Revealing the Winds under the Rain. Part I: Passive Microwave Rain Retrievals Using a New Observation-Based Parameterization of Subsatellite Rain Variability and Intensity—Algorithm Description

Citation	Hristova-Veleva, S. M., Callahan, P. S., Dunbar, R. S., Stiles, B. W., Yueh, S. H., Huddleston, J. N., ... & Weissman, D. E. (2013). Revealing the Winds under the Rain. Part I: Passive Microwave Rain Retrievals Using a New Observation-Based Parameterization of Subsatellite Rain Variability and Intensity—Algorithm Description. Journal of Applied Meteorology and Climatology, 52(12), 2828-2848. doi:10.1175/JAMC-D-12-0237.1
DOI	10.1175/JAMC-D-12-0237.1
Publisher	American Meteorological Society
Version	Version of Record
Terms of Use	http://cdss.library.oregonstate.edu/sa-termsofuse

Revealing the Winds under the Rain. Part I: Passive Microwave Rain Retrievals Using a New Observation-Based Parameterization of Subsatellite Rain Variability and Intensity—Algorithm Description

S. M. HRISTOVA-VELEVA, P. S. CALLAHAN, R. S. DUNBAR, B. W. STILES, S. H. YUEH, J. N. HUDDLESTON, S. V. HSIAO, AND G. NEUMANN

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

B. A. VANHOFF

College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, Oregon

R. W. GASTON AND E. RODRIGUEZ

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

D. E. WEISSMAN

Hofstra University, Hempstead, New York

(Manuscript received 7 July 2012, in final form 16 July 2013)

ABSTRACT

Scatterometer ocean surface winds have been providing very valuable information to researchers and operational weather forecasters for over 10 years. However, the scatterometer wind retrievals are compromised when rain is present. Merely flagging all rain-affected areas removes the most dynamic and interesting areas from the wind analysis. Fortunately, the *Advanced Earth Observing Satellite II* (*ADEOS-II*) mission carried a radiometer [the Advanced Microwave Scanning Radiometer (AMSR)] and a scatterometer, allowing for independent, collocated retrievals of rain. The authors developed an algorithm that uses AMSR observations to estimate the rain inside the scatterometer beam. This is the first in a series of papers that describe their approach to providing rain estimation and correction to scatterometer observations. This paper describes the retrieval algorithm and evaluates it using simulated data. Part II will present its validation when applied to AMSR observations. This passive microwave rain retrieval algorithm addresses the issues of nonuniform beam filling and hydrometeor uncertainty in a novel way by 1) using a large number of soundings to develop the retrieval database, thus accounting for the geographically varying atmospheric parameters; 2) addressing the spatial inhomogeneity of rain by developing multiple retrieval databases with different built-in inhomogeneity and rain intensity, along with a "rain indicator" to select the most appropriate database for each observed scene; 3) developing a new cloud-versus-rain partitioning that allows the use of a variety of drop size distribution assumptions to account for some of the natural variability diagnosed from the soundings; and 4) retrieving atmospheric and surface parameters just outside the rainy areas, thus providing information about the environment to help decrease the uncertainty of the rain estimates.

1. Introduction

Over the last 10 years, scatterometer ocean wind retrievals have proved very valuable to studies of air–sea interaction (e.g., Chelton et al. 2004; O'Neill et al. 2005; Minobe et al. 2008), ocean circulation (e.g., Lovenduski and Gruber 2005; Cunningham et al. 2007; Jiang et al. 2008), and weather phenomena such as El Niño and tropical storms (e.g., Chavas and Emanuel 2010). Assimilating scatterometer winds into numerical weather prediction models resulted in improving the forecast accuracy. Operational meteorologists have grown accustomed to basing their forecast of hurricane

Corresponding author address: Svetla M. Hristova-Veleva, Jet Propulsion Laboratory, Radar Science and Engineering, 4800 Oak Grove Dr., Pasadena, CA 91109.
E-mail: svetla.veleva@jpl.nasa.gov

DOI: 10.1175/JAMC-D-12-0237.1

formation, track, and intensity changes on scatterometer retrievals.

