-VIZ. The sensor for the temperature measurement is a 4.5-cm long white-coated rod (long rod) thermistor, and the sensor for humidity measurements is a

carbon hygristor (CH), whose characteristics are described in Blackmore and Taubvurtzel (1999). The VIZ-B2 sondes have manufactor-installed calibration factors for the humidity measments. For the NWS sondes, the H1 and H2 factors have been used at all NWS stations that use VIZ-B2 sondes since 1999.

c. Sippican Mark II with Meteolabor Snow White

During the experiment, eight successful dual-radiosonde launches were conducted at the ARM Duplex, two during the day and six during the night. Two radiosonde packages flew on the same balloon. The first package was the ARM DPLX-RS90, as discussed in Section 3a. The second was a radiosonde of the Sippican GPS Mark II type, operated by the National Aeronautics and Space Administration (NASA), which contained a VIZ 2.5 cm long white-coated rod (short rod) thermistor for temperature measurements and a VIZ carbon hygristor humidity sensor. In contrast to the NWS VIZ sondes, the H1 and H2 factors for the carbon hygristor were set to unity. The Mark II radiosonde had also attached a "Snow White" chilled mirror dew-point hygrometer, manufactured by Meteolabor AG, Switzerland [http://www.meteolabor.ch]. This instrument is a dew-point sensor designed for radiosonde application, whose performance has been evaluated in many studies (Fujiwara et al. 2003; Vöemel et al. 2003; Wang et al. 2003, Miloshevich et al. 2006). The accuracy of the mirror temperature measurement is better than 0.1°C. With accuracy in the air temperature measurement of 0.2°C, the RH uncertainty is about 2% of the %RH value (Vömel et al. 2003). The instrument has shown good performance in the tropical troposphere (Fujiwara et al. 2003) and in detecting cirrus clouds (Wang et al. 2003), although some deficiencies of the instrument were found in the presence of extremely dry layers (Vömel et al. 2003), in precipitation, heavy-water ladened clouds,

and sometimes in ice clouds (Schmidlin and Northam 2005). For convenience, we will refer to the Mark II humidity sensor and to the Snow White as MK2-CH and MK2-SW, respectively.

A GPS receiver and antenna were included in the Mark II system and integrated into the radiosonde electronics [www.sippican.com]. Values for wind and pressure are derived from the GPS data processing by applying the hydrostatic equation to the GPS altitude, so that a pressure sensor was not included. Three spare channels were also included in the system, and used to transmit the data from the attached MK2-SW. Samples were taken every 1.2 seconds. Each dual-radiosonde launch had therefore simultaneous and colocated humidity profiles from three humidity sensors (DPLX-RS90, MK2-CH, and MK2-SW) for inter-comparison.

4. Radiosonde data processing

Our first step in processing the data was simply to plot the soundings and to inspect them visually. In examining the RAOBs, we identified spurious values in the relative humidity measurements in the DPLX-RS90 soundings. An example of profile that was not accepted immediately is shown in Figure 2. This led to a series of quality control procedures, where soundings were filtered for the presence of the spikes in the humidity measurements (see Figure 2). Finally, for the statistical comparisons of the temperature and humidity profiles in Section 5, soundings were fitted to a regular grid over pressure, with a resolution of 5hPa. After inter-comparisons between the radiosoundings, and with remote sensing instruments (GPS, MWR, MWRP), outliers (entire profiles) were also identified and removed from the statistics (see Sections 6 and 7). Table 1 shows the number of launches and the available soundings after quality control.

We found that spurious values were present in the RH data in approximately 50% of the DPLX-RS90 soundings. The spurious data were observed generally above 10 km from the surface. However, after the start of the IOP, another meteorological experiment involving Aerosondes (Holland et al. 2001) was conducted at the Aerosonde Arctic Site in Barrow. We believe that the origin of the spikes is the result of radio frequency interference that affected the telemetry. In addition, the spikes were not present in the GW data. At this location, a spectrum analyzer was operated to search for an interference-free frequency before the radiosondes were launched. The fact that the spikes did not occur in the temperature (see Figure 2 (left)) may also be a result of the differing sensitivities of the temperature and RH calibration functions. Unfortunately, we could not obtain the raw sounding data to confirm our hypothesis. In general, above the tropopause, it was possible to recognize the behavior of the true profile behind the spikes, by assumptions on the vertical continuity of the measurements, for pressure values lower than 250-200 hPa. Following these considerations, we applied a non-linear filter to detect and replace these high-frequency interferences in the data. Noisy points were identified by using a median high-pass filter, and then corrected by interpolation over the adjacent points. The window of the filter varied with the noise magnitude and hence was varied from radiosonde to radiosonde. Figure 2 (right) shows the noisy RH measurements and the reconstructed profile. The filter works properly when the interferences are infrequent, and it is possible to estimate the missing value from contiguous points not affected by the noise. In few cases, some uncorrected spurious noise that still affected the RH measurement was eliminated by hand editing after intercomparisons.

5. Statistical comparison of RAOB temperature and relative humidity measurements

In this section, individual soundings, as well as statistical evaluations of the temperature and relative humidity measured by the different sensors are presented. Figure 3 shows the sounding profiles taken on March 26 at 2300 UTC, from the various radiosonde types. Among them, the Mark-II and the DPLX-RS90 were attached at the same balloon. First, we note excellent agreement between the four measurements of temperature up to 150 hPa (see Section 5b for discussion). In the humidity profiles, we note that the two RS90 measurements, as well as those of the MK2-SW are in good agreement. However, the two VIZ-like soundings differ considerably with the other types of sensors in vertical resolution below about 5 km, and in absolute magnitude above this altitude. At higher altitudes, the differences approach 30%. However, the two VIZ-like soundings agree roughly within 5% of each other. Thus, as mentioned above, dual launches provide unique opportunities to compare different radiosonde types, since the difference in the measurement is only due to the sensor type and not to temporal or spatial displacements. Although in this work we have chosen pressure as our vertical coordinate, plots of RH and T differences as a function of geopotential height were quite similar (Mattioli et al. 2005b).

a. GW-RS90 vs. DPLX-RS90 profiles

Figure 4 shows the profiles of temperature and relative humidity difference between the DPLX-RS90 and the GW-RS90 soundings. The comparisons are restricted during the daytime at 2300 UTC (2 p.m. AST), at the time of the GW-RS90 launches. The average

difference profile (black line) and standard deviation (std) profile (dark gray line) are also given. Usually, only small differences are present in the difference profiles, which can be attributed to the spatial baseline (2.4 km), temporal baselines (less than 15 min) and manufacturing tolerances for the sensor between the launches at the DPLX and at the GW. However, since many of the profiles have significant vertical structure in RH (an example of which can be seen in Figure 3), it is not surprising that differences of more than 20% were occasionally observed at the same pressure level. In our analysis, since the two stations have the same altitude above the sea level, we also evaluated the temperature differences near the surface. We noted a temperature bias near the surface (DPLX warmer than GW) of about 0.8 ºC, which persists for about 100 m. This is also not unexpected because of the presence of local sources of heating in the city of Barrow. Above 100 m, the temperature difference stays within 0.3 ºC, with a average std value of 0.32 ºC. For the RH, the mean difference stays within 2%, with average std value of 2%.

b. NWS-VIZ vs. DPLX-RS90 profiles

Figure 5 shows the temperature difference profiles between the NWS-VIZ soundings and the DPLX-RS90. The comparison is performed with the dataset divided into data taken at night at 1100 UTC (2 a.m. AST) and during the day at 2300 UTC (2 p.m. AST). Two features can be noticed. First, there is a gradient in temperature around 1000 hPa (corresponding to about 100–300 m above the surface), with the temperature over the NWS station higher than over the DPLX. The gradient in temperature is also present in the temperature comparison of NWS-VIZ and the GW-RS90 radiosondes (not shown). This phenomenon could again be explained by the presence of local heating in the town

of Barrow. As was shown in Figure 1, the NWS station is in the town, the DPLX is located in the periphery of the town, and GW is the farthest site from the town.

Second, our partition indicates the presence of a negative bias up to –3.5°C at 30 hPa for pressures lower than 250 hPa, between the NWS and the DPLX temperature, during the night, and from 915 to 50 hPa, almost no bias during the day. This latter behavior was also found in the comparison between the NWS-VIZ and the GW-RS90 (not shown). Also not shown, the nighttime bias reached –5°C at 20 hPa. A possible cause is in the magnitude of the radiation correction between the two types, since NWS does not apply such correction to the VIZ-B2 sondes. All the Vaisala radiosonde temperature data were corrected automatically for radiation errors by using the most recently supplied manufacturer's correction tables. Because the NWS-VIZ thermistor is a long white-coated rod, it has a very large IR error due to emission (emissivity 0.9), while its short-wave absorptivity is much less (~0.14). Both Vaisala sensor absorptivity and emissivity are quite small (<0.1). Therefore, the rod has a large IR error that is especially noticeable at night when compared with Vaisala type sensors (Schmidlin et al. 1986). The error of the rod is keyed to the background radiative environment and can be different depending on location and conditions.

Figure 6 shows statistical comparisons of the RH measurements. Since no specific difference was found in the day and night partitions, the comparison is shown for the entire dataset. At 1000 hPa, the RH profiles of NWS-VIZ radiosondes are 2% lower on average with respect to the DPLX-RS90. However, at P lower than 925 hPa, the RH from the NWS-VIZ carbon hygristor is on the average larger than the one from the DPLX-RS90 H-Humicap. Above about 250 hPa, this bias increases to values as large as

23%, with an average value of 17%. This is in contrast to the situation with P greater than 250 hPa, where the average RH difference is 4%. The corresponding std is 12% and 9%, respectively. The reason for this bias is due to the slow response time of the carbon hygristor at low temperatures (Blackmore and Taubvurtzel 1999), with no response below –60°C. Wang et al. (2003) found that the CH sensor stopped responding from –8°C to –55°C with a mean of –28°C. At the suggestion of an anonymous reviewer, we also examined the dependence of the VIZ-RS90 differences as a function of T. For the three ranges: A-T > –35°C; B- –35 < T < –50°C; C- T < –50°C, the biases were 5.1%, 10.3%, and 16.6%, respectively. However, by far the most important factor in determining the differences was the RH itself, with a bias of 18.5% for RH> 25%, while for RH> 25%, the bias was less than 2%.

c. Comparison of the NWS-VIZ and the DPLX-RS90 with the Mark-II

Figure 7 shows the difference between the NWS-VIZ radiosondes and the Mark-II. Because of the weight of the various packages on the same balloon, balloons for dual-radiosonde launches did not reach pressures lower than 150 hPa. Figure 7(left) gives the difference profiles taken by the NWS-VIZ and the MK2 VIZ temperature sensors. The analysis is shown for the data taken at night at 1100 UTC. As can be noted from the figure, there is a negative bias as large as 1°C, which resembles the negative bias shown in Figure 5(left) up to 150 hPa. The reason of such a difference in the temperature measurements of the two VIZ sensors is the use, by the Mark II radiosonde, of a short rod thermistor instead of the long rod. The short rod thermistor has a smaller IR error with respect to the long one (Schmidlin et al. 1986). Nevertheless, since the low-pressure values were never reached, the comparison is partially limited. Figure 7(right) gives the

comparison between two humidity sensors of the same type, the NWS-VIZ CH and the MK2-CH, for the entire dataset. The comparison shows generally good agreement between the measurements of the two carbon hygristors, with almost no bias and an average of the std profile less than 6%.

