A Neural Network for Real-Time Retrievals of PWV and LWP From Arctic Millimeter-Wave Ground-Based Observations

Maria P. Cadeddu, David D. Turner, and James C. Liljegren

Abstract—This paper presents a new neural network (NN) algorithm for real-time retrievals of low amounts of precipitable water vapor (PWV) and integrated liquid water from millimeter-wave ground-based observations. Measurements are collected by the 183.3-GHz G-band vapor radiometer (GVR) operating at the Atmospheric Radiation Measurement (ARM) Program Climate Research Facility, Barrow, AK. The NN provides the means to explore the nonlinear regime of the measurements and investigate the physical boundaries of the operability of the instrument. A methodology to compute individual error bars associated with the NN output is developed, and a detailed error analysis of the network output is provided. Through the error analysis, it is possible to isolate several components contributing to the overall retrieval errors and to analyze the dependence of the errors on the inputs. The network outputs and associated errors are then compared with results from a physical retrieval and with the ARM two-channel microwave radiometer (MWR) statistical retrieval. When the NN is trained with a seasonal training data set, the retrievals of water vapor yield results that are comparable to those obtained from a traditional physical retrieval, with a retrieval error percentage of ∼5% when the PWV is between 2 and 10 mm, but with the advantages that the NN algorithm does not require vertical profiles of temperature and humidity as input and is significantly faster computationally. Liquid water path (LWP) retrievals from the NN have a significantly improved clear-sky bias (mean of ∼2.4 g/m^2) and a retrieval error varying from 1 to about 10 g/m^2 when the PWV amount is between 1 and 10 mm. As an independent validation of the LWP retrieval, the longwave downwelling surface flux was computed and compared with observations. The comparison shows a significant improvement with respect to the MWR statistical retrievals, particularly for LWP amounts of less than 60 g/m^2. This paper shows that the GVR alone can provide overall improved PWV and LWP retrievals when the PWV amount is less than 10 mm, and, when combined with the MWR, can provide improved retrievals over the whole water-vapor range.

I. INTRODUCTION

RECENTLY, there has been an increased interest in the Arctic climate and its effects on the Earth's radiation budget. Several studies have focused on the retrieval of microphysical properties of mixed-phase clouds and on the effect of Arctic cloud forcing on the Earth's radiation budget [1]. One of the challenges posed by the study of the Arctic climate is that water vapor is very low during the winter, with integrated water-vapor amounts rarely exceeding 5 mm and cloud liquid water amounts rarely exceeding 100 g/m^2. In these conditions, the errors associated with retrievals from measurements in the 20- to 30-GHz region of the microwave spectrum have uncertainties that can be as high as 50% for water vapor and 30% or more for liquid water path (LWP). On the other hand, the longwave and shortwave downwelling fluxes at the surface have a very strong dependence on the amount of water in clouds and are particularly sensitive to LWP values of 30 g/m^2 or less [1]. It has also been established [2] that water, in the phase of supercooled liquid, exists in the majority of Arctic clouds at temperatures as low as −30 °C, strongly reinforcing the need for improved accuracy in the retrievals. Turner *et al.* [3] recently published an overview of the challenges associated with the retrieval of low amounts of liquid water.

Currently, real-time precipitable water vapor (PWV) and LWP retrievals at the Atmospheric Radiation Measurement Program Climate Research Facility (ACRF), Barrow, AK, are provided by the two-channel microwave radiometer (MWR) operating at center frequencies of 23.8 and 31.4 GHz. This portion of the atmospheric spectrum is mostly sensitive to water vapor and liquid water, with a residual influence coming from the 60-GHz oxygen absorption region. Since the integrated amounts of liquid and vapor are linearly related to atmospheric opacity, accurate retrievals from these channels can be achieved by linear regression. The MWRs have been proven to be reliable and accurate instruments under a wide range of atmospheric conditions; however, in the Arctic, the atmosphere is extremely dry during the winter, and the signal-to-noise ratio of the measurements is very low. An additional source of uncertainty associated with the current MWR retrievals is the effect of the modeling of the oxygen absorption line. This uncertainty, in the high-pressure and low-humidity conditions often present in the Arctic, can cause a positive bias as high as 25 g/m^2 in the clear-sky LWP retrievals [4].

Recent computations of the longwave and shortwave surface radiation budget have shown that when LWP retrievals from the MWR are used to parameterize the radiative properties of the atmosphere, the computed downwelling longwave and shortwave fluxes are considerably higher than observations [5], suggesting an overall overestimation of liquid water [2], [6]. The need for improved retrieval accuracy has driven a considerable effort that has led to the development of more sophisticated ground-based retrievals. Among them is the algorithm of Turner [5], which combines inputs from several instruments (including infrared and MWRs, a ceilometer, and radiosondes) to enhance the sensitivity to small amounts of liquid water. The deployment of the G-band vapor radiometer (GVR) in Barrow has also been a response to the need for increased accuracy in the retrievals. The instrument has been operating since its deployment in April 2005, providing continuous measurements of brightness temperatures in the region of the strong 183.3-GHz water-vapor absorption line.

This paper presents an neural network (NN) algorithm to retrieve low amounts of liquid water and water vapor from GVR measurements. The algorithm can provide real-time improved retrievals of vapor and liquid under most conditions encountered in the Arctic, and it can be used as an improved comparison point for more sophisticated multi-instrument retrievals, such as the one proposed in [5]. It shows that NN retrievals can take advantage of the high sensitivity of the GVR channels to small amounts of PWV, reaching percentage retrieval errors of 5% when the PWV is between 2 and 10 mm and substantially reducing the retrieval bias. The LWP retrievals show an improvement in the clear-sky discrimination by reducing the clear-sky bias to less than 5 g/m² (as compared to the 15 g/m² of the MWR retrievals) and reducing the differences between modeled and measured longwave downwelling surface fluxes.

The remainder of this paper is structured as follows. Section II provides a brief overview of the instrument and the measurements as they pertain to the retrieval algorithm (a more detailed description of the instrument and its calibration can be found in previous papers [7], [8]). A detailed discussion on the effect of stratospheric water vapor on the measurements [9] is not provided since this issue was extensively examined in [8] and [10] and found not to be a major concern at the GVR frequencies. In Section III, the retrieval methodology is presented for both the NN and the physical retrieval, and the advantages and disadvantages of the NN approach are briefly discussed. Section IV describes the network structure, training data set, and training procedure. The network is trained with GVR-only data, and with a combination of GVR and MWR data. A significant component of this paper is the realistic treatment and quantification of the retrieval errors associated with the NN; this is described in Section V. Finally, in Sections VI and VII, the retrieval results are presented together with a detailed error analysis. Water vapor retrievals are compared with physical retrievals and with the MWR linear retrievals. Liquid water retrievals are compared with MWR retrievals and tested by comparing the computed longwave downwelling radiation flux at the surface with observations. Section VIII provides a summary of this paper and describes future work plans.
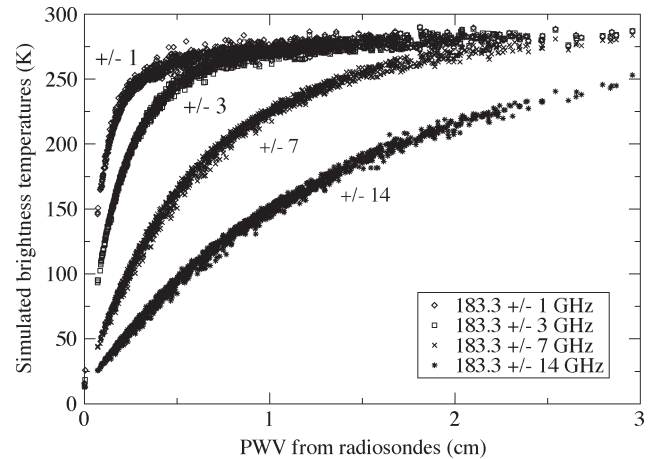

Fig. 1. Brightness temperatures simulated at the four GVR channels as a function of PWV.

II. INSTRUMENT AND MEASUREMENTS

The GVR, built by ProSensing under a U.S. Department of Energy small business innovation research grant, was deployed at the ACRF in April 2005. It has been fully operational since January 2006. The instrument is a double-sideband MWR with channels located ± 1, ± 3, ± 7, and ± 14 GHz away from the 183.3-GHz line center. A description of the instrument can be found in [7]. The instrument's parabolic mirror sequentially points to the sky, a hot load, and an ambient load to achieve accurate calibration. Calibration uncertainty on the GVR channels has been estimated to be less than 1 K for the channels close to the line center and less than 2 K for the more transparent channels in the wing of the absorption line [8].

The sensitivity of channels in the proximity of the 183.3-GHz line to PWV and LWP has been studied in [8], [9], and [11]. The findings showed that the sensitivity of GVR channels to water vapor is approximately 30 times higher than the sensitivity of the MWR when the PWV is less than 2.5 mm. Comparison of simulated and measured brightness temperatures also showed that two channels (± 7 and ± 14 GHz) have a sensitivity to liquid water that is about 2 and 3.5 time higher than that of the 31.4-GHz channel. Several retrieval algorithms to retrieve vapor and liquid from measurements at these frequencies have been recently developed. Specifically, a physical retrieval for liquid and vapor [8] and a linear retrieval that uses only channels in the wing of the line [11] have been developed. The physical retrieval has the disadvantage of requiring a temperature and humidity profile as input. The linear retrieval has the disadvantage of retrieving only in the linear region of the measurements when the PWV is very low (less than 3 mm). As shown in Fig. 1, the response of the ± 1- and ± 3-GHz GVR channels to PWV starts to become nonlinear when the PWV reaches about 2.5 mm, and the two channels saturate around 5 mm. The more transparent channels (± 7 and ± 14 GHz) still retain sensitivity until the PWV amount is about 20 mm. Above 20 mm, only the ± 14-GHz channel is appreciably sensitive. In this paper, we consider 20 mm as a physical water-vapor threshold above which the instrument will not be able to separate the vapor from the liquid phase, and we therefore limit our discussion to cases that fall below this threshold. It is important to notice that, when clouds are present, the 183.3 ± 3-GHz

channel will saturate when the water-vapor amount is less than 20 mm. From model simulations, it was determined that, when the water vapor is between 15 and 20 mm, an LWP of about 80 g/m^2 is sufficient to saturate this channel. Therefore, when the PWV is higher than 15 mm, the LWP retrievals will be reliable only for clouds with very small LWP.