However, Ku-band scatterometer ocean wind retrievals are compromised when rain is present within the sensor's field of view (FOV). Undetected rain can lead to the retrieval of winds that are erroneously oriented at cross-track with respect to the satellite's motion and have speeds that are larger than both models and buoys suggest—although it could also lead to severe underestimates of hurricane wind speeds. With scatterometer data alone, correcting the effects of rain contamination is exceedingly difficult. Even flagging rain-contaminated data can be problematic (Huddleston and Stiles 2000).

The six-month-long *Advanced Earth Observing Satellite II* (*ADEOS-II*) mission provided a unique opportunity where coincident SeaWinds scatterometer and Advanced Microwave Scanning Radiometer (AMSR) multichannel radiometer measurements are available. We have developed an algorithm to estimate atmospheric parameters, in particular rain, from AMSR's brightness temperatures to provide atmospheric correction for scatterometer wind retrieval from the collocated SeaWinds scatterometer on board *ADEOS-II*. Although our algorithm uses AMSR observations, it can be easily tailored to any conically scanning microwave radiometer with a similar set of channels [e.g., the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI), the Special Sensor Microwave Imager/Sounder (SSMI/S), WindSat, and the Global Precipitation Measurement (GPM) Microwave Imager (GMI)]. Still today, accurate detection and estimation of precipitation from spaceborne radiometers remains a significant challenge. As Smith et al. (1998) pointed out, many of the discrepancies in the rainfall estimates come from differences in the assumed by the algorithms spatial inhomogeneity inside the sensor's FOV and in the assumed microphysics of the precipitation.

Remote sensing retrievals of geophysical parameters employ a priori relationships between the set of parameters to be retrieved and the observables associated with them. For physically based retrievals these relationships are obtained from physical descriptions of the atmosphere, and the underlying surface, within a radiative transfer model and take the form of either model functions or retrieval databases that are used as the basis for the development of the inversion: the retrieval algorithm (e.g., Wilheit et al. 1977; Mugnai and Smith 1988; Smith and Mugnai 1988; Tao et al. 1990; Kummerow et al. 1991, 1996, 2001; Smith et al. 1992; Simpson and Tao 1993; Wentz and Spencer 1998; Panegrossi et al. 1998; Olson et al. 1999). As such, the retrievals are subject to the assumptions that went into building the relationship between the observables and the parameters of interest. Hence, a critical step toward producing realistic retrievals is assuring that the representation of the geophysical parameters in the retrieval relationships is physically accurate (Kummerow et al. 2001).

The critical assumptions that affect rain retrievals from microwave observations are 1) beam filling (horizontal inhomogeneity within the sensor's field of view; 2) vertical inhomogeneity of the precipitation; 3) cloud versus rain partitioning; 4) assumptions about the particle size distribution; and 5) assumptions about the underlying surface (sea surface temperature and wind). All of these parameters vary significantly as a function of rain type (convective versus stratiform), geographical location (midlatitude versus tropical), and season.

To avoid biases, the precipitation variability must be reflected in the databases and must be accounted for in the point-by-point retrievals (Kummerow et al. 2004). As Panegrossi et al. (1998) pointed out, it is important to identify the typology of an observed event and to associate it with the appropriate retrieval database generated from hydrometeor profiles that adequately represent the microphysical, macrophysical, and environmental conditions of the storm.