In Figure 8, the analysis is performed between the humidity sensors in the dual-radiosonde launches. Figure 8(left) shows the comparison between the two sensors, MK2-CH and MK2-SW, mounted on the package launched by NASA. The CH measurements, both MK2-CH and NWS-VIZ (not shown) have a significant bias with respect to the MK2-SW for pressures lower than 400-300 hPa (generally, above the tropopause). In particular, the average of the RH bias profile between MK2-CH and the MK2-SW is 7% for pressures greater than 400 hPa, while is 26% for pressures lower than 300 hPa. Similarly, the average of the RH bias profile between the NWS-VIZ and the MK2-SW was 6% for pressures greater than 400 hPa, and 20% for pressures lower than 300 hPa, respectively. This was also consistent with the results obtained in the comparison of the MK2-CH and the DPLX-RS90, also launched on the same balloon (not shown), where the average of the RH bias profile was 5% and 25%, respectively.

Finally, the dual-launch comparison between RH profiles obtained by the MK2-SW and the DPLX-RS90 H-Humicap shows a much better agreement, with an average value of the bias profile of 1%, and an average value of the std profile of 4%, as shown in Figure 8(right). Thus, both the SW and the RS90 humidity soundings are consistent with each other and both differ substantially from the VIZ.

6. PWV comparison among radiosonde types

In addition to the constant-pressure level comparisons, we compared the radiosonde types in terms of PWV. Figure 9 shows the PWV time series computed from the various radiosondes. Some features can be noticed: (a) even though only 30 days of observations were taken during the March-April 2004 time period, there was a 0.1 to 1.4 cm range in PWV; (b) the NWS-VIZ radiosondes measured in general higher values of PWV with respect to the Vaisala RS90 (launched at both DPLX and GW); (c) the PWV from the MK2-SW agrees very well with data from the Vaisala radiosondes, while the PWV computed from the MK2-CH is usually larger than the ones computed from both the Vaisala and the MK2-SW; (d) PWV computed from the MK2-CH data are closer to the values obtained from the NWS-VIZ radiosondes than are the DPLX-RS90 or MK2-SW measurements. Before presenting the statistical analysis of the data shown in Figure 9, it is necessary discuss outliers and how they were identified and removed from the data. Figure 10 shows the PWV from NWS-VIZ compared with PWV from the other radiosonde types. As can be noticed from Figure 10(a), there were four outliers in the DPLX-RS90 vs. NWS-VIZ comparison, shown in the scatterplots enclosed in dotted circle. Two PWV outliers with smaller value from the NWS-VIZ with respect the DPLX-RS90 (around 1 cm) are derived from NWS soundings taken on March 21, at 2300 UTC in which the RH from the carbon hygristor reached values as low as 1%, and on March 22, at 1100 UTC, in which the RH profile was consistently lower than the DPLX-RS90. The third PWV outlier (0.78 cm from NWS-VIZ vs. 0.57 cm from DPLX-RS90) is derived from NWS sounding taken on April 9, at 2300 UTC where the bias in the NWS-VIZ humidity measurements above the tropopause reached values as high as

40%. The fourth outlier, (around 0.1 cm), corresponds to the NWS sounding taken on March 18 at 11 UTC. On that occasion, the RH profile from the NWS-VIZ resembled the RH profile from the MK2-CH, but was consistently lower by 28-30% than all the other soundings.

The outliers affected considerably the statistical comparisons between the data sets of Figure 10 (a-d). Statistical analyses of the cleaned data quantify the features observed in Figure 9, and are given in Table 2. Here, the values of bias, std, correlation coefficient (corr), slope and intercept (int) of a linear fit are reported. The number of available dual-radiosonde launches is not sufficient for a significant statistical comparison, but the analysis can give a useful indication of the general behavior. On the average, the PWV from the NWS-VIZ radiosondes is about 0.05 cm higher than the PWV from the DPLX-RS90 radiosondes. Good agreement is found between the PWV from the MK2-SW and the DPLX-RS90 for the dual-radiosonde launches, with a root mean square (rms) difference better than 0.01 cm. The PWV from MK2-CH shows a consistent bias (on the order of 0.03 cm) with respect to the PWV from the other sensor types, and except for the apparently erroneous sounding of March 18, at 1100 UTC, it is closer to PWV from NWS-VIZ (see Figure 10(d)). When we compare soundings of similar instrument type, we note that the rms difference between the PWV values from the GW-RS90 and the DPLX-RS90 is 0.011 cm, and the rms difference between NWS-VIZ and MK2-CH sensor is 0.013 cm. These small differences are due to spatial and temporal baseline, and sensor accuracy. The rms difference between PWV data from NWS-VIZ and MK2-SW sensor is 0.039 cm. These observations about the integrated quantity PWV are in agreement with the differences observed in the humidity profiles that were presented in

Section 5. No significant difference is found between the MK2-CH and the NWS-VIZ in the RH measurements and in PWV.

Diurnal differences were also observed. As can be seen from Table 2, the bias in the PWV comparison of the NWS-VIZ with the GW-RS90 is 0.061 cm, and with the DPLX-RS90 is 0.046 cm. However, the GW-RS90 is only taken during the day. When the comparison between NWS-VIZ and DPLX-RS90 involved data taken at 2300 UTC, the bias (NWS-DPLX) is 0.058 cm, but for data taken at night at 1100 UTC, the bias is 0.031 cm.

7. PWV comparisons between MWR, MWRP, GPS and the radiosondes

In this section, we summarize the comparisons of PWV obtained from the MWR and the MWRP operating at the GW field site, with the PWV from a GPS receiver located near the GW, and from the radiosondes. Although these remote sensors represent a mature technology, there is a dearth of information about their accuracy during cold and arctic conditions.

a. MWR

The MWR that operates at the GW is a dual-frequency water vapor radiometer of the WVR-1100 series from Radiometrics Corporation [www.radiometrics.com], operating at 23.8 and 31.4 GHz. The MWR scans at five elevation angles (19.35, 23.4, 30.15, 41.85 and 90.0 degrees) in the east–west direction at about 1-minute intervals. During clear conditions, brightness temperature (Tb) measurements at these angles are used to calibrate the radiometer by the tipping curve method (Liljegren 2000; Han and Westwater 2000). Based on the comparison of three similar radiometers that operated at the ARM

Southern Great Plains (SGP) Central Facility, the accuracy of these radiometers is about 0.4 K rms (Mattioli et al. 2005a). However, during the IOP, we were able to obtain a large number of good tipping curves. Furthermore, forward model Tb calculations given in Racette et al. (2005) showed comparisons with radiosondes at the 0.2 K rms level. Thus we believe, that our accuracy in Tb determinations by the MWR is at least as good as the 0.4 K rms value determined at the SGP. Here, PWV is retrieved from the Tb measurements using a physically based statistical inversion algorithm whose forward-model component is that of Liljegren et al. (2005). Considering that the sensitivity of 23.8 GHz channel to PWV is about 1.3 K/mm (see Racette et al. 2005), an absolute accuracy of 0.4 K in Tb corresponds to an absolute accuracy of 0.3 mm in PWV. The manufacturer's estimates of "repeatability" of 0.25 K (on a 30 s sample) leads to a similar number for PWV of 0.2 mm on a 30-s basis. For the much more humid environment during a Spring IOP at the SGP (1.0 to 2.0 cm of PWV), the rms accuracy in measuring PWV, determined by comparison with radiosondes, GPS, and between different radiometers, was found to be about 0.75 mm (Mattioli et al. 2005a).

b. MWRP

The MWRP is a twelve-channel radiometer of the TP/WVP-3000 series from Radiometrics Corporation, with five K-band channels (22.235, 23.035, 23.835, 26.235, and 30.0 GHz) and seven V-band channels (51.25, 52.28, 53.85, 54.94, 56.66, 57.29, and 58.8 GHz), which provides humidity and temperature profiles at about 5-minute intervals. The K-band channels and the lowest V-band channel are used from humidity sensing, while the remaining V-band channels are used for temperature profiling. The system

includes also an infrared broad-band radiometer and PTU sensors for the measurements of cloud base temperature and surface pressure, temperature, and humidity, respectively.

PWV can be retrieved from MWRP observations by using different combinations of channels (Liljegren and Lesht 2004). For example, in this study we compare results obtained using only 2 channels, 23.835 and 30 GHz (2ch), the five K-band channels (5ch), and 6 channels, which are the five K-band plus the 51.25 GHz channel (6ch). Both the MWRP and the MWR are provided with a wet window sensor mounted on the top of the radiometer that turns a heater on during condensing or precipitating conditions to promote the evaporation of rain or snow. In our comparisons, PWV data from the radiometers are accepted when the wet window sensor indicated that the heater was not activated. However, as we show later in this section, in terms of PWV, the retrievals obtained when the heater was activated seemed consistent with the observations from other instruments (GPS, RAOBS). The manufacture estimate of the absolute accuracy of the V-band channels is 0.5 K, with a repeatability of 0.25 K on a 5-min basis.

c. GPS (Near real-time and Reprocessed)

A SuomiNet (Ware at al. 2000) GPS receiver (SG27) located near the GW provided PWV measurements throughout the experiment. The site is also incorporated into the NOAA Forecast Systems Laboratory (FSL) ground-based GPS Meteorology network (http://gpsmet.noaa.gov). FSL and the Scripps Institution of Oceanography collaborated to produce the near real-time (NRT) and reprocessed data used in this experiment, by using the method described in Gutman et al. (2004a). PWV is retrieved from the tropospheric GPS signal delay by first parsing it into its wet and dry components by subtracting a hydrostatic delay that is calculated from surface pressure (Saastamoinen

1972). Then, the wet component is mapped into PWV using a transfer function that is nearly proportional to surface temperature (Bevis et al. 1994). The temporal resolution of the measurements is 30 min.