Although the instrument was specifically designed for retrieving low amounts of water vapor, the enhanced sensitivity to liquid water displayed by the two most transparent channels enables us to develop an algorithm for the retrieval of LWP as well. Of course, the retrieval of liquid water is influenced by the instrument's sensitivity to vapor.

The measurements analyzed in this paper were collected in the winter and summer of 2007. Radio-frequency interference (RFI) from a nearby U.S. Navy Radar strongly affected the GVR measurement. Although several filters have been applied to the raw and processed data, the effect of the interference is still present. Data used in the comparison with radiosonde measurements were filtered with a 5-min running median filter to smooth interference spikes. Additional instruments used in this paper include an infrared thermometer (IRT) to identify clear-sky situations and a sky radiometer equipped with a shaded pyrgeometer with a hemispheric field of view to measure the longwave downwelling irradiance.

The MWR is a two-channel microwave radiometer that provides continuous retrievals of integrated water vapor and liquid water. The instrument is continuously calibrated through an automatic calibration algorithm that performs continuous tip curve collection when the sky is clear and that continuously updates the calibration parameters. The estimated calibration accuracy of this instrument is 0.3 K [12].

III. Retrieval Algorithms

NNs have been extensively used to retrieve geophysical quantities from satellite measurements [13], [14]. They have proven to be reliable statistical techniques that, in some cases, are preferable to computationally expensive variational retrievals. For example, once the network presented in this paper is trained, it takes approximately 20 min to process a full month of data at the highest temporal resolution (about 300 000 measurements), whereas it would take about one month to process the same amount of data with the physical retrieval developed in [8]. In the case of water vapor and liquid water retrievals from GVR measurements, the use of an NN allows us to overcome some of the shortcomings of the already existent retrievals; namely, the poor time resolution of the physical retrieval (due to the need for radiosonde profiles) and the nonlinear nature of the problem. With an NN, we can fully explore the nonlinear regime of the measurements, investigating the performance of the retrieval up to a PWV amount of 20 mm. Of course, the NN itself is a statistical algorithm, and it has some of the same shortcomings as those associated with the linear statistical retrievals. For example, for the network developed in this paper that uses only brightness temperatures as input, we expect the standard deviation of the results to be higher than what would be expected from a physical retrieval, since the latter is more constrained by the vertical profiles used as input. When we evaluate the performance of the network, it is therefore important to

assess whether the advantage of real-time retrievals is offset by a degradation of the accuracy of the results.

One of the problems encountered by the remote sensing community in using NN algorithms is the difficulty in assessing the errors associated with the network output. This fact has traditionally constituted one of the reasons that a physical retrieval is preferred, since error bars can be promptly estimated by analyzing the covariance associated with the output [15]. Usually, errors on the NN outputs are estimated by computing the total root mean square error (rmse) over the training data set [16]. This approach, while legitimate in a linear case, does not allow the computations of error bars associated with the single output, and it therefore fails to properly represent the algorithm performance in a case when we expect a deterioration of the performance for a certain range of the input variables. The GVR retrievals certainly fall in this category. Since the GVR channels progressively lose sensitivity to water vapor as PWV increases, we correspondingly expect a deterioration of the retrievals. To address this point, we developed a procedure to compute error bars associated with the network outputs by treating the various sources of errors independently, and accounting for sources that were considered to be the largest contributors to the overall uncertainty. Although (as we will show later) the methodology might slightly underestimate some of the errors, the estimated values are in good agreement with the errors associated with the physical retrieval algorithm.

The physical algorithm used in this paper was presented in [8]. It uses a Gauss–Newton method that finds the zeros of the gradient of a cost function. The minimization is achieved by successive iterations, starting from a first-guess profile of temperature, relative humidity, and liquid water. Although continuously retrieved temperature and humidity profiles are available from a colocated 12-channel microwave profiler, they were not used in this paper because uncertainties in the upper tropospheric layers of the retrieved temperature profile could introduce considerable errors in the physical retrievals [8]. Therefore, the use of the physical retrieval was limited to the times of radiosonde ascent.

The third retrieval algorithm used in this paper is the MWR RETrieval (MWRRET) [17]. The algorithm is a linear statistical regression that dynamically estimates possible biases present in the 23.8- and 31.4-GHz brightness temperatures before proceeding with the retrievals. The MWRRET uses all-year coefficients derived from an extensive set of radiosonde profiles. Estimated retrieval rms uncertainty with this algorithm is around 0.4 mm for water vapor and around 15 g/m^2 for the LWP.

IV. NN Description

An NN provides a nonlinear approach to map a set of input variables $x_i (i = 1, \ldots, N_i)$ to a set of output variables y_j, provided a set of targets $t_j (j = 1, \ldots, N_o)$

$$y_{ij} = F(x_{ik}, w_{ij}) + \varepsilon_{ij} \tag{1}$$

where the index i indicates the elements of the training set $(i = 1, \ldots, N)$, j varies between one and the number of output variables N_o, and k varies between one and the number of input variables N_i. The vector ε represents the output errors

TABLE I
NN ARCHITECTURE AND FINAL rms ERROR

	GVR NN retrieval			GVR + MWR NN retrieval	MWRRET retrieval
Output	PWV		LWP	PWV	PWV and LWP
Input frequencies (GHz)	183.3 ±1, ±3, ±7, ±14		183.3 ±1, ±3, ±7, ±14	183.3 ±1, ±3, ±7, ±14, 23.8, 31.4	23.8, 31.4
Number of hidden nodes	20		20	20	-
Number of weights	100		100	140	-
PWV RMS error (mm)	All 0.3 Winter 0.2 Summer 0.3		-	0.1	~ 0.4
LWP RMS error (g/m²)	-		Winter 6.4 Summer 8.9	-	~15.0

and is discussed in greater detail later. The vector $\mathbf{w}$ represents the network weights that will be optimized during the learning process (w_{ij} being the weight connecting the element j to the element i). A network is usually composed of one input layer, some intermediate or hidden layers, and one output layer. In our case, each element of the input layer is connected to each element of the hidden layer through a logistic function

$$h_i = \frac{1}{1 + \exp(-a_i)}, \qquad i = 1, \ldots, N_{\mathrm{h}} \qquad (2)$$

where N_{h} is the number of hidden nodes and a_i is

$$a_i = \sum_{j=1}^{N_{\mathrm{i}}} w_{ij} x_j + b_i. \qquad (3)$$

Each element of the hidden layer is connected to the output layer through a function as in (2) or through a linear function

$$y_j = \sum_{k=1}^{N_{\mathrm{h}}} w_{jk} h_k + b_j \qquad (4)$$

where the vector $\mathbf{b}$ is called bias. The estimation of the optimal solution is achieved by optimizing the set of the weights $w_m (m = 1, \ldots, N_{\mathrm{w}})$ by minimizing a quality criterion, which, in this case, is the least squares criterion

$$E = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N_{\mathrm{o}}} (y_{ij} - t_{ij})^2. \qquad (5)$$

A general network architecture with one input layer (with four or six nodes depending on the number of channels), one output layer with one node, and one hidden layer with 20 nodes was considered sufficient for the purpose of this paper. The choice of this simple architecture is based on previous work on similar data (e.g., [13]). PWV and LWP were retrieved with two separate networks, since the use of a one-output network noticeably simplifies the error computations. For the PWV, two separate networks were tested. One network was trained by using only GVR measurements and one network was trained by using a combination of GVR and MWR measurements. The LWP network was trained with only GVR measurements. Following the initial analysis, it was deemed important to train

the GVR network (for both PWV and LWP) by using a seasonal rather than an all-year training data set. The structure of the four networks tested in this paper is shown in Table I. The elements of the input layer are brightness temperatures for various selected channels. The output layer has one element that is either PWV or LWP. In both networks, all layers are connected through the logistic function (2). The choice of the number of hidden nodes and connections is somewhat arbitrary as long as the network has enough connections to reproduce the complexity of the process it has to learn. Several network configurations, with different numbers of hidden nodes were tried. The network was trained repeatedly starting with five hidden nodes and progressively adding five nodes up to 25. For each network configuration, the following parameters were compared: the final rmse, the covariance of the final weights (which defines the network uncertainty as explained later) and the time needed for training. The rmse did not change noticeably when more nodes were added; however, the variability of the weights and the network uncertainty decreased. On the other hand, the time needed to train the network increased as the number of hidden nodes increased. At the end of the simulations, a network with 20 nodes was considered a good compromise between the three parameters. After the training, network weights that have very small values can be eliminated. In any case, the choice of network architecture should not have a large impact on the results, and we account for the uncertainty associated with the network weights in the computation of the output error. The Stuttgart NN Simulator (SNNS) (available at http://www-ra.informatik.uni-tuebingen.de/software/snns/welcome_e.html) was used to train the network. The SNNS was widely used in previous work involving retrievals of geophysical parameters from ground-based and satellite data [16], [18]. A conjugate descent algorithm was chosen to find the optimal solution [19].