For our algorithm, we built retrieval databases containing the atmospheric variables of interest, diagnosed from soundings obtained globally, along with the corresponding microwave brightness temperatures calculated using a radiative transfer model. Our algorithm addresses in a new way the issues of nonuniform beam filling and hydrometeor uncertainty by using a number of retrieval databases, each capturing a mode of the relation between the spatial inhomogeneity of rain and its intensity, as determined from our analysis of TRMM observations. The novelty of the algorithm is in its ability to select the retrieval database that is most appropriate for a particular observational scene by using a specially developed rain indicator to determine the presence, intensity, and degree of homogeneity of the rain within the sensor's FOV. We build our retrieval databases from globally observed atmospheric structures. This is in contrast with algorithms that use cloud model simulations of midlatitude squall lines and/or tropical convection.

Section 2 describes the algorithm in five subsections, section 3 discusses the benefits of the adopted approach, and section 4 illustrates the performance of the algorithm when applied to independent simulated data. Section 5 provides a summary and points to the areas of future work.

2. Description of the algorithm

In both nonrainy and rainy conditions the algorithm uses piecewise linear regression to retrieve the sea

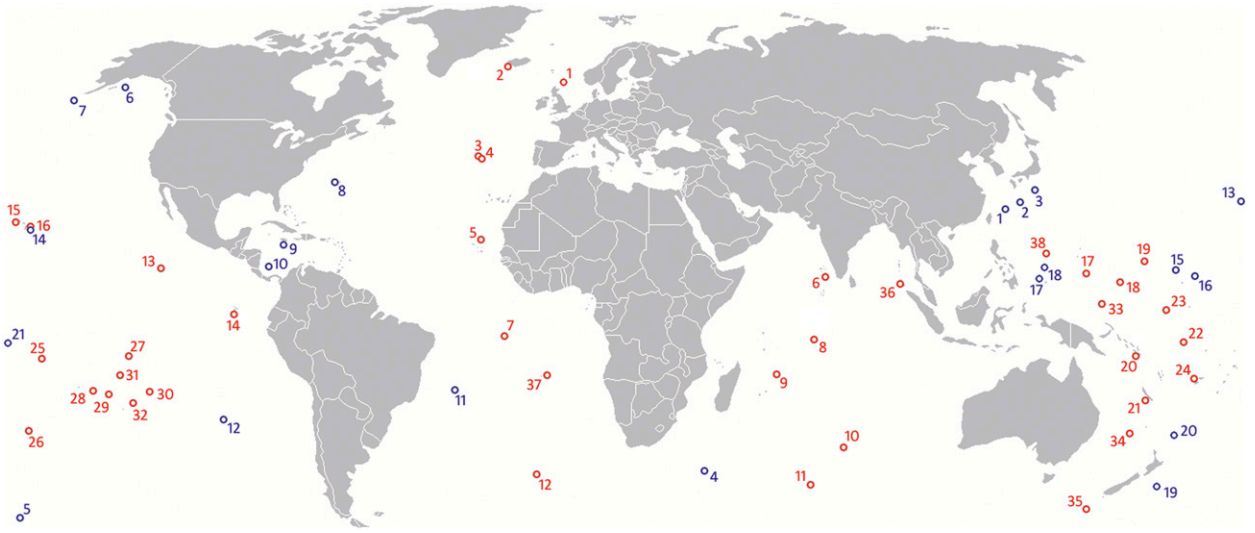

FIG. 1. Locations of the island soundings. Marked in red are the 38 sounding sites that were used to develop the retrieval databases. The 43 000 soundings were collected over a 3-yr period between 1998 and 2001. Marked in blue are the locations of the 20 sites that were used in testing the representativeness of the original sounding database. Data from these sites (20 000 soundings) were collected over a 3-yr period (1978–82) 20 years earlier.

surface temperature (SST; °C), the near-surface wind speed W (m s^{-1}), the columnar vapor V (mm), and the columnar liquid L (mm) in a scheme very similar to the one proposed by Wentz and Meissner (2000) for use in nonrainy situations. However, the approach we took has two distinct pieces: the development of four databases (one for nonrainy conditions and three for rainy conditions, each with different levels of beam inhomogeneity and rain intensity), and the development of a detector index we call the ''rain indicator'' (RI), whose value (along with its local sample variance) is used to select the most appropriate of the four retrieval databases to estimate the physical parameters, including column-average rain rate.