When using NRT data, because of the time constraints imposed on data availability from operational weather forecasting, no special effort is made to reduce the impact of various sources of processing errors, which appear as noise or scatter about the mean in the NRT solution. As described in Gutman et al. (2004b), errors occur in the estimation of the zenith tropospheric signal delay (Bevis et al., 1992), in the estimation of the zenith hydrostatic delay (Saastamoinen 1972, Elgered 1993), and in the estimation of the wet delay transfer function (Bevis et al. 1994). These errors are independent, and correspond to an equivalent PWV retrieval error of 0.75 mm, 0.25 mm, and 0.1 mm respectively, for a total estimated PWV retrieval error of 0.8 mm rms. However, initial comparisons of the GPS NRT data with those of the MWR revealed a substantial amount of scatter in the GPS data.

For the meteorological conditions of this experiment, we analyzed the errors arising from estimating the weighted mean temperature of the atmosphere Tm (Davis et al. 1985) in the wet delay transfer function to PWV, either from radiosonde or surface data. Here, Tm was calculated from 30-min interpolations of the four-times-per-day DPLX-RS90 radiosondes. Although, the once-a-day radiosondes launched at GW were not sufficient to represent the diurnal differences, the rms difference between Tm computed at the DPLX and GW for the daytime dataset is 0.24 K, which is negligible in terms of PWV. For the month of the experiment, we found that the rms difference between Tm computed from RAOBs and from the surface temperature Ts was less than 5 K, with a relative error

in the transfer function of about 2%, and an rms difference in terms of PWV of 0.012 cm. These errors are small relative to the scatter. We believe the scatter is mainly associated with site-dependent multipath, which affects the zenith total delay estimation, which was associated with low satellite elevation angles. A contributing factor was high reflectivity properties associated with frozen ice and snow of the surface. Another possible source of error is in the Niell mapping function (Niell 1996), which relates the zenith delay to the signal slant path. (The use of an elevation cut-off angle of 7.5 deg increased the accuracy of its use for an arctic environment.)

An attempt to minimize the scatter in the NRT was made by: (1) implementing the post fit residual option in the GAMIT (GPS at Massachusetts Institute of Technology) data software package (King and Boch 1996) used by FSL (2) using precise orbits from the GNSS (Global Navigation Satellite Systems) Service (rather than real-time hourly orbits provided to FSL by Scripps), and (3) identifying and removing repetitive aspects of the tropospheric delay time series associated with multipath. The reprocessing steps reduced most, but not all of the scatter in the GPS data, as can be noticed from Figure 11, which shows the time series of PWV-NRT and PWV reprocessed for the duration of the experiment. For the two data sets, the bias (PWV-reprocessed - PWV-NRT) is 0.012 cm and the std is 0.064 cm. Considering that the average PWV value during the experiment was 0.4 cm, the percentage difference is substantial, and is over 15%. In the following comparisons, only reprocessed GPS data will be used.

d. PWV from MWRP

The PWV from MWRP is retrieved by using measurements at 2 channels, 23.835 and 30 GHz (2ch), at 5 channels, 22.235–30 GHz (5ch), and at 6 channels, 22.235–30 and the

51.25 GHz (6ch). Monthly retrieval coefficients, including the mean radiating temperature Tmr, have been specifically computed from past radiosoundings using Vaisala radiosondes (RS80-H and RS90) launched at the Great White between 1998 and 2004, assuming a 0.3 K rms error in brightness temperature. All are based on the modified Rosenkranz model (Hitran width at 22 GHz, CKD 2.4 continuum; see Liljegren et al. 2005).

Table 3 gives the statistical comparison of the PWV from the MWRP (from the three set of channels) with the PWV retrieved from the MWR. The 2-channel retrieval showed the larger scatter in the PWV, which is apparently due to the less frequent sampling of each channel of the MWRP (5 min) with respect to the MWR (30 s). The use of 5 and 6 channels reduced this scatter considerably, with the same rms difference of 0.033 cm.

e. Intercomparison of remote and radiosonde PWV measurements

Figure 12 shows the PWV time series from the MWRP, retrieved by using six channels, the MWR, the radiosondes launched at the DPLX, and the reprocessed GPS data. In general, there is good qualitative agreement between the measurements except at times when the GPS differs from the other sensors by about 1.5 mm. However, one of the strengths of GPS (i. e. the availability of data during all weather conditions) is noted between days 80 and 82, when the wet rain flags eliminated radiometer data. The simultaneous presence of the MWR, the MWRP, and the GPS allowed us to further investigate the peculiar NWS-VIZ soundings identified in the radiosonde analysis (see also Figure 10), that occurred on March 21 at 2300 UTC, on March 22 at 1100 UTC and on April 9 at 2300 UTC. In the comparison with the MWR, the MWRP and the GPS, we also found an outlier in the DPLX-RS90 radiosondes that occurred during March 22 at

1700 UTC. These cases are shown in Figure 13. Even though in some occasions the wet window flags were on, the data are of sufficient quality that the three RAOB outliers are easily recognized, and hence have been eliminated in the following statistical comparisons. It was pointed out by an anonymous reviewer that there is an apparent low bias in Figure 13a of the GPS relative to the MWR. Further examination of the data showed that for the period from day 80.3 to 83.4, the surface temperature Ts was warmer than –20 ºC, and that Tmr's calculated from radiosondes differed by about 13 ºC, for both channels, from the monthly averages that were used in the retrievals. As a test, we recomputed the MWR and the MWRP retrievals of PWV using temporally interpolated Tmr and for the warmest periods (Ts > – 20 ºC), the radiometrically derived PWV was lowered by as much as 0.09 cm. For the time period shown in Figure 13a, the average MWR-GPS difference was 0.035 cm with a standard deviation of 0.04 cm by using interpolated Tmr, and 0.08 cm with a standard deviation of 0.03 cm by using monthly averaged values. Because we wanted to use only the ARM operational retrieval algorithm in our complete analysis, we did not use interpolated Tmr's for any subsequent statistical analysis of the data set. Data identified by the wet window flag were also not used.

The statistical analyses of the GPS comparisons with the MWR and the MWRP are given in Table 4. The comparisons were performed using 30-minute-averaged radiometer data centered on GPS time. In general, good agreement is found between the radiometers and the GPS. The bias is negligible in the comparison with the MWR, and is 0.026 cm in the comparison with the MWRP, and the std is on the order of 0.05 cm.

These rms differences (0.058 cm) are somewhat better than the estimated accuracy of 0.08 cm given in Section 7c.

Table 5 summarizes the results of the comparison of the MWR, the MWRP and the GPS with the DPLX-RS90, the GW-RS90 and the NWS-VIZ. In the comparison with the radiosondes, both the PWV from GPS and the MWR are greater than the PWV from the DPLX-RS90 by about 0.03-0.04 cm, and less that the NWS-VIZ by about 0.015 cm. The MWRP is greater than the DPLX-RS90 by 0.013 cm and less than the NWS-VIZ by 0.036 cm. These results are consistent with the radiosonde comparisons shown previously in Section 4 and 5, in which the VIZ carbon hygristor (both the NWS-VIZ and the MK2-CH) show an apparent bias (16-20% higher) in the RH measurements above the troposphere with respect to the other sensors. The bias difference (about 0.015 cm) of the MWR with respect to the MWRP in the comparison with the radiosonde is consistent with the results of Table 3. Because of the substantial bias between the remote sensors and the GW RS90 data, and, remembering that the GW soundings were only taken during the day, we examined various diurnal trends in the data. The results are given in Table 6. This table shows the existence of two diurnal biases. First, we observe that there is no diurnal difference between the MWR and the NWS-VIZ, the MWRP and the NWS-VIZ, and the MWR vs. the MWRP. However, there is a diurnal difference between the (MWRP, MWR) and the DPLX-RS90 of about 0.02 cm. Since the bias is consistent with that one found in the comparison with the GW-RS90, we believe that these differences are not due to the spatial baseline but to a day-time dry bias shown by the Vaisala RS90. Although in a different climatic region, a dry bias in daytime RS90 measurements was also found by Miloshevich et al. (2006) and attributed to solar heating of the RS90

humidity sensor. As a second diurnal bias, we noticed a 0.04 bias between the (MWRP, MWR) and the GPS, all of which are located at the Great White. We found that this bias is mainly associated with repetitive aspects of multipath that still affected GPS data after reprocessing. The use of the Tm-Ts relationship may also lead to diurnal differences in PWV from GPS of about 1-2%, as found by Wang et al. (2005). In our case, this error was evaluated of the order of 1%.

8. Summary and discussion

One of the major goals of the 2004 NSA Radiometric experiment was to evaluate the performance of radiosondes and operational remote sensing systems during cold (0 to –40ºC) and dry (PWV < 0.5 cm) conditions. This was motivated because of the necessity to develop accurate forward models of both infrared and millimeter wavelength radiometers, during these conditions. Previously, experience by ARM (Revercomb et al. 2003) at the Mid-latitude field station in Oklahoma, or in the Tropics (Westwater et al. 2003) have been able to use MWRs to check the quality of radiosondes, and, to scale radiosonde-humidity profiles to be consistent with PWV derived from the MWR. However, for the Arctic Winter conditions, with PWV frequently less than 0.3 cm and surface temperatures less than -35ºC, it has been questioned if MWRs have the required sensitivity to perform the necessary scaling and/or quality checks. To evaluate both radiosonde and remote sensor performance, we deployed five different radiosonde packages as well as measurements of PWV by the MWR, MWRP, and GPS. Comparisons between temperature and relative humidity profiles have been presented, as well as comparisons in PWV between data taken by the remote sensors and each of the various radiosonde types. Our conclusions can be summarized as follows:

(1) Relative humidity measurements from the VIZ carbon hygristor (both the NWS-VIZ and the MK2-CH) show an apparent bias with respect to the other instruments (VIZ being higher) above the troposphere, with an average bias of the order of 16-20%. We also examined these differences as a function of temperature and relative humidity. Although, the differences generally became larger at temperatures lower than -35ºC, the most important factor was RH; for RH < 25%, the bias was 18.5%. In addition, no significant difference was found between the MK2-CH and the NWS-VIZ, either in the RH profiles or in PWV. Given that surface temperatures are frequently less that – 35ºC, this lack of reliability is a serious problem for the climate record at Barrow. Similar conclusions have been made by Wang et al. (2003) and by Ferrare et al. (2004) for upper tropospheric and lower stratospheric measurements at mid-latitudes.