A. Training Data Set

The training data set consists of brightness temperatures simulated with a forward radiative transfer model [20] from a six-year ensemble (1633 cases) of radiosonde soundings (years 2000 to 2006) collected in Barrow. Three liquid path amounts were simulated for each temperature and humidity profile when the relative humidity at a given layer exceeded 95%. The final training data set so obtained consists of 4854 patterns. In
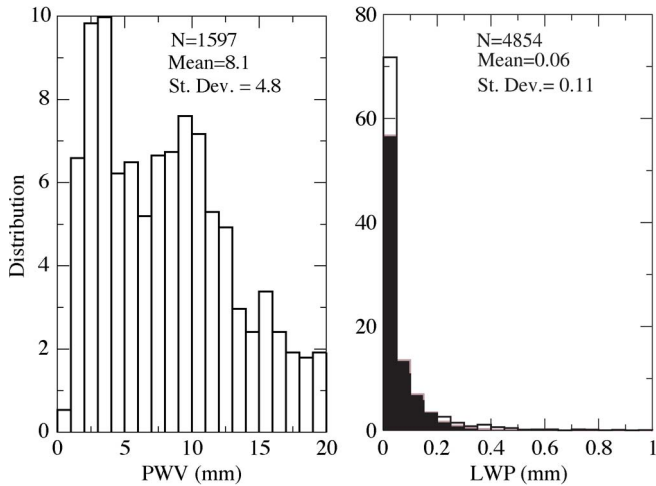
Fig. 2. (Black line) Histograms of PWV and LWP in the training data set expressed as a percentage. The gray distribution in the right panel is from one year of MWR-retrieved LWP.
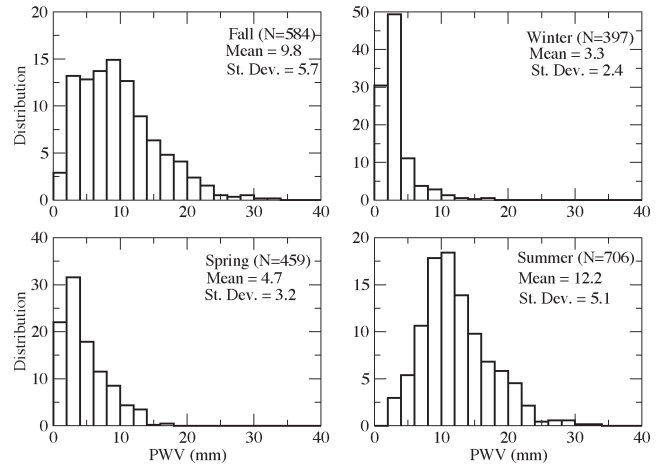

Fig. 3. Seasonal histograms of PWV in the training data set expressed as a percentage. The four partially overlapping subsets include four months of data: November through February (winter), February through May (spring), May through August (summer), and August through November (fall).

addition to the training data set, an independent testing data set was used to monitor the optimization procedure. This testing set was computed from 550 radiosondes collected in 2007, for a total of 1638 simulations, including cloudy cases. Gaussian noise was added to the simulated brightness temperatures in the training and testing sets (1 K for the GVR channels and 0.5 K for the MWR channels).

The training data set is important because it should represent the climatic variability of the site of interest very well. The performance of the retrieval algorithm depends partially on how well the events are represented in the training set, since the network will poorly reproduce rare or infrequent events. As shown in Fig. 1, the four GVR channels progressively lose their sensitivity as the PWV increases, and, as explained in Section II, when the PWV amount reaches 20 mm, three of the four channels are saturated. For this reason, a physical limit of 20 mm was set as the maximum retrievable PWV from GVR-only measurements, and only these cases were used to train the network ($N = 1597$).

Fig. 2 shows the probability distribution of PWV (left side) and LWP (right side) for the whole training data set, expressed as a percentage. About 65% of cases in the training data set have PWV amounts of less than 10 mm (half of which have a PWV of less than 5 mm). This is also the PWV range where the GVR channels have the best sensitivity. To assess how realistically the LWP is represented in the training data set, we show in Fig. 2 (right panel) the probability distribution of one year of LWP retrievals from the two-channel MWR at the NSA (histogram in brown). In about 83% of cases, the LWP amount of the training data set is less than 0.1 mm (100 g/m^2), which compares well with the 72% of cases of the MWR-retrieved LWP distribution. While the LWP does not display a strong seasonal variation (except for a slightly higher average in summer and fall), the PWV distribution has a marked seasonality, as shown in Fig. 3. Winter (November–February) and spring (February–May) are drier, with an average PWV amount of 3 and 5 mm, respectively. This seasonal variation of water vapor is due to increased local evaporation in summer and fall, caused by more solar radiation, as well as changes in the atmospheric circulation and ice cover over the nearby ocean.

The PWV seasonality affects the results of the training. When trained with the full data set, the NN attempts to find the best fit over the whole PWV range. We can therefore expect a generally higher rms error. However, since the GVR is specifically designed for low amounts of PWV, seasonal training will provide better results during dry conditions. For this reason, the GVR network was first trained with the full (all-year) data set and then with two seasonal data sets, one for winter–spring and one for summer–fall. The seasonal training is expected to give better retrievals for low amounts of PWV by taking better advantage of the highly sensitive channels closer to the line center and improving the overall rms error. The LWP network was trained seasonally as well, since the response of the network depends on the amount of water vapor. The combined GVR + MWR network was trained with the all-year data set.

B. Training Procedure

When training of the network is completed, the resulting vector of network weights $\mathbf{w}^*$ is the optimal solution that minimizes the rms error defined in (5). To avoid "overtraining" during the minimization procedure, we periodically tested the network on the independent testing set mentioned in Section IV-A. Fig. 4 (top panel) shows an example of such a training curve for the GVR PWV network, trained with the seasonal winter data set. The curve is shown up to the point when the training was stopped ($\sim 7 \cdot 10^5$ iterations). The dashed line represents the testing data set. In the bottom panel, the backward differences of the rms errors are displayed to better illustrate the flattening of the training curve up to the point when the training was stopped. Similar curves are obtained for all the other networks. To stop the training, the descent of the testing data set was monitored and the training was stopped when the rms error of the testing set decreased of less than a fixed threshold (usually between 1% and 2% depending on the network) for a selected number of iterations. An additional parameter that was used to assess the completeness of the training was the behavior of the Hessian matrix computed
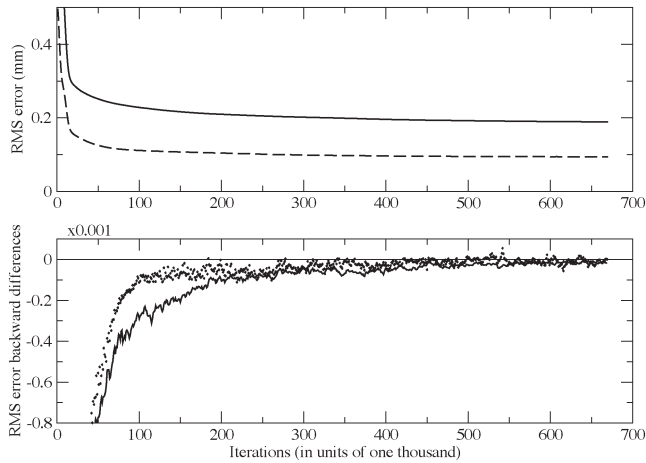
Fig. 4. (Top) Training curve of the GVR-only network trained with the seasonal winter data set. The solid line is the training data set, and the dashed line is the testing data set. (Bottom) Backward differences between elements of the rms error curve for the (solid) training and (dots) testing sets.

after the training. The Hessian matrix represents the second derivative of the error function with respect to the network weights. It is discussed in more detail on Section V-B. It should be definite positive for a well-trained network.

The final rms errors for all tested networks are shown in Table I. For PWV, the combination of the GVR and MWR measurements improves the error computed over the whole data set. This improvement is due to the fact that MWR channels retain their sensitivity through a larger range of water-vapor conditions, supplementing the loss of sensitivity of GVR channels. However, as is shown later, the GVR retrieval errors are input dependent, while the GVR + MWR errors display less sensitivity to the inputs. Therefore, an error estimate over the whole data set does not properly represent the performance of the GVR retrieval, and it is important to introduce a more complete treatment that can better illustrate the dependence of the errors on the amount of water vapor.

V. NETWORK RETRIEVAL ERRORS

The results of the NN retrievals are affected by a variety of factors. In our analysis, we partially follow [21] and assume that the total retrieval error σ_T^2 is the sum of three independent components: the data target variance σ_t^2, the model uncertainty σ_m^2, and the instrument noise σ_N^2. The data target variance is usually the largest contributor to the overall retrieval uncertainty, and it is assumed here to be input dependent. It is directly related to the intrinsic variability of the training data set and includes radiative transfer and radiosonde uncertainties as well. Radiative transfer uncertainties derive from the spectral modeling of the line absorption parameters and water-vapor continuum. The radiative transfer model used in this paper uses the HITRAN database line spectral file [22] and the CKDv2.4 water-vapor continuum [23], where the latter is most likely the largest source of uncertainty for the ±7- and ±14-GHz simulations. The liquid water model used is Liebe 93 [24]. The accuracy of the liquid water absorption model at frequencies higher than 90 GHz suffers from lack of experimental data in supercooled conditions. Uncertainties in the liquid absorption model will affect the retrievals more when the cloud tempera-

ture is below $-20\ °C$, which is not an uncommon occurrence in winter. Uncertainty due to radiosonde soundings can derive from incorrect soundings of stratospheric layers (where the humidity is very low) and from the effect of a dry bias. Errors in the sounding of stratospheric humidity of Vaisala RS90 radiosondes have been extensively investigated in [8] and found to have a minimal effect on the measurements, when the water vapor exceeds 1 mm. The effect of the dry bias will impact the training of the network and will have a larger effect for daytime soundings.

Model uncertainties represent the errors associated with the determination of the optimal weights and are usually a small part of the total retrieval errors. This error component is estimated by computing the covariance matrix of the weights, as explained in better detail in the subsequent sections. Finally, the instrument noise error is estimated by perturbing the input measurements and computing the resulting change in the network output. Although the three components of the error are theoretically independent, there will be a degree of correlation among them due to the way they are estimated. This is because both the target noise and the instrument noise are estimated through the network itself and will therefore contain a component related to the network uncertainty. However, this network component is much smaller than the remaining components and will therefore have a small effect, if anything it will lead to a small overestimation of the errors.

Another important approximation introduced in this analysis is the assumption that the water vapor and liquid water follow a Gaussian distribution. This is not the case in reality. In particular, the LWP, has an approximately log-normal distribution that is truncated at LWP $= 0$. However, the same assumption of a Gaussian distribution is made when retrieving these parameters by means of optimal estimation technique and linear regression. Assuming that the variables follow a Gaussian distribution when in fact they do not will result in a bias in the retrievals.