The retrieval databases were built using a large number of globally distributed soundings of atmospheric structures, whose underlying physics were diagnosed, and that were then composed according to our quantitative characterization of their spatial inhomogeneity. Section 2a describes the flow of the algorithm, as well as our derivation of precipitation profiles from soundings that we then input to the second step of the database construction. Section 2b defines the rain indicator, and section 2c completes the description of the second step of the database construction (which uses the rain indicator). Section 2d describes the details of vertically integrated liquid retrievals, and section 2e does the same for the vertically averaged rain-rate retrievals.

a. *Overview of the approach*

1) OFFLINE

Offline, we start with approximately 43 000 soundings from a number of island locations covering an ~3-yr period between 1998 and 2001 (Fig. 1), providing vertical profiles of temperature, pressure, and water vapor (V_v). A condensed-water profile (L_v) was diagnosed from each sounding following the method of Keihm et al. (1995) [see section 2c(1) below for details] after very satisfactory validation. Next, each of these profiles was used to create 60 different scenes, with the same temperature, pressure, water vapor, and condensed water profiles, by placing these over an ocean surface with SST and near-surface wind speed (W) that were varied in the ranges from −2° to 34°C and from 0 to 25 m s^{-1}, respectively. The resulting $60 \times 43\,000 = 2\,580\,000$ scenes include the expected range in all geophysical variables of interest (SST, W, V_v, and L_v). However, the expected correlations between the surface and atmospheric parameters are left out. In nature such correlations do exist. We chose not to incorporate them in the ensemble of scenes in order to allow for the retrieval of the usual and the unusual with no a priori preferential probability (e.g., winter cold outbreaks over the warm SSTs of the Gulf Stream). The ability of our algorithm to properly handle such extremes is discussed in section 4.

Once the ensemble of scenes was created, a one-dimensional Eddington approximation radiative transfer model (Kummerow 1993) and a surface emissivity model

(Wentz and Meissner 2000) were used to compute the top-of-atmosphere (TOA) brightness temperatures (TBs) at the AMSR frequencies and polarizations for each of the 2 580 000 scenes in the ensemble. Mie theory (Mie 1908) was used for computing the microwave scattering signatures of the rain. Attenuation by vapor and oxygen was also included. A database that contains the simulated brightness temperatures and the underlying geophysical variables (SST, W, V, and L, where V and L are the vertical integrals of V_v and L_v) was created in that manner.

We then split off those scenes where the column-integrated liquid L (the vertical integral of L_v) is less than 0.2 mm. This is our first database, the clear-air database. We then proceed to develop retrieval coefficients using linear regression approach. Acknowledging the high nonlinearity in relationships between brightness temperatures and geophysical variables, we make use of linearizing functions of the form $F(\mathrm{Tb}) = -\ln(290 - \mathrm{Tb})$. To address the remaining nonlinearity, we apply the iterative Wentz and Meissner (2000) multiple-regression procedure to this database, first deriving global linear-regression coefficients for each of the variables SST, W, L, and V and then deriving piecewise-linear refined regressions that are valid in subregions of the total four-dimensional parameter space. These piecewise linear regressions were computed using restricted datasets that were generated in the following manner: for each of the four geophysical parameters (SST, W, V, and L) the range of possible values was divided into up to five subregimes (including the global one), making for a total of 400 regimes; 400 restricted datasets were then generated by selecting from the general dataset only the members of the ensemble for which the four geophysical parameters fall within the specified bounds for this regime. Finally, multiple linear regression was used to derive sets of coefficients for each of the 400 regimes.