(2) Based on eight dual-radiosonde launches, the Vaisala RS90 H-Humicap and the Snow White chilled mirror showed good agreement in the RH measurements, with an average of the bias profile of 1% and an average rms difference of the entire profile of the order of 5%. On a much larger data sample of RS90 launches at the ARM Duplex, there were some large positive spikes (20 to 30%) in the RS90 measurements. These were associated with radio frequency interference from another meteorological experiment that was being conducted at the same time, and were corrected by a filtering procedure. These spikes were not observed for launches at the GW.

(3) There was a bias in night-time temperature soundings for pressures lower than 300 hPa between the NWS-VIZ and the DPLX-RS90 soundings, which is due to

a large IR radiation error for the VIZ-B2 sondes not routinely corrected by NWS. This bias reached about 3.5°C at 30 hPa and 5ºC at 20 hPa.

(4) GPS measurements taken at Barrow provided the general pattern of PWV, including cloudy or snowy conditions, and the continuous availability (every 30 minutes) is very useful. However, Near Real Time data were affected by a substantial scatter compared to the MWRs. Reprocessing of the GPS data reduced most but not all this scatter. For very dry conditions, when the PWV was less than 2mm, the scatter exceeded 1 mm. Most of the scatter was associated with multi-path propagation that was associated with low satellite elevation angles. Contributing to this was high reflectivity properties associated with frozen ice and snow. Other meteorological errors were shown to be small relative to the multi-path effects.

(5) In using the MWRP, the use of all five channels in the water vapor band provided very good agreement with the MWR, with 0.033 cm rms. PWV retrieved by using 6 channels (five K-band channel and the 51.25 GHz) was also analyzed and provided the same rms difference. The PWV retrieved from the MWRP by using 2 channels (22.235 and 30 GHz) provided PWV values with a slightly larger dispersion (0.042 cm rms), due to a less frequent sampling of each channel of the profiler (5 min) with respect to the MWR (30 s). All of these retrievals used the new absorption algorithms of Liljegren et al. (2005).

(6) The use of the MWRP and the MWR data to identify spurious radiosonde data was especially useful, and should be considered in automated quality control. In addition, biases in the radiosonde observations can be identified.

(7) Over sample sizes that ranged from about 1000 to 7000 data points, the average values of the GPS, MWR, and MWRP (5-channel) PWV retrievals were close. The bias of the GPS with the MWR and MWRP was 0.005 and 0.026 cm (better than 8.6% of the mean PWV of 0.3 cm), and the two regression slopes better than 0.98. When comparing PWV retrievals of the MWR with 2-channel, 5-channel, and 6-channel retrievals of the MWRP, the maximum bias was 0.018 cm (6% of the mean PWV). The rms differences of the MWR vs. the 5- and 6-channel retrievals were both 0.033 cm (11% of the mean PWV).

(8) We found diurnal differences in PWV when comparing RS90 radiosonde data with three independent measurements: MWR, MWRP, and NWS-VIZ. These diurnal differences were, respectively, 0.025, 0.016, and 0.027 cm. Such differences, although in a different climatic regime, were also found by Miloshevich et al. (2006) and were attributed to heating of the RS90 humidity elements. The once-a-day operational soundings at the GW were not sufficient to study further these biases. We would suggest that flying Vaisala radiosondes from the GW at synoptic times would be helpful in exploring these diurnal differences.

(9) In an anomalously warm period, a small bias was found in radiometric retrievals of PWV (as large as 0.09 cm). This was due to the use of a monthly averaged mean radiating temperature that was used in the retrievals. To overcome such biases, temperature profile retrievals from the MWRP could be used. For locations in which such information is not available, the use of temporally

interpolated values computed from radiosondes or from numerical forecast models is suggested.

Starting on April 1, 2006, ARM is now launching two extra soundings-a-day (0600 and 1800 UTC), but not at the synoptic times of the NWS. However, the increased soundings should still be valuable in studying diurnal effects. Other work in progress will be focused on developing forward models for the millimeter wavelength radiometric observations that were taken during then 2004 IOP and are currently being taken by ARM at the NSA.

Acknowledgements

The work presented in this paper was sponsored by the Environmental Sciences Division of the Department of Energy under the Atmospheric Radiation Measurement Program.

The authors wish to thank Dr. Janet Machol and Dr. P. Ola Persson of NOAA/ETL for their useful comments on the manuscript.

References

Bevis, M., S. Businger, T. Herring, C. Rocken, R. Anthes, R. Ware, 1992: GPS meteorology: remote sensing of the atmospheric water vapor using the global positioning system. *J. Geophys. Res.*, **97**(14), 75-94.

——, S. Businger, S. Chiswell, T. A. Herring and R. A Anthes, C. Rocken and R.H. Ware, 1994: GPS meteorology: mapping zenith wet delays onto precipitable water. *J. Appl. Meteor.*, **33**(3), 379-386.

Blackmore, W. H., and B. Taubvurtzel, 1999: Environmental chamber tests on NWS radiosonde humidity sensors. Preprints, *11th Symposium on Meteorological Observations and Instrumentation*, Dallas, TX, Amer. Meteor. Soc., 259–262.

Cimini, D., A. J. Gasiewski, M. Klein, E. R. Westwater, V. Leuski, and S. Dowlatshahi, 2005: Ground-based Scanning Radiometer Measurements during the Water Vapor IOP 2004: a valuable new data set for the study of the Arctic atmosphere. Proceedings. *15th Atmospheric Radiation Measurement (ARM) Science Team Meeting*, Daytona Beach, Florida. U.S. Department of Energy. [Available at http://www.arm.gov/publications/proceedings/conf15/extended_abs/cimini_d.pdf]

Clough, S. A., P. D. Brown, 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. Proceedings. *9th*

Atmospheric Radiation Measurement (ARM) Science Team Meeting, Washington, D.C. U.S. Department of Energy. [Available at: http://www.arm.gov/publications/proceedings/conf09/abstracts/clough-99.pdf].

Davis, J. L., L. T .A. Herring, I. I. Shapiro, A .E. Rogers, and G. Elgered, 1985: Geodesy by radio interferometry: Effects of atmospheric modelling errors on estimates of baseline length. *Radio Sci*., **20**, 1593–1607.

Elgered, G., 1993: Tropospheric radio-path delay from ground-based microwave radiometry. *Atmospheric Remote Sensing by Microwave Radiometry*. M. A. Janssen, Ed. New York: J. Wiley & Sons, Inc., 218-258.

Ferrare, R. A., E. V. Browell, S. Ismail, S. A. Kooi, L. H. Brasseur, V. G. Brackett, M. B. Clayton, J. D. W. Barrick, G. S. Diskin, J. E. M. Goldsmith, B. M. Lesht, J. R. Podolske, G. W. Sachse, F. J. Schmidlin, D. D. Turner, D. N.Whiteman, D. Tobin, L. M. Miloshevich, H. E. Revercomb, B. B. Demoz, and P. Di Girolamo, 2004: Characterization of upper-troposphere water vapor measurements during AFWEX using LASE, *J. Atmos. Oceanic Technol.,* **21**(12), 1790– 1808.

Gutman, S. I., S. R. Sahm, S. G. Benjamin, B. E. Schwartz, K. L. Holub, J. Q. Stewart, and T. L. Smith, 2004a: Rapid Retrieval and Assimilation of Ground Based GPS Precipitable Water Observations at the NOAA Forecast Systems Laboratory: Impact on Weather Forecasts. *J. Meteor. Soc. Japan*, **82** (1B), 351-360.

——, S. I., S. R. Sahm, S. G. Benjamin, T. L. Smith, 2004b: GPS water vapor observation errors. *8th Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface*, paper 8.3, Seattle, WA, Amer. Meteor. Soc.

Fujiwara, M., M. Shiotani, F. Hasebe, H. Vömel, S. J. Oltmans, P. W. Ruppert, T. Horinouchi, and T. Tsuda, 2003: Performance of the Meteolabor "Snow White" chilled-mirror hygrometer in the tropical troposphere: Comparisons with the Vaisala RS80 A/H-Humicap sensors. *J. Atmos. Ocean. Technol.*, **20**, 1534–1542.

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

Hewison, T., D. Cimini, L. Martin, C. Gaffard, and J. Nash, 2006: Validating clear air absorption models using ground-based microwave radiometers and vice-versa, Meteorologische Zeitschrift.

Holland, G. J., P. J. Webster, J. A. Curry, G. Tyrell, D. Gauntlett, G. Brett, J. Becker, R. Hoag, and W. Vaglienti, 2001: The Aerosonde Robotic Aircraft: A New Paradigm for Environmental Observations: *Bull. Amer. Meteorol. Soc.* **82(5)**, pp. 889–901.

King, R.W. and Y. Bock. 1996: Documentation of the GAMIT GPS analysis software, version 9.4. Massachusetts Institute of Technology and Scripps Institution of Oceanography, Cambridge.

Liljegren, J. C. 2000: Automatic self-calibration of ARM microwave radiometers. *Microwave Radiometry and Remote Sensing of the Earth's Surface and Atmosphere*. Utrecht: P. Pampaloni and S. Paloscia, Eds., VSP Press, 433–443.

——, J. C. and B.M. Lesht, 2004: Preliminary results with the twelve-channel Microwave Radiometer Profiler at the North Slope of Alaska Climate Research Facility. Proceedings. *14th Atmospheric Radiation Measurement (ARM) Science Team Meeting*. Albuquerque, New Mexico. U.S. Department of Energy. [Available at: http://www.arm.gov/publications/proceedings/conf14/extended_abs/liljegren-jc.pdf]

——, J. C., S. –A Boukabara, K. Cady-Pereira, and S. A. Clough, 2005: The Effect of the Half-Width of the 22-GHz Water Vapor Line on Retrievals of Temperature and Water Vapor Profiles with a Twelve-Channel Microwave Radiometer. *IEEE Trans. Geosci. Remote Sens.,* **43** (5), 1102-1108.