A. Target Noise

We saw earlier that the training procedure minimizes the rms error defined in (5). When only GVR data are used, we can reasonably assume that the uncertainty in the retrieval depends on the accuracy of the brightness temperatures used as input. The methodology used here to estimate this input-dependent target noise is a variation of the approach suggested by [21] and successfully used in [25]. After the network is trained, the absolute residuals $\sigma = |\mathbf{y} - \mathbf{t}|$ are computed. A secondary network, identical to the original one, is then trained by using the elements of σ as targets. The resulting trained error network is then applied to the radiometric measurements at the moment of the retrieval to estimate an input-dependent $\sigma_t^2(\mathbf{x})$. Although this methodology may underestimate the magnitude of the target noise, it gave reasonable results in the cases analyzed here. As previously mentioned, since we are training the output target noise with the same network used for the retrieval, the target noise will not be completely independent from the network errors themselves. The upper panel of Fig. 5 shows the mean differences between real residuals and residuals estimated (over the training set) with the auxiliary network. Residuals are binned in 0.1-mm bins, and they are displayed together with
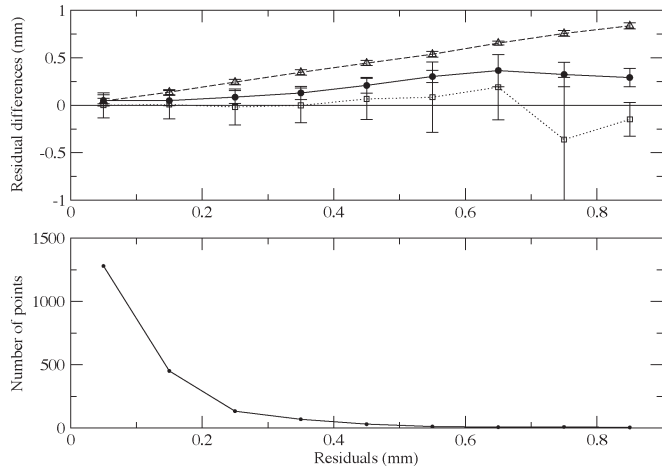
Fig. 5. (Top panel) Mean differences between real and estimated residuals for the PWV GVR-only network computed over the winter data set (circles joined by solid line). Squares joined by dotted line were obtained after correction coefficients were applied to the estimated residuals. Residuals are binned in bins of 0.1 mm, and the ± 1 standard deviation in within each bin is shown as a vertical bar. (Bottom panel) Number of points inside each bin. The triangles joined by dashed lines are obtained by setting the estimated residuals to zero.
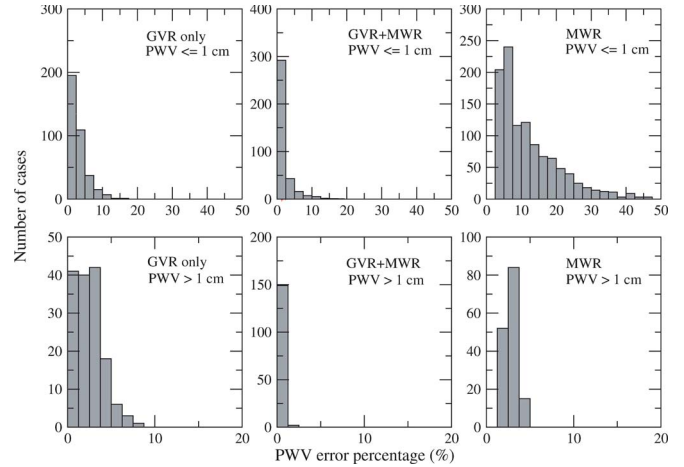

Fig. 6. (Left and central columns) Distribution of target error percentage over the testing data set (clear-sky cases) for the GVR-only and GVR + MWR networks. (Right column) MWR error distribution for an assumed rms error = 0.4 mm. The top three panels show cases with a PWV of less than 10 mm ($N = 365$). The bottom three panels show cases with a PWV greater than 10 mm ($N = 151$).

their standard deviation inside each bin. Residuals estimated by the secondary network (black points joined by solid line) are generally lower than the real residuals [26]. From a scatterplot of real and estimated residuals, a correction was determined (slope and offset) to correct for this underestimation. Corrected residuals are shown in Fig. 5 as squares joined by a dotted line. Correction coefficients were derived for all networks and for all training data sets and were used to estimate the errors when the network was applied to real data.

In the case of the combined MWR + GVR retrieval, the resultant residuals depend less on the input (as expected). This is due to the fact that the MWR sensitivity does not depend on the water-vapor amount. This can be seen in Fig. 6, which shows the percentage error distribution (target noise) over the testing data set for the GVR and GVR + MWR networks compared to the distribution of the MWR retrieval errors assumed to be equal to 0.4 mm. Residuals are shown for clear-sky cases only. Plots on the top display target noise for cases with a PWV of < 10 mm ($N = 365$), while plots on the bottom are for cases with a PWV greater than 10 mm ($N = 151$). Residuals are shown for the GVR, the MWR, and the GVR + MWR networks from left to right. When the PWV is less than 10 mm, the target noise for the GVR-only and the GVR + MWR networks is less than 5% in ∼90% of the cases. When the PWV is more than 10 mm, the target noise for the GVR + MWR and MWR networks is less than 5%. The MWR percentage error distribution is shown for an assumed constant error of 0.4 mm. The histograms in Fig. 6 show that the network is capable of weighing the varying sensitivity of the channels achieving low percentage errors when the ± 1- and ± 3-GHz channels are most sensitive (PWV < 10 mm) and using the remaining sensitivity of two-wing channels when the two most sensitive channels are saturated. The NN errors shown in Fig. 6 only refer to the target noise that is a single component of the total NN retrieval uncertainty, so the comparison with the MWR errors, which represent the total uncertainty, is not equivalent. It is shown later that the instrument noise plays a major role in

the combined GVR + MWR algorithm but it has less impact in the GVR-only retrieval.

B. Network Model Uncertainty

The second component of the retrieval error is due to model uncertainty. When the optimization procedure is complete, the network has determined the optimal weights $\mathbf{w}^*$. If we assume that the targets follow a Gaussian distribution, the network weights follow a Gaussian distribution, with mean $\mathbf{w}^*$ and covariance $\boldsymbol{C_w}$. It can be shown [21], [26], [27] that the covariance of the network weights is the inverse of the Hessian matrix

$$H = \sum_{i=1}^{N} \frac{1}{\sigma_{\mathrm{t}}^2(x_i)} \frac{\partial E_i^2}{\partial w^2}. \tag{6}$$

This gives us an opportunity to compute the effect of the network uncertainties on the network output as

$$\sigma_{\mathrm{m}}^2(\mathbf{x}) = \mathbf{g}^{\mathrm{T}}(\mathbf{x}) \mathbf{H}^{-1} \mathbf{g}(\mathbf{x}). \tag{7}$$

As previously mentioned, the fact that the distribution of the target is not actually Gaussian will introduce a bias in the estimation of this error as well. If we indicate the network output with $y(\mathbf{x}, \mathbf{w})$, then $\mathbf{g}(\mathbf{x})$ can be computed as $\mathbf{g}(\mathbf{x}) = \partial y(\mathbf{x}, \mathbf{w})/\partial \mathbf{w}$, a vector whose dimension equals the number of weights N_{w}. The Hessian matrix $\mathbf{H}$ is a square symmetric matrix of dimensions N_{w} that can be computed by finite differences or (as in this case) computed by exact computation following [28]. Before inversion, the Hessian matrix was regularized by adding a small scalar λ to the diagonal elements. The scalar was incremented in small steps until the matrix was positive definite. The criteria monitored to determine the optimal regularization were the number of negative eigenvalues, the rms difference between the diagonal elements of the original and regularized matrices, and the condition number of the regularized matrix [25]. The left panels and top right
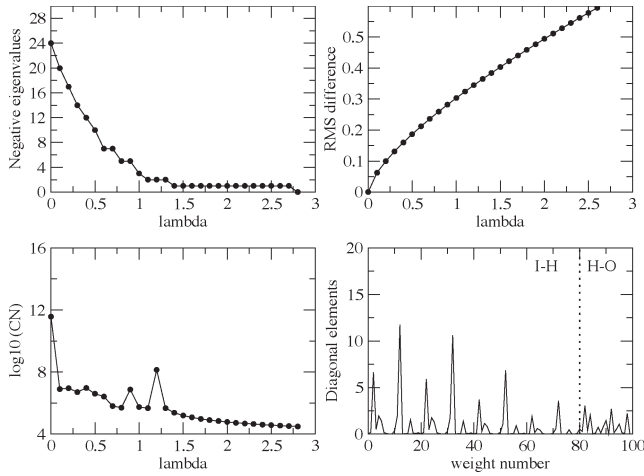
Fig. 7. Parameters used to monitor the regularization of the Hessian matrix before inversion. (Top left) Number of negative eigenvalues. (Bottom left) Conditon number. (Top right) RMS difference of matrix diagonal elements. The bottom right panel shows the diagonal elements of the Hessian matrix for the GVR-only network. The first 80 weights (to the left of the dotted line) connect input and hidden layers. The last 20 weights (to the right of the dotted line) connect hidden and output layers.
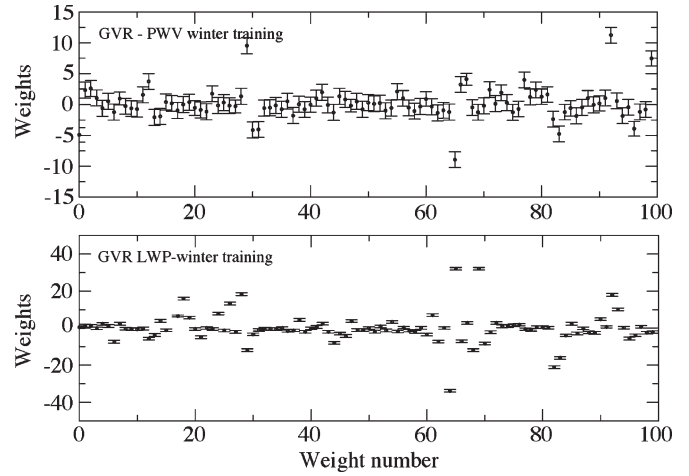

Fig. 8. Optimal weights for the GVR-only network trained with the winter data set, with the associated standard deviation computed as the square root of the Hessian diagonal elements. The first 80 weights are between the input and hidden layer. The remaining 20 weights are between the hidden and output layers. The top panel shows the PWV network weights, and the bottom panel shows the LWP network weights.

panel of Fig. 7 show, as an example, the variation of these three parameters with the increase of the small regularization scalar λ. The bottom right plot of Fig. 7 shows the diagonal elements of the Hessian matrix. The values in Fig. 7 were computed for the GVR vapor network in which the first 80 weights connect the input and hidden layers, while the remaining 20 weights connect the hidden and the output layers. Once the Hessian matrix is regularized, the covariance matrix of the network weights can be computed. The diagonal elements of this matrix can give us an idea of the uncertainty associated with the network weights. One example is shown in Fig. 8, which shows the network weights and associated error bars (± 1 standard deviation) computed as the square root of the diagonal elements of the covariance matrix. The uncertainty on the network weight appears to be quite low. This source of error is not the dominant contributor to the overall retrieval error. Once the network is applied to real measurements, the effect of network uncertainty on individual output points can be evaluated by using (7).