With those profiles whose L value exceeds 0.2 mm, three separate rain retrieval databases were created, each database consisting of ''composed'' observations (sets of nine scenes) having different degrees of homogeneity and rain intensity [see sections 2c(1) and 2c(2) for details]. As we did for the clear-air database, an iterative multiple regression procedure was applied to derive the coefficients of the piecewise-linear regressions to estimate L, V, W, and SST from the observed brightness temperatures. It should be noted that the 89-GHz brightness temperatures are not considered because of their high sensitivity to frozen particles and the associated high uncertainty in their modeling. The composed scenes in the three rainy databases reflect the three different modes that we observed in our analysis of the relation between the spatial inhomogeneity of the rain and its intensity as evident in the TRMM data, resulting in one reasonably homogeneous–low-intensity, one reasonably homogeneous–high-intensity, and one inhomogeneous database. The criterion and details are in section 2c(2) below.

Independently of the construction of the databases, we define a rain indicator scalar that can be computed from a given set of six coincident brightness temperatures [(19 GHz + 37 GHz + 89 GHz) × 2 polarizations, all at their native resolutions] along with the six brightness temperatures over the nearest clear-air pixel. In practice, a clear-air pixel is one whose column-integrated liquid L is less than 0.2 mm (and to calculate RI, we use for L the rough value given by the first global linear regression as explained in section 2b). Large values of RI correspond to higher rain rates, and larger values of its local sample variance (calculated from the eight closest neighbors) correspond to increased inhomogeneity: that is why we use both values to select the appropriate retrieval database.

2) In real time

In real time, we start by determining which of the four databases we will use to interpret a specific vector of brightness temperatures, using RI. We then apply the appropriate regression formulas depending on the database and proceed with the retrieval in a stepwise and iterative manner (as in Wentz and Meissner 2000).

The stepwise approach retrieves different geophysical parameters in consecutive steps, using different sets of channels. If various frequencies are to be used together to arrive at a certain geophysical parameter, all of the frequencies must be reduced to a common basis so that the measurements apply to the same area (these are the so-called resampled AMSR observations). This limits the resolution of the retrieved parameters to the spatial resolution of the lowest frequency used. Since the two lowest-frequency AMSR channels (the 6.9- and the 10.7-GHz channels) are the ones that are most sensitive to the surface parameters (SST and W), the spatial resolution of the retrieved SST and W is limited to the resolution of these two channels. During the retrieval of SST and W one could simultaneously retrieve also V and L. However, such an approach is not desirable since achieving the highest possible spatial resolution is always of interest. Indeed, this is why we use the stepwise approach. In particular, in the case of clear air, the surface parameters (SST and surface wind W) are retrieved first, using 10–37-GHz observations, all resampled to the 10-GHz footprint resolution. Next, columnar vapor (V) and columnar cloud liquid (CCL) are retrieved by using only 19–37-GHz channels all resampled to the 19-GHz footprint resolution. The already retrieved SST and W

values are used in an iterative manner to specify the surface conditions (the subregimes in SST and W) in which the V and L are to be retrieved.

The iterative nature of the algorithm is reflected in the following steps: when a geophysical parameter is retrieved for the first time, coefficients derived from the global unrestricted database are used. The initial estimate is used as guidance in selecting two sets of retrieval coefficients derived from the restricted databases that included only values encompassing our first-guess value. Then two new retrievals are obtained using the two restricted-set coefficients. A final estimate is then obtained by averaging the two restricted-set retrieved values, each one weighted by the relative distance between the initial estimate and the central value for each of the restricted sets. This procedure helps avoid discontinuities in the retrieved geophysical parameters.

In case of detected rainy conditions the retrieval proceeds in a similar manner with the following exceptions: 1) selecting the appropriate rain-specific dataset, and 2) SST, W, and V are not retrieved. In that case the values from the neighboring clear-air cells are interpolated to provide the environmental conditions in which the retrieval should proceed. Also, 3) the retrieval of the columnar liquid L is followed by cloud-rain partitioning and by the retrieval of column-averaged rain rate.