Mattioli, V., E. R. Westwater, S. I. Gutman, and V. R. Morris, 2005a: Forward Model Studies of Water Vapor using Scanning Microwave Radiometers, Global Positioning System, and Radiosondes during the Cloudiness Inter-Comparison Experiment. *IEEE Trans. Geosci. Remote Sens.*, **43** (5), 1012-1021.

——, V., E. R. Westwater, D. Cimini, J. S. Liljegren, B. M. Lesht, S. I. Gutman, and F. Schmidlin, 2005b: "Analysis of Radiosonde and Precipitable Water Vapor Data from the 2004 North Slope of Alaska Arctic Winter Radiometric Experiment", Proceedings. *15th Atmospheric Radiation Measurement (ARM) Science Team Meeting*. Daytona Beach, Florida, U.S. Department of Energy. [Available on-line at: http://www.arm.gov/publications/proceedings/conf15/extended_abs/mattioli_v.pdf].

Miloshevich, L. M., H. Vömel, D. N. Whiteman, B. M. Lesht, F. J. Schmidlin, and F. Russo, 2006: Absolute accuracy of water vapor measurements from six operational radiosonde types launched during AWEX-G and implications for AIRS validation, *J. Geophys. Res.*, 111, D09S10, doi:10.1029/2005JD006083.

Niell, A. E. (1996), Global mapping functions for the atmosphere delay at radio wavelengths, *J. Geophys. Res.*, 101 (B2), 3227–3246.

Paukkunen, A., V. Antikainen, and H. Jauhiainen, 2001: The accuracy and performance of the new Vaisala RS90 radiosonde in operational use. *11th Symposium on Meteorological Observations and Instrumentation.* Albuquerque, NM. Amer. Meteor. Soc.

Racette, E. P., E. R. Westwater, Y. Han, A. J. Gasiewski, M. Klein, D. Cimini, D. C. Jones, W. Manning, E. J. Kim, J. R. Wang, V. Leuski, and P. Kiedron, 2005:

Measurement of Low Amounts of Precipitable Water Vapor Using Ground-Based Millimeterwave Radiometry. *J. Atmos. Ocean. Tech*., **22** (4), 317 – 337.

Revercomb H. E., 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, 2003: The ARM programs's water vapor intensive observation periods: overview, initial accomplishments, and future challenges. *Bull. Amer. Meteorol. Soc.* **84(2)**, 217-236.

Saastamoinen, J., 1972: Introduction to practical computation of astronomical refraction. *Bull. Geod.,* 106, 383-397.

Schmidlin, F. J., J. K. Luers, and P. D. Huffman, 1986. Preliminary estimates of radiosonde thermistor errors. NASA Technical Paper 2637. Wallops Island, Virginia.

——, F. J., and E. T. Northam, 2005: Standards for evaluating radiosonde measurements. *9th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface.* San Diego, CA. Amer. Meteor. Soc.

Turner, D. D., and J. E. M. Goldsmith, 1999: Twentyfour hour Raman lidar water vapor measurements during the Atmospheric Radiation Measurement program's 1996

and 1997 water vapor intensive observation periods. *J. Atmos. Oceanic Technol.,* **16,** 1062–1076.

——, B. M. 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.

Vömel, H., M. Fujiwara, M. Shiotani, F. Hasebe, S. J. Oltmans, and J. E. Barnes, 2003: The behavior of the Snow White chilled-mirror hygrometer in extremely dry conditions. *J. Atmos. Oceanic Technol.*, **20**, 1560–1567.

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

——, D. J. Carlson, D. B. Parsons, T. F. Hock, D. Lauritsen, H. L. Cole, K. Beierle, and E. Chamberlain, 2003: Performance of operational radiosonde humidity sensors in direct comparison with a chilled mirror dew-point hygrometer and its climate implication. *Geophys. Res. Lett.*, 30 (16), 1860, 10.1029/2003GL016985.

——, L. Zhang, and A. Dai, 2005: Global estimates of water-vapor-weighted mean temperature of the atmosphere for GPS applications. J. *Geophys. Res.*, 110, D21101, 10.1029/2005JD006215.

Ware, R. H., D. W. Fulker, S. A. Stein, D. N. Anderson, S. K. Avery, R. D. Clark, K. Droegemeier, J. P. Kuettner, and J. B. Minster, 2000: SuomiNet: A real-time national GPS network for atmospheric research and education. Bull. Amer. Meteor. Soc., **81**, 677-694.

Westwater E. R., 1993: Ground-based Microwave Remote Sensing of Meteorological Variables. *Atmospheric Remote Sensing by Microwave Radiometry*, M.A. Janssen, Ed. New York: J. Wiley & Sons, Inc, 145–213.

——, E. R., 1997: Remote sensing of tropospheric temperature and moisture by integrated observing systems. *Bull. Amer. Meteor. Soc.,* **78**, 1991-2006.

——, E. R., B. Stankov, D. Cimini, Y. Han, J. A. Shaw, B. M. Lesht, and C. N. Long, 2003: Radiosonde Humidity Soundings and Microwave Radiometers during Nauru99. *J. Atmos. Oceanic. Technol.*, **20** (7), 953-971.

——, E.R., Klein, M., Gasiewski, A., Leuski, V., Shaw, J.A., Mattioli, V., Cimini, D., Liljegren, J.C., Lesht, B.M., Zak, B.D., Uttal, T., Hazen, D.A., Weber, B.L., and Dowlatshahi, S., 2004: The 2004 North Slope of Alaska Arctic Winter Radiometric

Experiment. Proceedings. *14th ARM Science Team Meeting*. Albuquerque, New Mexico. U.S. Department of Energy. [Available on-line at: http://www.arm.gov/publications/proceedings/conf14/extended_abs/westwater-er.pdf].

——, D. Cimini, V. Mattioli, M. Klein, V. Leuski, A. J. Gasiewski, S. Dowlatshahi, J. S. Liljegren, B. M. Lesht, and J. A. Shaw, 2005: Microwave and Millimeter Wave Forward Modeling Results from the 2004 North Slope of Alaska Arctic Winter Radiometric Experiment. Proceedings. *15th ARM Science Team Meeting*, Daytona Beach, Florida. U.S. Department of Energy. [Available on line at: http://www.arm.gov/publications/proceedings/conf15/extended_abs/westwater_er.pdf]

FIGURE AND TABLE CAPTIONS

Figure 1. Location and coordinates of ARM "Great White" site (GW), ARM Duplex (DPLX), and NWS upper-air station in Barrow, Alaska.

Figure 2. (left) Temperature profile in the DPLX-RS90 sounding launched on April 4, 2004 at 2300 UTC. (Right) Relative humidity affected by unrealistic noise (gray dots) and reconstructed profiles (black dots) in the same sounding. Here, large excursions occur at pressures below 70 hPa.

Figure 3. Temperature (left) and relative humidity (right) soundings on March 26 at 2300 UTC.

Figure 4. Temperature difference profiles (left) and relative humidity difference profiles (right) between the Vaisala RS90 radiosondes launched at the DPLX and at the GW.

Figure 5. Temperature difference profiles between the NWS-VIZ radiosondes and the DPLX-RS90. (Left) Dataset taken at night (at 1100 UTC, 2 a.m. AST); (right) dataset taken during the day (at 2300 UTC, 2 p.m. AST).

Figure 6. RH difference profiles between the NWS-VIZ radiosondes and the DPLX-RS90, the comparison performed for the entire dataset.

Figure 7. Profiles of temperature and relative humidity difference between the NWS-VIZ and the Mark-II. (Left) NWS-VIZ long rod thermistor vs. MK2 VIZ short rod thermistor, for the dataset taken at night. (Right) NWS-VIZ carbon hygristor vs. MK2-CH, for the entire dataset.

Figure 8. Profiles of RH difference between the VIZ carbon hygristor, the Vaisala capacitor, and the Snow White chilled mirror, for the dual-radiosonde launches. (Left) MK2-CH vs. MK2-SW. (Right) DPLX-RS90 vs. MK2-SW.

Figure 9. PWV time series computed from the radiosondes that were operating during the 2004 NSA experiment.

Figure 10. PWV from NWS-VIZ compared with PWV from the other radiosonde types. Outliers are removed from the statistics and shown in the scatterplot enclosed in dotted circles. (a) PWV from NWS-VIZ compared to PWV from DPLX-RS90; (b) to GW-RS90; (c) to MK2-SW; (d) to the MK2-CH.

Figure 11. PWV time series from the near real-time GPS data (black dots), and reprocessed GPS data (gray dots).

Figure 12. PWV time series of PWV from the MWRP, retrieved by using 6 channels, from the MWR, the reprocessed GPS and the DPLX-RS90 radiosondes.

Figure 13. PWV time series from the MWRP, the MWR, the GPS and the DPLX-RS90 and NWS-VIZ radiosondes. Radiosonde outliers are indicated by the arrows.

Table 1. Number of soundings deployed and available after the processing.

Table 2. PWV comparisons among the RAOBs: bias, std, slope, intercept, corr, and sample size. Biases in the comparison of RAOB Y vs RAOB X refer to $PWV_Y - PWV_X$.

Table 3. PWV from MWRP retrieved by using 2 channels (2ch), 5 channels (5ch) and 6 channels (6ch) compared with PWV from MWR. Biases refer to (MWRP-MWR). Sample size is 7221.

Table 4. PWV from GPS compared with PWV from MWR, and the MWRP (retrieved by using 5 channels). Biases refer to PWV(GPS)-PWV(radiometer).

Table 5. PWV from the remote sensors compared with PWV from the radiosondes. Biases refer to PWV(remote sensor)-PWV(RAOB). PWV from the MWRP is retrieved by using the five K-band channels.

Table 6. PWV comparisons during the day (at 2300 UTC) and night (at 1100 UTC) among the RAOBs and the MWR, the GPS, and the MWRP (5-channel retrieval), and among MWR and the other remote sensors. Biases refer to PWV(remote sensor) – PWV(RAOB), and PWV(remote sensor)–PWV(MWR). Bias, std, and int are in cm.