C. Instrument Noise

The final source of uncertainty that we take into account is due to the instrument noise. We apply a perturbation $\partial x = \pm 1$ K to each GVR measurement and $\partial x = \pm 0.5$ K to each MWR measurement used in the retrieval

$$\frac{\partial y}{\partial x_i} = \frac{1}{2} \left(\left| \frac{\partial y}{\partial x_{i+}} \right| + \left| \frac{\partial y}{\partial x_{i-}} \right| \right) \tag{8}$$

where the index $i(= 1, \ldots, N_c)$ indicates the radiometric channel used in the retrieval and ∂x indicates a perturbation around the measured value. Assuming that the radiometric noise is independent for the four channels, the total radiometric noise contribution is

$$\sigma_N^2 = \sum_{i=1}^{N_c} \sigma_i^2 \left(\frac{\partial y}{\partial x_i} \right)^2 \tag{9}$$

where σ_i^2 is the channel variance. It was mentioned in Section II that measurements from this instrument were affected by the presence of RFI. The interference will introduce a certain degree of noise correlation among the channels. Since the RF environment at the radiometer location changes continuously because new instruments are often added (and some are removed), it is desirable not to model the interference as a permanent feature of the data set and the retrieval. However, the RFI effects were estimated (although not included in the subsequent analysis) on the subset of data analyzed in this paper, and a nonnegligible covariance was found between the ± 1- and ± 14-GHz channels. The effect of this covariance [when added to (9)] is to increase the instrument noise component by about 10%.

VI. PWV RETRIEVAL RESULTS

The NN coefficients were applied to a total of four months of GVR and MWR measurements in 2007, two months in winter (January, February) and two months in summer (July, August). GVR brightness temperatures were smoothed with a 5-min running median filter to eliminate possible artifacts due to RFI. Retrievals from the GVR network that was trained with the all-year data set (not shown) have little advantage with respect to the MWR statistical retrieval. The NN retrievals noticeably deteriorate when the PWV amount increases.

The real advantages of using the GVR to retrieve PWV can be better seen when we train the NN by using a seasonal training set. In this way, the high sensitivity of the water-vapor channels can be used to focus on the retrieval of low amounts of PWV. A scatter-plot of radiosonde PWV versus retrieved PWV is shown in Fig. 9. In the left panel, all cases are shown, while in the right panel, only cases with a PWV of less than 5 mm are shown. When the PWV is less than 5 mm, the NN output has less scatter than the MWRRET output. Although the physical retrieval is probably the best retrieval technique for small amounts of PWV, the NN retrieval has comparably good results. The slightly superior performance of the physical retrieval is
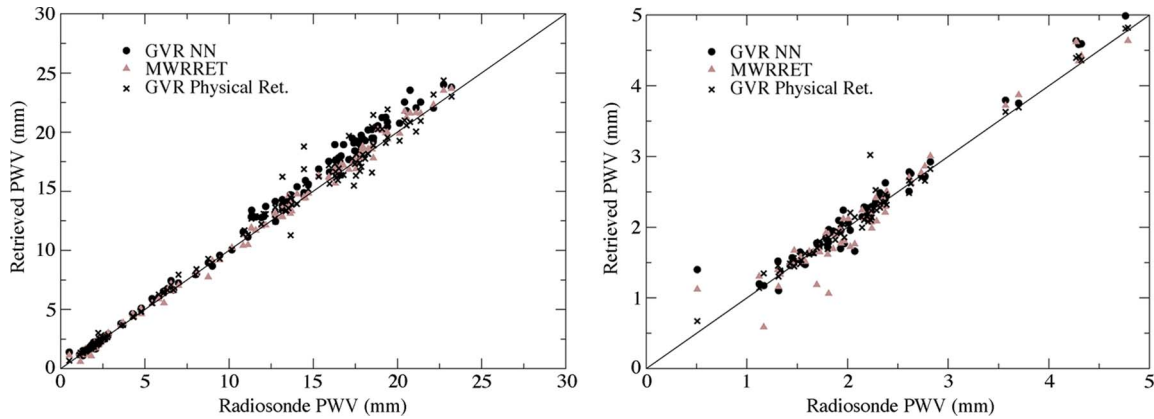
Fig. 9. Scatter plot of PWV measured by radiosonde (X-axis) and retrieved by using (black circles) the GVR-only NN seasonal retrieval, (crosses) physical retrieval, and (brown triangles) MWRRET linear retrieval. (Left panel) All cases ($N = 145$). (Right panel) Cases with PWV of $<= 5$ mm ($N = 62$).

TABLE II
REGRESSION PARAMETERS BETWEEN RADIOSONDE PWV AND PWV FROM NN, PHYSICAL
RETRIEVAL, AND LINEAR RETRIEVAL. ALL UNITS ARE IN MILLIMETERS

Retrieval	ALL DATA (N = 145) RS mean = 9.39 RS st. dev = 7.10				PWV <=5 mm (N = 62) RS mean = 2.21 RS st. dev = 0.8				PWV >10 mm (N = 70) RS mean = 16.18 RS st. dev = 3.21			
	GVR-Seasonal (NN)	Physical	GVR+MWR	MWRRET	GVR-Seasonal (NN)	Physical	GVR+MWR	MWRRET	GVR-Seasonal (NN)	Physical	GVR+MWR	MWRRET
Mean	9.96	9.64	9.65	9.51	2.27	2.23	2.34	2.19	17.26	16.60	16.56	16.43
Standard dev.	7.62	7.29	7.26	7.26	0.91	0.87	0.88	0.93	3.46	3.26	3.39	3.37
Correlation coefficient	0.998	0.994	0.999	0.998	0.984	0.988	0.987	0.976	0.983	0.936	0.992	0.992
(Slope)	1.070±0.005	1.019±0.009	1.021±0.004	1.021±0.004	1.03±0.02	0.997±0.019	0.998±0.02	1.04±0.03	1.05±0.02	0.95±0.04	1.05±0.02	1.04±0.02
(Intercept)	−0.09±0.06	0.05±0.11	0.06±0.05	−0.09±0.05	−0.02±0.057	0.03±0.05	0.14±0.05	−0.11±0.07	0.13±0.4	1.22±0.7	−0.43±0.3	−0.4±0.2
Sonde-ret mean	−0.57	−0.23	−0.25	−0.10	−0.07	−0.03	−0.13	0.02	−1.08	−0.42	−0.38	−0.25
Sonde-ret st dev	0.68	0.83	0.37	0.38	0.17	0.13	0.14	0.21	0.66	1.16	0.47	0.45
RMS	0.89	0.86	0.45	0.39	0.18	0.13	0.19	0.21	1.27	1.23	0.60	0.51

achieved by utilizing the radiosonde temperature profiles as input. Interestingly, the physical retrieval has a larger variability when the PWV is higher. The reason for this variability is not known at this point. Note that when we estimate the rms error with respect to radiosonde, we are including uncertainties in the radiosonde estimation as well. These uncertainties are more pronounced during very dry conditions. The numerical results for the various networks are shown in Table II. Results presented in this table were obtained by computing the linear regression coefficients between the retrieved PWV and PWV measured by the radiosondes Vaisala RS92 launched from the same location where the radiometer is situated. Although the accuracy of Vaisala RS92 radiosondes is generally considered better than the accuracy of RS80, recent studies indicates the possible presence of a dry bias in the radiosonde measurements that is dependent on the solar zenith angle [29], [30]. This study is also supported by Rowe *et al.* [31], which found significant diurnal biases in Vaisala radiosondes launched from Antarctica. Table II clearly shows that MWRRET retrievals perform very well across the whole range of PWV. While the rms error of the MWRRET retrieval does not sensibly change across the PWV range, the GVR-only retrieval varies from 0.18 mm to more than 1 mm as the PWV increases. The combined GVR + MWR retrieval, on the other hand, has an rms error comparable to that of the MWR-only retrieval.
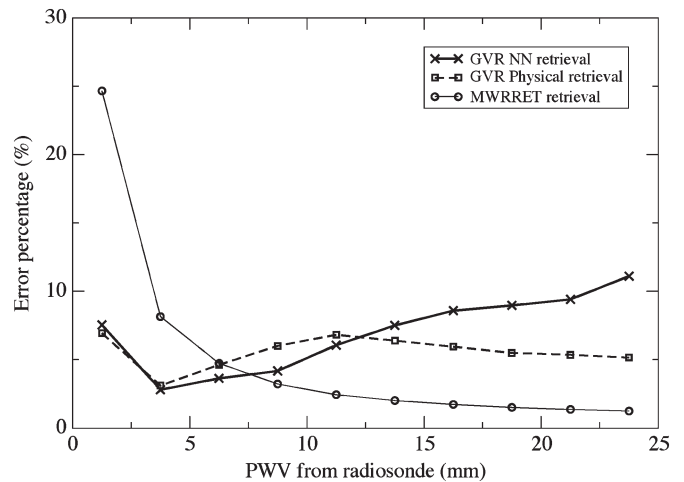

Fig. 10. Total PWV retrieval error percentage for (crosses) the NN seasonal retrieval, (squares) physical retrieval, and (solid line) MWRRET linear retrieval. Errors are averaged over 2.5-mm bins.