The algorithm flow is summarized in Fig. 2. What makes this algorithm different from other approaches is 1) the use of multiple retrieval databases; 2) the use of the RI to select the most appropriate database for a predicted inhomogeneity–intensity mode; and 3) a new cloud-versus-rain partitioning that allows the use of a variety of drop size distribution (DSD) assumptions to account for some of the natural variability in the cloud/rain partitioning diagnosed from the soundings.

b. Rain indicator

Cloud-model simulations of midlatitude and tropical storms (Hristova-Veleva 2000; Biggerstaff et al. 2006) were used to develop a passive microwave rain indicator that not only flags the observations for the presence of rain but also provides an approximate first estimate of the intensity of the rain.

1) DEVELOPMENT APPROACH

We used the modeled hydrometeor and thermodynamic profiles as inputs to instrument simulators to generate satellite observables and radar reflectivity at the TRMM precipitation radar (PR) and Quick Scatterometer (QuikSCAT) frequency, and microwave brightness temperatures at the frequency and polarizations of AMSR. Mie theory (Mie 1908) was employed for computing all single-particle scattering properties. The microwave brightness temperatures were computed using a one-dimensional Eddington's second approximation radiative transfer model (Weinman and Davies 1978; Kummerow 1993). Both liquid and frozen hydrometeors were accounted for in the forward computations of TBs and radar reflectivity. As a next step, the brightness temperatures at the model resolution were convolved with AMSR-like antenna patterns to simulate satellite-like observables. The so-generated synthetic passive microwave brightness temperatures were then related to the total liquid and the columnar maximum radar reflectivity (both from the model) to develop the rain indicator, as discussed below. The RI and its components were formulated to optimize its performance as both a rain flag and a tool for first assessment of the rain intensity and spatial variability.

The rain indicator incorporates emission and scattering signals and borrows ideas from Petty (1994a,b) and Liu et al. (1995). Microwave signals at the top of the atmosphere can be classified into two categories depending on how the microwave field interacts with the hydrometeors: 1) an emission signal that is dominant at lower frequencies and shows a warming signature in the presence of rain, and 2) a scattering signal that is dominant at higher frequencies and exhibits a cooling signature in the case of heavy precipitation. Because the emission signal tends to dominate in light rain conditions while the scattering signal tends to dominate in heavy rain cases, both have to be incorporated in the algorithms to cover the entire rainfall spectrum.

2) EMISSION SIGNAL

Strong emission in the atmosphere reduces the polarization difference in the highly polarized radiation emitted from the ocean surface. Hence, the polarization difference is representative of the atmospheric emission itself. Following Petty and Katsaros (1992), we define a measure of the normalized polarization difference (PD) as

$$\mathrm{PD} = \frac{\mathrm{TB}_V - \mathrm{TB}_H}{\mathrm{TB}_{\mathrm{VBackGround}} - \mathrm{TB}_{\mathrm{HBackGround}}}, \qquad (1)$$

where TB_V and TB_H are, respectively, the vertically and the horizontally polarized brightness temperatures at a particular channel and for a particular observation, while $\mathrm{TB}_{\mathrm{VBackGround}}$ and $\mathrm{TB}_{\mathrm{HBackGround}}$ represent the corresponding temperatures of the nonrainy surrounding areas. These background temperatures have to be continuously updated along the satellite track in order to properly reflect the geographically varying environmental conditions (see below).