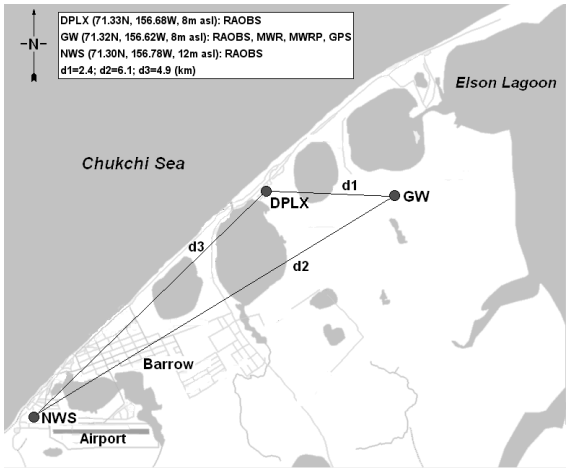

Figure 1. Location and coordinates of ARM "Great White" site (GW), ARM Duplex (DPLX), and NWS upper-air station in Barrow, Alaska.

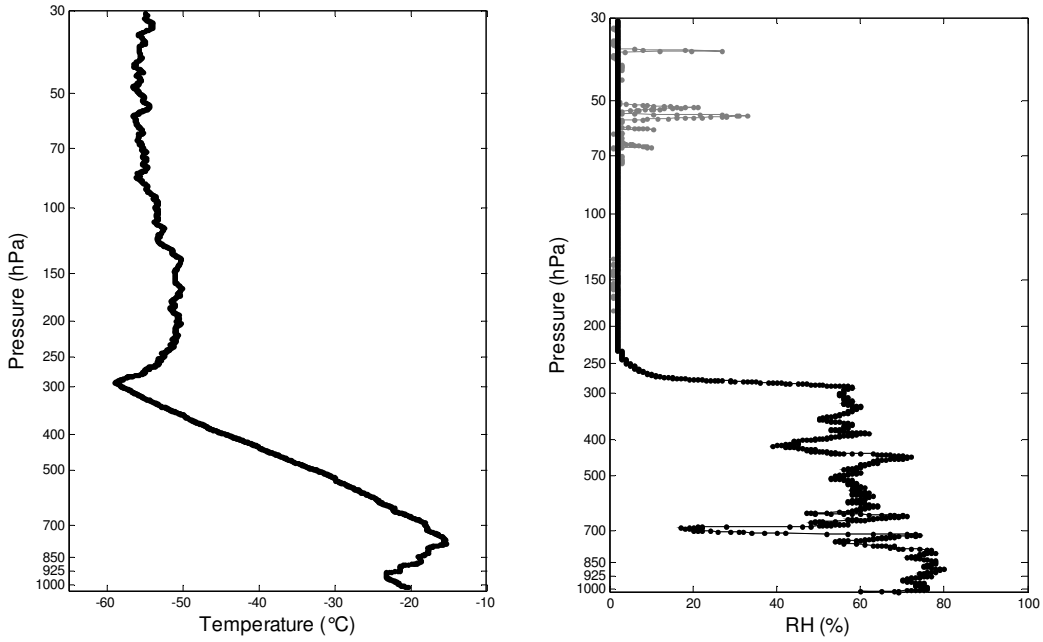

Figure 2. (left) Temperature profile in the DPLX-RS90 sounding launched on April 4, 2004 at 2300 UTC. (Right) Relative humidity affected by unrealistic noise (gray dots) and reconstructed profiles (black dots) in the same sounding. Here, large excursions occur at pressures below 70 hPa.

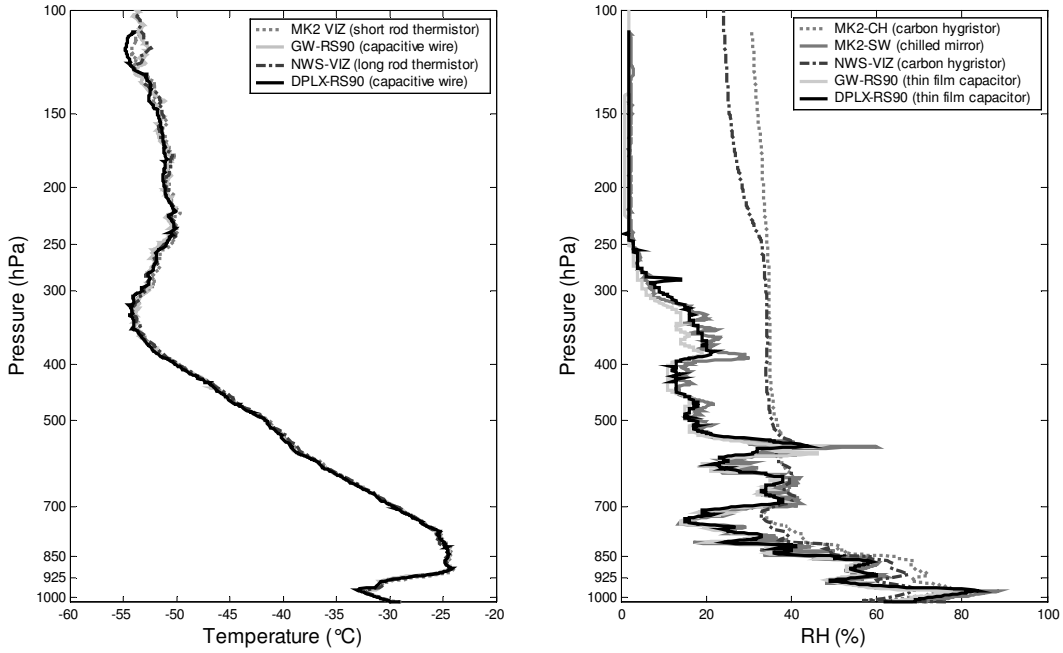

Figure 3. Temperature (left) and relative humidity (right) soundings on March 26 at 2300 UTC.

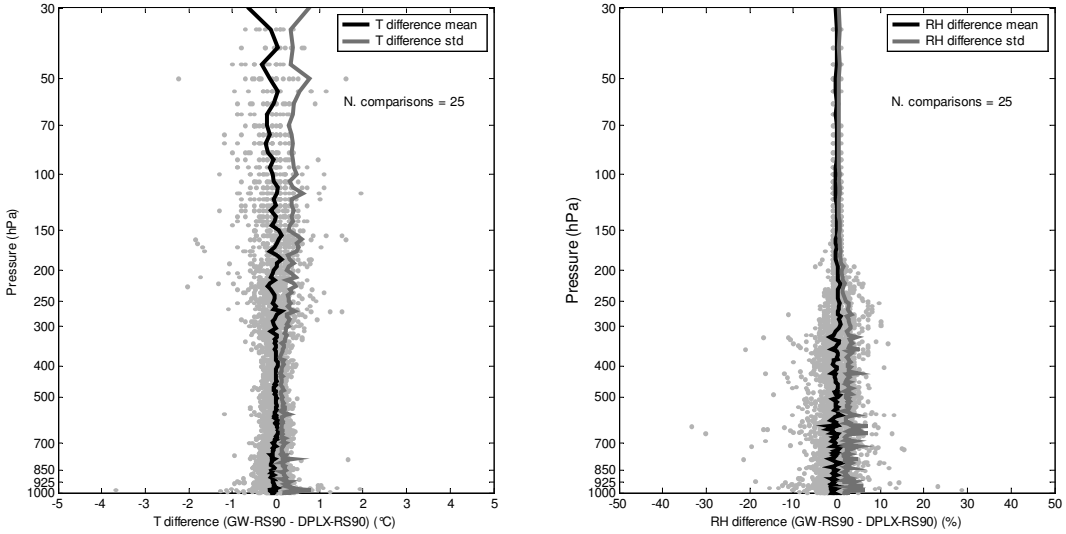

Figure 4. Temperature difference profiles (left) and relative humidity difference profiles (right) between the Vaisala RS90 radiosondes launched at the DPLX and at the GW.

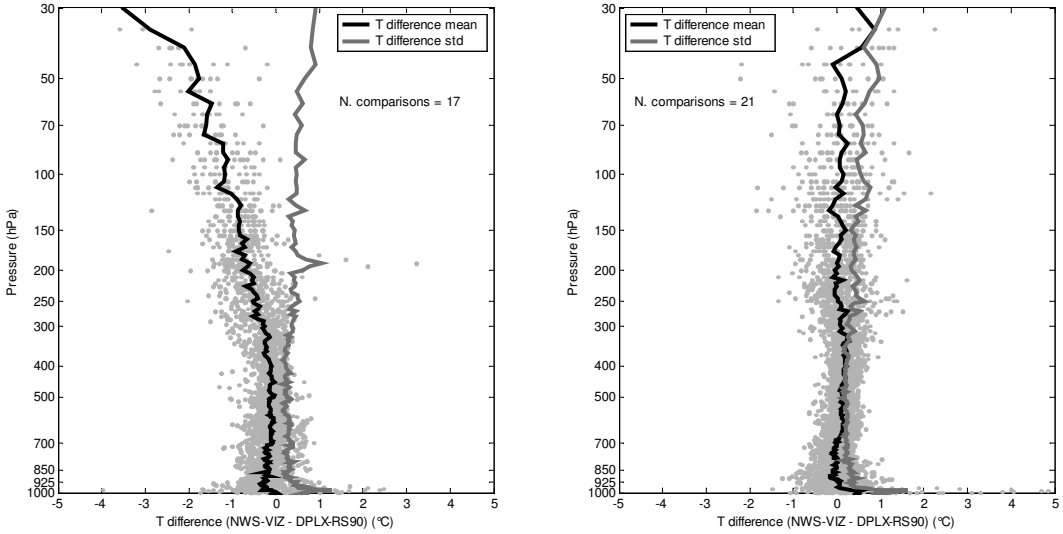

Figure 5. Temperature difference profiles between the NWS-VIZ radiosondes and the DPLX-RS90. (Left) Dataset taken at night (at 1100 UTC, 2 a.m. AST); (right) dataset taken during the day (at 2300 UTC, 2 p.m. AST).

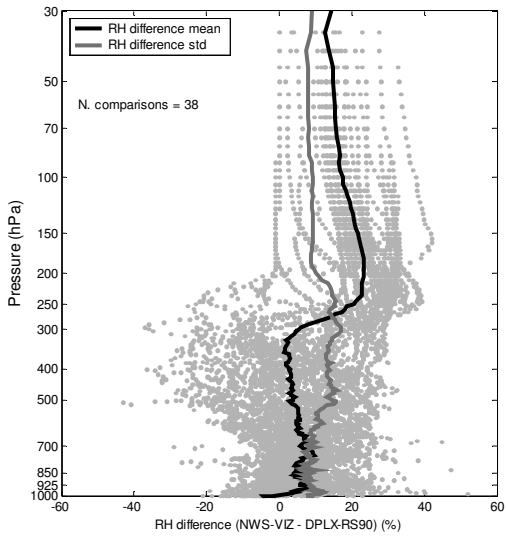

Figure 6. RH difference profiles between the NWS-VIZ radiosondes and the DPLX-RS90, the comparison performed for the entire dataset.