The total percentage errors of the physical and NN retrievals are shown in Fig. 10 as a function of radiosonde PWV. From this figure, it is evident that the physical retrieval errors are comparable to those of the NN and that both errors increase when the PWV increases, to become larger than the MWR
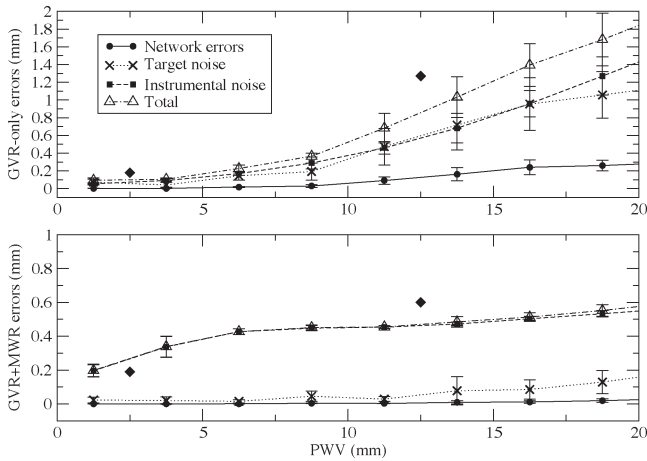
Fig. 11.   Contribution to the total PWV retrieval error (indicated by triangles) due to the three error components (target noise, network error, and instrument noise). (Top panel) GVR-only network. (Bottom panel) GVR + MWR network. The disconnected black diamonds (centered at 2.5 and 12.5) are the rms values from Table II for PWV < 5 mm and PWV > 10 mm.
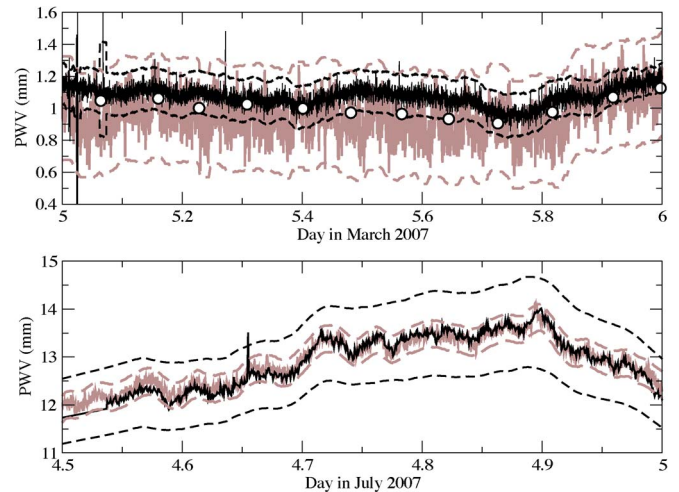

Fig. 12.   Time series of PWV as retrieved by (black lines) the GVR seasonal NN, and by (brown lines) the MWRRET linear retrieval and as measured by (circles) radiosondes. The dashed lines of corresponding colors represent ±1 standard deviation of (smoothed) the retrieval. (Top panel) One day of data collected on March 5, 2007. (Bottom panel) Data collected during part of July 4, 2007.

linear retrieval errors when the PWV amount is higher than ~10 mm. Both the physical retrieval and NN errors are in the 5% range for PWV amounts between 2 and 7 mm. Single contributions to the total retrieval errors from the three components are shown in Fig. 11. Here, retrieval errors averaged over bins of PWV are displayed together with their standard deviation. This figure shows that in the GVR-only network, the largest contributors to the overall error are the target noise and the instrument errors that rapidly increases above 10 mm. It is also clear from Fig. 11 that, since for the GVR-only network the instrument noise increases as the PWV increases, a 1-K perturbation in the input brightness temperature will cause a larger error in the retrievals when the channels are close to saturation. In the MWR + GVR network, the instrument noise is the largest contributor to the overall error that remains almost constant over the whole PWV range. For comparison, the rms errors from Table II are shown in Fig. 11 as black diamonds. The agreement between the theoretical and observed errors is acceptable. Discrepancies between the modeled and observed errors in the GVR-only network could derive from the RFI effects not accounted for in the model (as we discussed in Section V-C the inclusion of RFI effects in the computation of the instrument noise component will increase the modeled error of about 10%). In the GVR + MWR network, on the other hand, it seems that the instrumental noise for the MWR may be slightly overestimated.

To conclude the PWV comparison, Fig. 12 shows a time series of PWV retrieved with the GVR seasonal algorithm during March (top panel) and July (bottom panel). The NN retrievals are in black, while the MWR retrievals are in brown. The smooth lines are the corresponding error boundaries (±1 standard deviation), and the circles are radiosonde measurements. Only a short portion of the data is shown for clarity. During the month of March 2007, the Radiative Heating in Underexplored Bands Campaign (RHUBC) intensive operation period (IOP) was under way at the NSA. As part of this IOP, a large number of radiosonde soundings were launched daily. The top panel of Fig. 12 shows data collected on March 5. It can be seen that although the MWR retrievals perform remarkably

well in such dry conditions, the noise in the MWR retrievals is much larger than it is in the GVR retrievals. The bottom panel of Fig. 12 shows a portion of data collected during July 4, with a PWV above 10 mm. In this case, we see that uncertainty in the GVR-only retrieval is higher than it is in the MWR retrieval.

VII. LWP RETRIEVAL RESULTS

Accurate retrievals of LWP in the Arctic environment represent a challenging task because the liquid amount is very low in the majority of cases. As previously mentioned, although the GVR was specifically designed for the retrieval of small amounts of water vapor, the sensitivity to liquid of the two-wing channels is about three times the sensitivity of MWR's channels [8], [11]. Errors for the MWR statistical retrieval are estimated to be around 15 g/m^2; however, an LWP in this range is not infrequent in the Arctic. The statistical distribution of one year of LWP retrievals from the MWR is shown in Fig. 2 (right panel). The figure shows that about 80% of cases have an LWP amount of less than 100 g/m^2, resulting in a retrieval uncertainty of at least 30% or more in the majority of cases. The high uncertainty of MWR retrievals is evident when we examine the retrievals during clear-sky condition, as is shown later.

The LWP network was trained with the same seasonal training sets used in the vapor retrieval. Fig. 13 shows a time series of retrieved LWP similar to the cases shown in Fig. 12. The top panel shows LWP retrieved on March 5, 2007 (a complete clear-sky day). The bottom panel shows a low-liquid cloud case on February 15, 2007. One of the first conclusions that can be drawn from Fig. 13 is that GVR retrievals return less liquid than do the MWR retrievals.

Clear-sky liquid retrievals for the MWR and GVR were analyzed by selecting cases where the readings from the colocated IRT were less than 215 K. Distributions of clear-sky retrievals from the two instruments are shown in Fig. 14. During clear-sky conditions, the GVR retrieves an LWP amount of less than
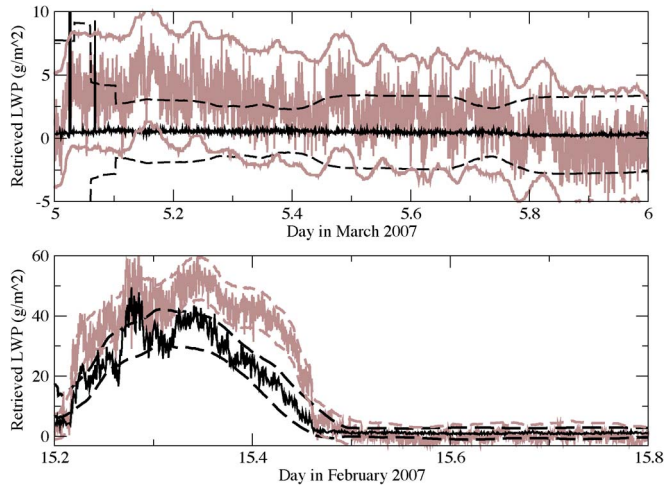
Fig. 13. Time series of LWP as retrieved by (black lines) the GVR seasonal NN and by (brown lines) the MWRRET linear retrieval. The dashed lines of corresponding colors represent ± 1 standard deviation of (smoothed) the retrieval. (Top panel) One clear-sky day of data collected on March 5, 2007. (Bottom panel) Data collected during part of February 15, 2007.
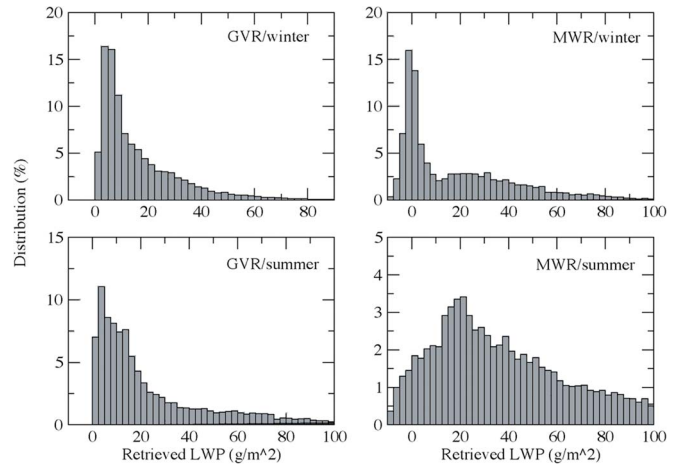

Fig. 15. Distribution of retrieved LWP during cloudy (IRT of > 230 K) conditions for the month of January (winter, $N = 16\,112$) and July (summer, $N = 21\,063$). The left panel shows retrievals from the GVR, and the right panel shows retrievals from the MWR.

Fig. 16. Downwelling longwave surface flux residuals (measurements minus model) for cloudy cases obtained by using (black solid circles) the GVR-retrieved LWP and (open squares) the MWR-retrieved LWP as the model parameter. Cases are classified as cloudy if the IRT temperature is > 230 K.
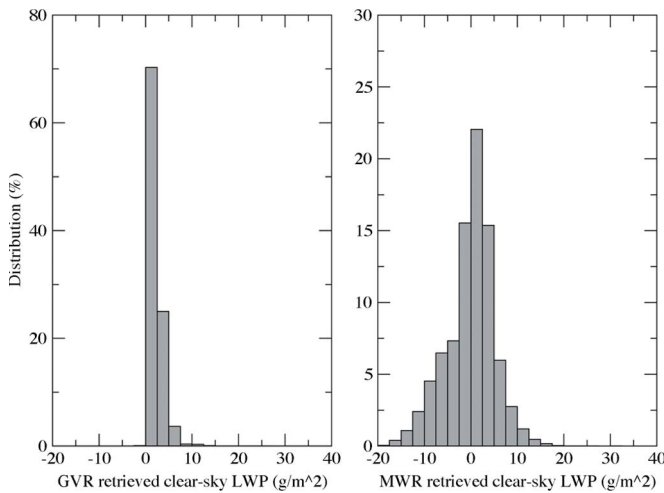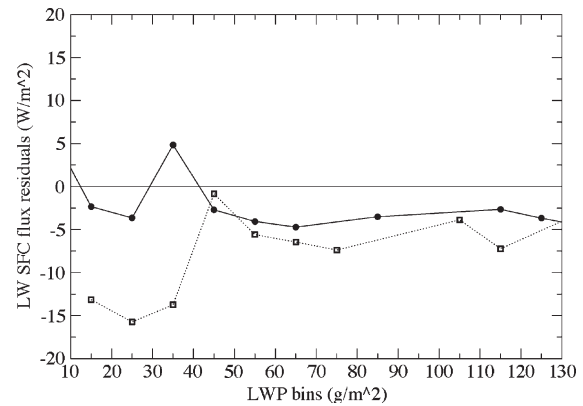

Fig. 14. Distribution of retrieved LWP during clear-sky (IRT of < 215 K) conditions for the month of January and July combined. The left panel shows retrievals from the GVR. The right panel shows retrievals from the MWR ($N = 32\,444$).