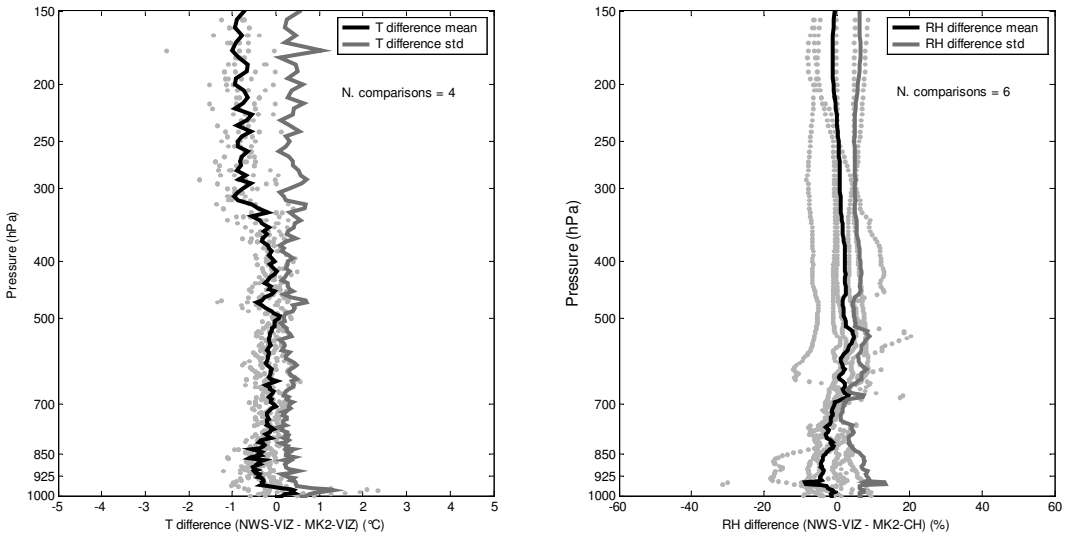

Figure 7. Profiles of temperature and relative humidity difference between the NWS-VIZ and the Mark-II. (Left) NWS-VIZ long rod thermistor vs. MK2 VIZ short rod thermistor, for the dataset taken at night. (Right) NWS-VIZ carbon hygristor vs. MK2-CH, for the entire dataset.

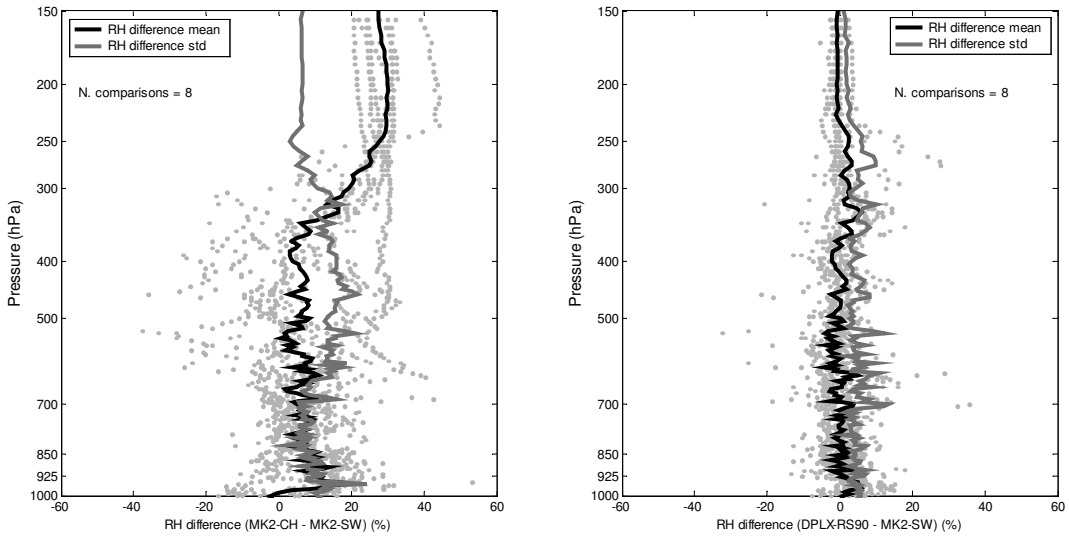

Figure 8. Profiles of RH difference between the VIZ carbon hygristor, the Vaisala capacitor, and the Snow White chilled mirror, for the dual-radiosonde launches. (Left) MK2-CH vs. MK2-SW. (Right) DPLX-RS90 vs. MK2-SW.

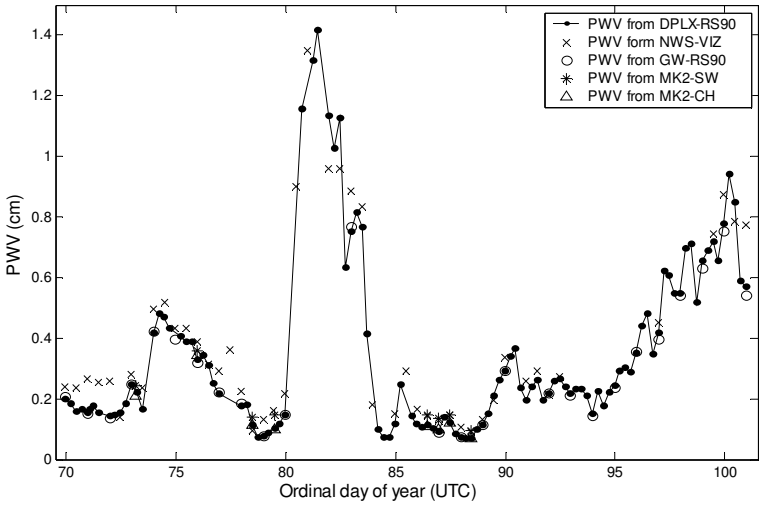

Figure 9. PWV time series computed from the radiosondes that were operating during the 2004 NSA experiment.

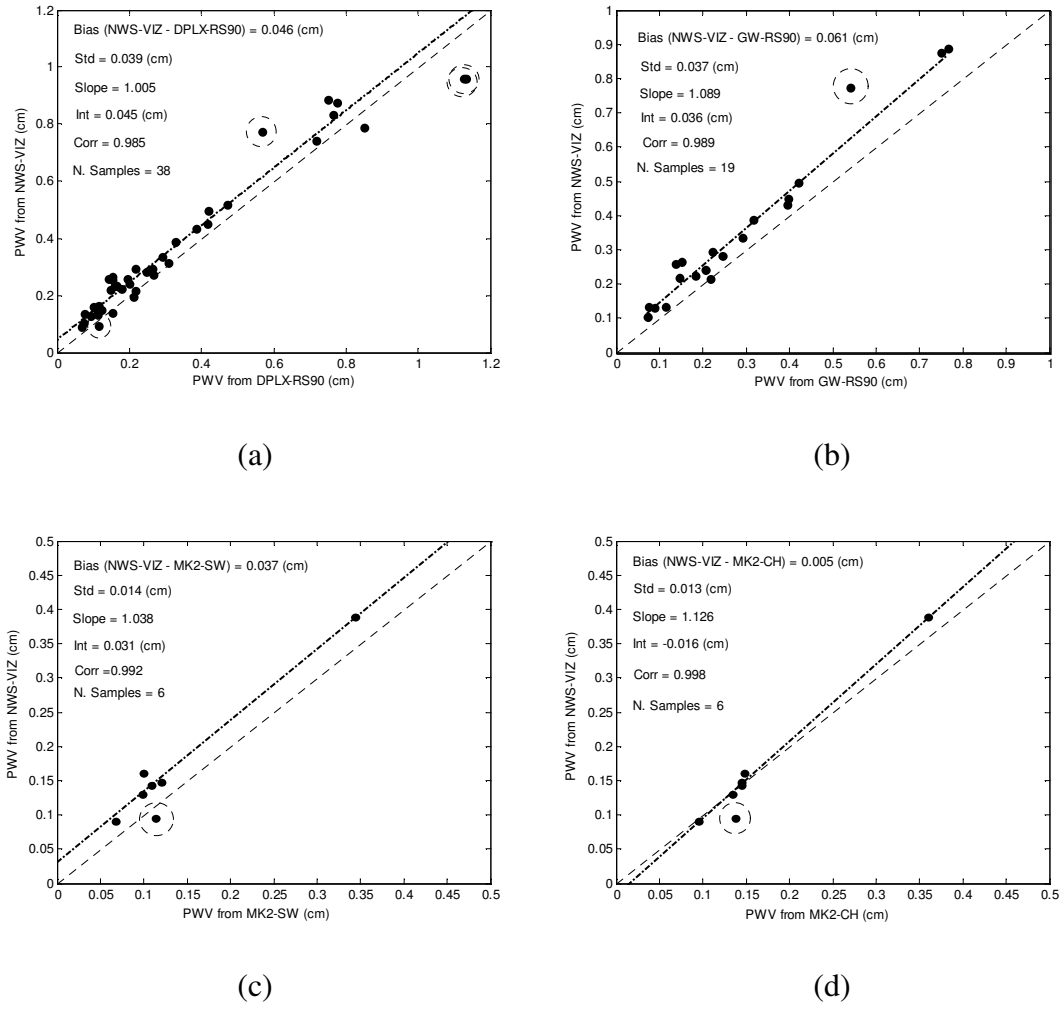

Figure 10. PWV from NWS-VIZ compared with PWV from the other radiosonde types. Outliers are removed from the statistics and shown in the scatterplot enclosed in dotted circles. (a) PWV from NWS-VIZ compared to PWV from DPLX-RS90; (b) to GW-RS90; (c) to MK2-SW; (d) to the MK2-CH.

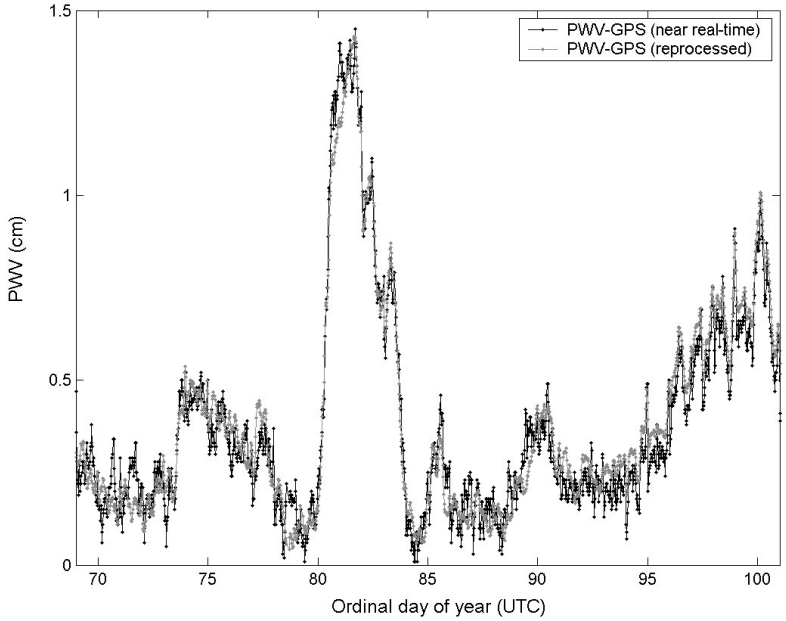

Figure 11. PWV time series from the near real-time GPS data (black dots), and reprocessed GPS data (gray dots).

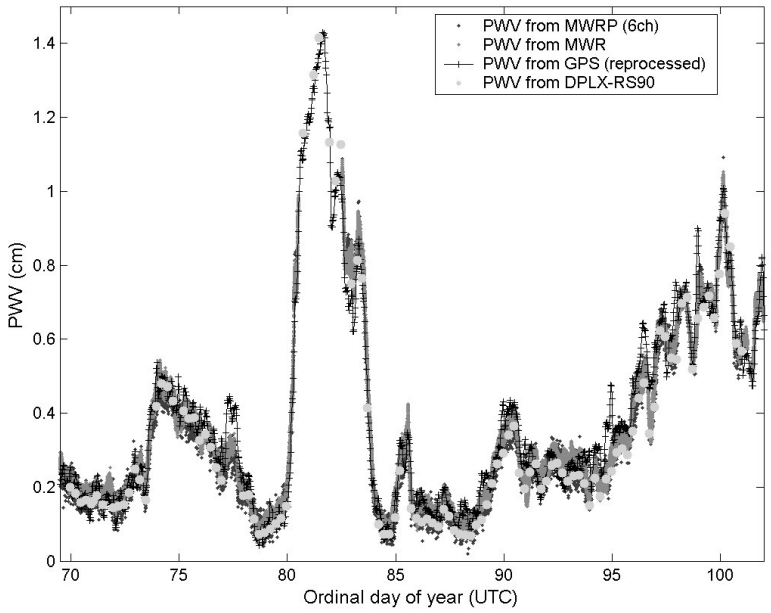

Figure 12. PWV time series of PWV from the MWRP, retrieved by using 6 channels, from the MWR, the reprocessed GPS and the DPLX-RS90 radiosondes.

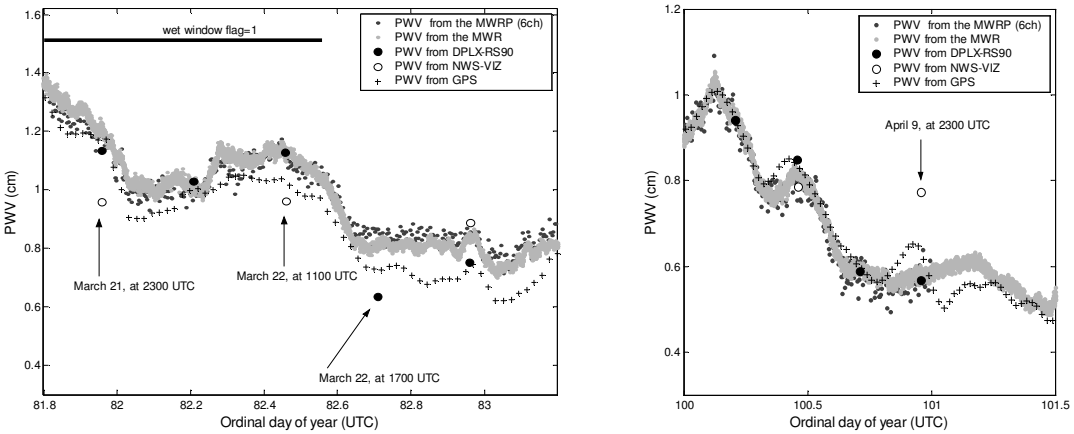

Figure 13. PWV time series from the MWRP, the MWR, the GPS and the DPLX-RS90 and NWS-VIZ radiosondes. Radiosonde outliers are indicated by the arrows.

Table 1. Number of soundings deployed and available after the processing.

RAOBs type	N. launches	Post-processed soundings
DPLX-RS90	124	113
GW-RS90	28	26
MK2 with SW	10	8
NWS-VIZ	51	44

Table 2. PWV comparisons among the RAOBs: bias, std, slope, intercept, corr, and sample size. Biases in the comparison of RAOB Y vs. RAOB X refer to $PWV_Y - PWV_X$.

	bias (cm)	std (cm)	slope	int (cm)	corr	samples
NWS-VIZ vs DPLX-RS90	0.046	0.039	1.005	0.045	0.985	38
NWS-VIZ vs GW-RS90	0.061	0.037	1.089	0.036	0.989	19
GW-RS90 vs DPLX-RS90	-0.005	0.010	0.979	0.002	0.999	25
NWS-VIZ vs MK2-SW	0.037	0.014	1.038	0.031	0.992	6
NWS-VIZ vs MK2-CH	0.005	0.013	1.126	-0.016	0.998	6
MK2-CH vs MK2-SW	0.031	0.010	0.934	0.040	0.995	8
MK2-CH vs DPLX-RS90	0.031	0.009	0.984	0.034	0.994	8
DPLX-RS90 vs MK2-SW	-0.0004	0.009	0.951	0.007	0.996	8
NWS-VIZ vs DPLX-RS90 (Daytime dataset)	0.058	0.034	1.088	0.036	0.991	21
NWS-VIZ vs DPLX-RS90 (Nighttime dataset)	0.031	0.040	0.955	0.045	0.987	17

Table 3. PWV from MWRP retrieved by using 2 channels (2ch), 5 channels (5ch) and 6 channels (6ch) compared with PWV from MWR. Biases refer to (MWRP-MWR). Sample size is 7221.

	MWR				
	Bias (cm)	Std (cm)	Slope	Int (cm)	Corr
MWRP (2ch)	-0.018	0.038	0.975	-0.011	0.979
MWRP (5ch)	-0.018	0.028	0.983	-0.013	0.989
MWRP (6ch)	-0.015	0.029	1.009	-0.017	0.989

Table 4. PWV from GPS compared with PWV from MWR, and the MWRP (retrieved by using 5 channels). Biases refer to PWV(GPS)-PWV(radiometer).

	GPS					
	Bias (cm)	Std (cm)	Slope	Int (cm)	Corr	Samples
MWR	0.005	0.049	0.981	0.011	0.973	1419
MWRP	0.026	0.052	0.994	0.028	0.963	1295

Table 5. PWV from the remote sensors compared with PWV from the radiosondes. Biases refer to PWV(remote sensor)-PWV(RAOB). PWV from the MWRP is retrieved by using the five K-band channels.

			MWR			
	Bias (cm)	Std (cm)	Slope	Int (cm)	Corr	Samples
NWS-VIZ	-0.014	0.039	0.968	-0.004	0.984	42
DPLX-RS90	0.030	0.022	0.988	0.034	0.994	104
GW-RS90	0.048	0.030	1.077	0.025	0.993	25
			MWRP			
NWS-VIZ	-0.036	0.039	0.931	-0.016	0.981	40
DPLX-RS90	0.013	0.025	1.001	0.012	0.992	100
GW-RS90	0.025	0.034	1.097	-0.005	0.992	26
			GPS			
NWS-VIZ	-0.013	0.064	0.913	0.019	0.973	44
DPLX-RS90	0.037	0.059	0.939	0.058	0.979	112
GW-RS90	0.092	0.070	1.111	0.058	0.958	26

Table 6. PWV comparisons during the day (at 2300 UTC) and night (at 1100 UTC) among the RAOBs and the MWR, the GPS, and the MWRP (5-channel retrieval), and among MWR and the other remote sensors. Biases refer to PWV(remote sensor) – PWV(RAOB), and PWV(remote sensor)–PWV(MWR). Bias, std, and int are in cm.

	DPLX-RS90		NWS-VIZ		MWR	
	day	night	day	night	day	night
MWR	Bias= 0.045 Std= 0.023 Slope= 1.065 Int= 0.027 Samples= 27	Bias= 0.020 Std= 0.024 Slope= 0.947 Int= 0.037 Samples= 26	Bias= -0.019 Std= 0.040 Slope= 0.954 Int= -0.005 Samples= 23	Bias= -0.009 Std= 0.037 Slope= 0.978 Int= -0.001 Samples= 19		
GPS	Bias= 0.081 Std= 0.065 Slope= 1.021 Int= 0.075 Samples= 30	Bias= -0.0004 Std= 0.049 Slope= 0.936 Int= 0.025 Samples= 28	Bias= 0.009 Std= 0.072 Slope= 0.891 Int= 0.047 Samples= 24	Bias= -0.038 Std= 0.041 Slope= 0.954 Int= -0.021 Samples= 20	Bias=0.026 Std= 0.054 Slope=1.004 Int= 0.025 Samples=116	Bias= -0.019 Std= 0.046 Slope= 0.972 Int= -0.008 Samples= 120
MWRP	Bias= 0.022 Std= 0.027 Slope= 1.075 Int= 0.001 Samples= 29	Bias= 0.006 Std= 0.022 Slope= 0.933 Int= 0.024 Samples= 23	Bias= -0.037 Std= 0.037 Slope= 0.966 Int= -0.026 Samples= 23	Bias= -0.035 Std= 0.042 Slope= 0.867 Int= 0.003 Samples= 17	Bias= -0.018 Std= 0.030 Slope=1.020 Int= -0.024 Samples= 638	Bias= -0.020 Std= 0.027 Slope= 0.955 Int= -0.007 Samples= 590