2.5 g/m^2 and less than 5 g/m^2 in approximately 70% and 95% of cases, respectively, while the MWR has a broader distribution centered on 0 g/m^2 but a high number ($\sim$45%) of negative cases. When the IR temperature is higher than 215 K, clouds are likely present in the IRT field of view. A cloud base temperature threshold of 230 K was set as the minimum temperature to classify the scene as cloudy (with some form of liquid water), although the cloud could be completely made of ice. The distribution of retrieved LWP for January and July 2007 for cases classified as cloudy is shown in Fig. 15. In January, $\sim$2.5% (and in July $\sim$12%) of GVR retrievals have an LWP amount of less than 2.5 g/m^2 in the region of retrieval uncertainty. As previously mentioned, the MWR retrievals might be biased toward larger values because of the modeling of the oxygen line. Previous studies showed that LWP retrievals from the MWR are generally higher than retrievals obtained from IR and aircraft measurements [32]. The right side of Fig. 15 indicates

that the percentage of negative LWP values is higher in winter, when the LWP amounts are usually lower.

As an independent test for assessing the retrievals, we used the retrieved LWP as a parameter in the radiative transfer model RRTM that computes longwave downwelling surface fluxes [33]. The cloud liquid water has a large impact on the calculations of the longwave downwelling flux, particularly in the region of low LWP [3], and we expect that a better LWP retrieval will result in a better agreement between measured and modeled surface fluxes. We performed the computations for 135 cases coincident with the times of radiosonde ascent. Radiosonde temperature and humidity profiles were used as input to the radiative transfer model, together with the estimated LWP. Clouds were assumed to be entirely liquid, with a liquid droplet effective radius of 7.0 μm and cloud fraction of one. A cloud layer was assumed to be present when the relative humidity was higher than 95% and the liquid water profile was assumed to be constant. The retrieved LWP were binned in 10 g/m^2 intervals. The mean of the longwave radiative flux residuals (measurements minus model) in each bin are shown Fig. 16 for cloudy cases. The corresponding numerical values are displayed in Table III. The number of cases

TABLE III
COMPARISON OF MEASURED AND MODELED DOWNWELLING LONGWAVE SURFACE FLUXES

LWP bin (g/m²)	All data						Cloudy data					
	N		Meas–model mean (W/m²)		Meas–model st. dev. (W/m²)		N		Meas–model mean (W/m²)		Meas–model st. dev. (W/m²)	
	GVR	MWR	GVR	MWR	GVR	MWR	GVR	MWR	GVR	MWR	GVR	MWR
0–10	116	45	−11.12	−10.41	20.40	25.68	24	8	6.6	26.86	18.93	14.16
10–20	19	6	−7.91	−18.35	18.62	12.92	17	5	-2.34	−13.15	9.61	6.17
20–30	5	13	−3.64	−23.10	6.31	19.91	5	10	−3.64	−15.74	6.31	12.16
30–40	4	8	4.83	−13.74	10.47	15.08	4	8	4.83	−13.75	10.47	15.08
40–50	4	4	−2.71	−2.73	1.64	3.75	4	3	−2.70	−0.84	1.64	2.09
50–60	5	7	−3.93	−5.59	1.44	4.57	4	7	−4.06	−5.59	1.58	4.57
60–70	1	2	−4.71	−6.46	-	4.53	1	2	−4.72	−6.45	-	4.52
70–80	1	2	−3.51	−6.97	-	0.41	1	1	−3.51	−7.37	-	-
80–90	1	1	−2.65	−3.74	-	-	1	0	−2.65	-	-	-
90–100	1	0	−3.67	-	-	-	1	0	−3.67	-	-	-

in each LWP bin (N), the mean, and the standard deviation of the differences between measurements and model are displayed for the two retrievals. MWR retrievals are limited to cases with an LWP > 0. In spite of the few number of cases, the behavior of the residuals is consistent with the statistics of LWP shown in Figs. 14 and 16. The GVR-based LWP retrievals clearly give smaller biases and standard deviations in the radiative flux comparison than do the MWR retrievals.

When all cases are considered (not shown in the figure), residuals involving the GVR-retrieved LWP are much smaller than residuals computed with the MWR-retrieved LWP. Surface fluxes obtained with MWR retrievals are constantly higher than the observed ones, suggesting that the MWR overall overestimates LWP. This behavior of MWR residuals is consistent with what is found in [5]. Residuals obtained with GVR retrievals are generally in the uncertainty of the longwave radiative flux measurements themselves (± 4 W/m²). If we use the IRT readings to split the data into cloudy cases (Fig. 16) and clear-sky cases, we see that a large portion of the overestimation in the MWR retrievals is due to clear-sky LWP bias. This bias is caused partially by the weakness of the signal and partially by the modeling of the wing of the oxygen absorption region. The distribution of residuals during the clear-sky and cloudy scenes is very consistent with the distribution of LWP shown in Figs. 15 and 16, with the GVR uncertainty increasing when the LWP is below 5 g/m² and with the MWR retrieving too much over the LWP range of 0–60 g/m². The overall GVR residuals are better than those from the MWR, indicating a more realistic representation of the LWP distribution.

VIII. SUMMARY AND CONCLUSION

Retrievals of small amounts of water vapor and liquid water have become an essential need to assess the effect of clouds and water vapor on the Earth radiation budget. The Arctic environment in particular presents challenging conditions in which traditional 20- to 30-GHz MWRs operate at the limit of

their sensitivity. The use of the stronger 183.3-GHz absorption line poses a challenge in itself, because the brightness temperatures are nonlinearly related to the water-vapor amounts. This paper develops an NN for the retrieval of low amounts of water vapor and liquid water in the Arctic from a ground-based 183.3-GHz MWR. A previous physical algorithm developed for this instrument required vertical profiles of temperature and humidity as input. By using an NN, it is possible to provide real-time retrievals and, at the same time, exploit the nonlinear region of response of the instrument. The network was trained first with an all-year data set and then with a seasonal data set. A seasonal training data set was deemed preferable since it enables the network to focus on cases with smaller PWV amounts, therefore taking better advantage of the higher sensitivity of the GVR channels. Particular attention was devoted to the realistic treatment of the NN output errors. Three independent sources of errors were identified and estimated, and their individual contribution to the overall retrieval error was analyzed. The estimation errors were found to be of the same magnitude as those obtained from a traditional physical retrieval, confirming the reliability of the NN outputs.

Outputs from the NN retrieval were compared with those from a physical retrieval that uses GVR-only measurements and those from a linear statistical retrieval (MWRRET) that uses only MWR measurements. Results show that a seasonally trained NN can retrieve PWV with a 5% uncertainty between about 2 and 7 mm and with about 10% uncertainty when the PWV is more than 10 mm. The rms error computed with reference to radiosonde measurements for which PWV < 5 mm showed a very low bias (-0.08 mm) and a standard deviation of about 0.18 mm. A combination of GVR and MWR measurements resulted in an improved rms error across the whole PWV range.

LWP retrievals display a markedly improved clear-sky detection capability and a retrieval error between 1 and 10 g/m², depending on the water-vapor amount. LWP retrieved with the NN is generally lower than the LWP retrieved from

MWR measurements. Radiative transfer computations showed that longwave downwelling surface fluxes, computed by using GVR-retrieved LWP, are closer to the measured fluxes than those computed by using MWR-retrieved LWP, indicating that a lower LWP amount is probably more realistic.

Overall, the NN results can be seen as a step forward in providing improved real-time retrievals of water vapor and cloud liquid water. These retrievals can be used as a starting point or as a comparison point for more sophisticated retrievals combining multiple active and passive instruments. In the near future, a radiometer operating at 90/150-GHz will be deployed at the ACRF to improve the retrieval of small LWP amounts. Once this additional instrument is deployed, we plan to extend the NN architecture to include measurements at 90 and 150 GHz and to explore the possibility of retrieving cloud liquid content from microwave measurements between 30 and 176 GHz.

ACKNOWLEDGMENT

REFERENCES

[1] M. D. Shupe and J. M. Intrieri, "Cloud radiative forcing of the Arctic surface: The influence of cloud properties, surface albedo, and solar zenith angle," *J. Clim.*, vol. 17, no. 3, pp. 616–628, Feb. 2004.

[2] J. Verlinde, J. Y. Harrington, G. M. McFarquhar, V. T. Yannuzzi, A. Avramov, S. Greenberg, N. Johnson, G. Zhang, M. R. Poellot, J. H. Mather, D. D. Turner, E. W. Eloranta, B. D. Zak, A. J. Prenni, J. S. Daniel, G. L. Kok, D. C. Tobin, R. Holz, K. Sassen, D. Spangenberg, P. Minnis, T. P. Tooman, M. D. Ivey, S. J. Richardson, C. P. Bahrmann, M. Shupe, P. J. Demott, A. J. Heymsfield, and R. Schofield, "The mixed-phase Arctic cloud experiment," *Bull. Amer. Meteorol. Soc.*, vol. 88, no. 2, pp. 205–221, Feb. 2007.

[3] D. D. Turner, A. M. Vogelmann, R. T. Austin, J. C. Barnard, K. Cady-Pereira, J. C. Chiu, S. A. Clough, C. Flynn, M. M. Khaiyer, J. Liljegren, K. Johnson, B. Lin, C. Long, A. Marshak, S. Y. Matrosov, S. A. McFarlane, M. Miller, Q. Min, P. Minnis, W. O'Hirok, Z. Wang, and W. Wiscombe, "Thin liquid water clouds: Their importance and our challenge," *Bull. Amer. Meteorol. Soc.*, vol. 88, no. 22, pp. 177–190, Feb. 2007.

[4] M. P. Cadeddu, V. H. Payne, S. A. Clough, K. Cady-Pereira, and J. C. Liljegren, "Effect of the oxygen line-parameter modeling on temperature and humidity retrievals from ground-based microwave radiometers," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 7, pt. 2, pp. 2216–2223, Jul. 2007.

[5] D. D. Turner, "Improved ground-based liquid water path retrievals using a combined infrared and microwave approach," *J. Geophys. Res.—Atmos.*, vol. 112, no. D15, p. D15 204, Aug. 2007.

[6] D. D. Turner, "Arctic mixed-phase cloud properties from AERI Lidar observations: Algorithm and results from SHEBA," *J. Appl. Meteorol.*, vol. 44, no. 4, pp. 427–444, Apr. 2005.

[7] A. L. Pazmany, "A compact 183-GHz radiometer for water vapor and liquid water sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 7, pp. 2202–2206, Jul. 2007.

[8] M. P. Cadeddu, J. C. Liljegren, and A. L. Pazmany, "Measurements and retrievals from a new 183-GHz water-vapor radiometer in the arctic," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 7, pp. 2207–2215, Jul. 2007.

[9] P. E. Racette, 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, "Measurements of low amounts of precipitable water vapor using ground-based millimeterwave radiometry," *J. Atmos. Ocean. Technol.*, vol. 22, no. 4, pp. 317–333, Apr. 2005.

[10] V. H. Payne, J. S. Delamere, K. E. Cady-Pereira, R. R. Gamache, J.-L. Moncet, E. J. Mlawer, and S. A. Clough, "Air-broadened half-widths of the 22- GHz and 183-GHz water-vapor lines," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3601–3617, Nov. 2008.

[11] D. Cimini, A. J. Gasiewski, M. Klein, V. Y. Leusky, and J. C. Liljegren, "Ground-based millimeter- and submillimeter-wave observations of low vapor and liquid water contents," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 7, pt. 2, pp. 2169–2180, Jul. 2007.

[12] J. C. Liljegren, "Automatic self-calibration of ARM microwave radiometers," in *Microwave Radiometry and Remote Sensing of the Earth's Surface and Atmosphere*, P. Pampaloni and S. Paloscia, Eds. Lorton, VA: VSP Book, 2000, pp. 433–443.

[13] F. C. Aires, C. Prigent, W. B. Rossow, and M. Rothstein, "A new neural network approach including first guess for retrieval of atmospheric water vapor, cloud liquid water path, surface temperature, and emissivities over land from satellite microwave observations," *J. Geophys. Res.—Atmos.*, vol. 106, no. D14, pp. 14 887–14 907, Jul. 2001.

[14] W. J. Blackwell, "A neural-network technique for the retrieval of atmospheric temperature and moisture profiles from high spectral resolution sounding data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 11, pp. 2535–2546, Nov. 2005.

[15] C. D. Rodgers, "Inverse methods for atmospheric sounding," in *Theory and Practice*. Hackensack, NJ: World Scientific, 2000.

[16] F. Del Frate and G. Schiavon, "Neural networks for the retrieval of water vapor and liquid water from radiometric data," *Radio Sci.*, vol. 33, no. 5, pp. 1373–1386, Sep./Oct. 1998.

[17] D. D. Turner, S. A. Clough, J. C. Liljegren, E. E. Clothiaux, K. E. Cady-Pereira, and K. L. Gaustad, "Retrieving liquid water path and precipitable water vapor from the Atmospheric Radiation Measurement (ARM) microwave radiometers," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pt. 2, pp. 3680–3690, Nov. 2007.

[18] P. Bauer, P. Amayenc, C. D. Kummerow, and E. A. Smith, "Over-ocean rainfall retrieval from multisensor data of the Tropical Rainfall Measuring Mission. Part II: Algorithm implementation," *J. Atmos. Ocean. Technol.*, vol. 18, no. 11, pp. 1838–1855, Nov. 2001.

[19] M. F. Moller, "A scaled conjugate-gradient algorithm for fast supervised learning," *Neural Netw.*, vol. 6, no. 4, pp. 525–533, 1993.

[20] S. A. Clough, M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara, and P. D. Brown, "Atmospheric radiative transfer modeling: A summary of the AER codes," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 91, no. 2, pp. 233–244, Mar. 2005.

[21] G. Papadopoulos, P. J. Edwards, and A. F. Murray, "Confidence estimation methods for neural networks: A practical comparison," *IEEE Trans. Neural Netw.*, vol. 12, no. 6, pp. 1278–1287, Nov. 2001.

[22] L. S. Rothman, D. Jacquemart, A. Barbe, D. Chris Benner, M. Birk, L. R. Brown, M. R. Carleer, C. Chackerian, Jr., K. Chance, L. H. Coudert, V. Dana, V. M. Devi, J.-M. Flaud, R. R. Gamache, A. Goldman, J.-M. Hartmann, K. W. Jucks, A. G. Maki, J.-Y. Mandin, S. T. Massie, J. Orphal, A. Perrin, C. P. Rinsland, M. A. H. Smith, J. Tennyson, R. N. Tolchenov, R. A. Toth, J. Vander Auwera, P. Varanasi, and G. Wagner, "The HITRAN 2004 molecular spectroscopic database," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 96, no. 2, pp. 139–204, Dec. 2005.

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

[24] H. J. Liebe, G. A. Hufford, and M. G. Cotton, "Propagation modeling of moist air and suspended water/ice particle at frequencies below 1000 GHz," in *Proc. AGARD Meeting Atmos. Propag. Effects Through Natural Man-Made Obscurants Visible MM-Wave Radiation*, Palma de Mallorca, Spain, May 1993.

[25] M. D. Müller, A. K. Kaifel, M. Weber, S. Tellmann, J. P. Burrows, and D. Loyola, "Ozone profile retrieval from Global Ozone Monitoring Experiment (GOME) data using a neural network approach (Neural Network Ozone Retrieval System (NNORSY))," *J. Geophys. Res.*, vol. 108, no. D16, p. 4497, Aug. 2003. DOI: 10.1029/2002JD002784.

[26] R. Tibshirani, "A comparison of some error estimates for neural network models," *Neural Comput.*, vol. 8, no. 1, pp. 152–163, Jan. 1996.

[27] F. Aires, "Neural Network uncertainty assessment using Bayesian statistics with applications to remote sensing: 1. Network Weights," *J. Geophys. Res.*, vol. 109, no. D10, p. D10 303, May 2004. DOI: 10.1029/2003JD004173.

[28] C. M. Bishop, "Exact calculation of the Hessian matrix for the multi-layer perceptron," *Neural Comput.*, vol. 4, no. 4, pp. 494–501, Jul. 1992.

[29] J. C. Liljegren and M. P. Cadeddu, "Application of microwave radiometry to improve climate records," presented at the 87th AMS Meeting,

St. Antonio, TX, Jan. 14–18, 2007. [Online]. Available: http://ams.confex.com/ams/pdfpapers/120430.pdf

[30] K. E. Cady-Pereira, M. W. Shephard, D. D. Turner, E. J. Mlawer, S. A. Clough, and T. J. Wagner, "Improved daytime column-integrated precipitable water vapor from Vaisala radiosonde humidity sensors," *J. Atmos. Ocean. Technol.*, vol. 25, no. 6, pp. 873–883, 2008.

[31] P. M. Rowe, L. M. Miloshevich, D. D. Turner, and V. P. Walden, "Dry bias in Vaisala RS90 radiosonde humidity profiles over Antarctica," *J. Atmos. Ocean. Technol.*, vol. 25, no. 9, pp. 1529–1541, Sep. 2008.

[32] B. Lin, P. Minnis, A. Fan, J. A. Curry, and H. Gerber, "Comparison of cloud liquid water paths derived from in situ and microwave radiometer data taken during the SHEBA/FIREACE," *Geophys. Res. Lett.*, vol. 28, no. 6, pp. 975–978, 2001.

[33] E. J. Mlawer, S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, "Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave," *J. Geophys. Res.—Atmos.*, vol. 102, no. D14, pp. 16663–16682, Jul. 1997.

Maria P. Cadeddu received the Laurea degree in physics from the University of Cagliari, Cagliari, Italy, in 1994 and the Ph.D. degree in physics from Heriot-Watt University, Edinburgh, U.K., in 2002.

Since 2005, she has been the Instrument Mentor for the U.S. Department of Energy Atmospheric Radiation Measurement Program Climate Research Facility microwave instrumentation. She is currently with Argonne National Laboratory, Argonne, IL. Her research interests include microwave instrumentations and application of remote sensing to atmospheric research.

Dr. Cadeddu is a member of the American Geophysical Union.

David D. Turner received the B.A. and M.S. degrees in mathematics from Eastern Washington University, Cheney, in 1992 and 1994, respectively, and the Ph.D. degree in atmospheric science from the University of Wisconsin—Madison, Madison, in 2003.

He is currently a Researcher with the Space Science and Engineering Center, University of Wisconsin. He is actively involved in the Department of Energy's Atmospheric Radiation Measurement (ARM) program, and is currently the Chair of the ARM Radiative Processes working group and is a member of the ARM Science and Infrastructure Steering Committee. His current research interests include infrared and microwave remote sensing, longwave radiative transfer, and retrieving water vapor and aerosol properties from active and passive remote sensors.

James C. Liljegren received the B.S., M.S., and Ph.D. degrees in mechanical engineering from the University of Illinois at Urbana–Champaign, Urbana, in 1980, 1983, and 1989, respectively. His dissertation focused on stochastic modeling of turbulent dispersion in the convective boundary layer.

In 1990, he was with the Pacific Northwest National Laboratory, where he became involved in ground-based microwave radiometry with the U.S. Department of Energy Atmospheric Radiation Measurement (ARM) program. He was with the Ames Laboratory, Ames, IA, in 1997, where he continued to pursue research related to microwave radiometry and retrieval algorithm development. Since 1999, he has been with Argonne National Laboratory, Argonne, IL. In 2000, he became the Site Program Manager for the ARM Southern Great Plains Climate Research Facilities in Oklahoma and Kansas. From 2003 to 2008, he has also served as the ARM Instrument Team Coordinator.

Dr. Liljegren is a member of the American Meteorological Society and the American Geophysical Union.