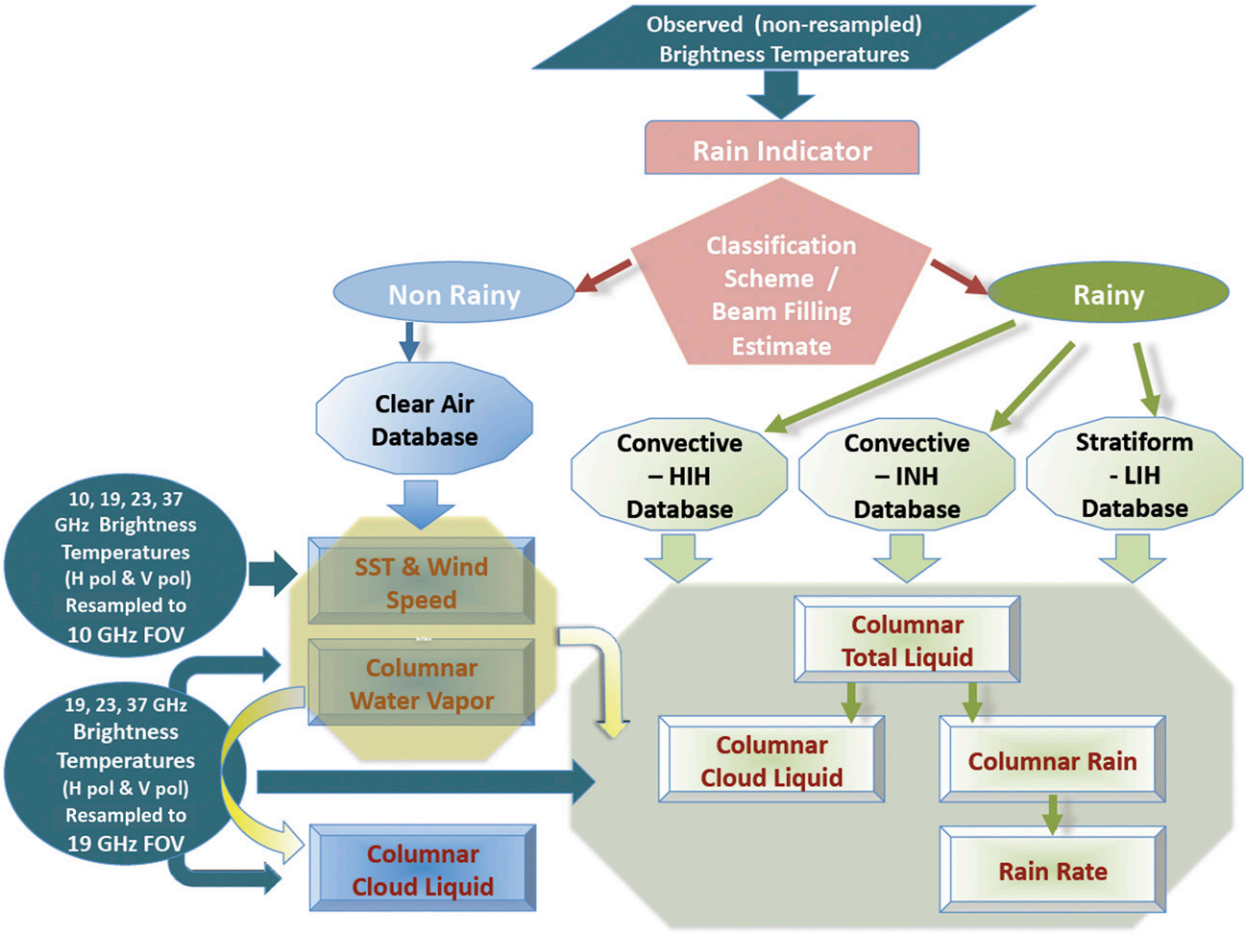

FIG. 2. Algorithm flow. As a first step, the rain indicator is computed from the observed brightness temperatures, using their native resolutions. A retrieval database is selected based on the value and the spatial variability of the rain indicator. In the next step, the geophysical retrievals are obtained using brightness temperatures that are resampled to the lowest spatial resolution, depending on the set of channels that are needed. In the case of clear air, a stepwise approach is used in which the surface parameters (SST and surface wind W) are retrieved first, using 10–37-GHz observations, all resampled to the 10-GHz footprint resolution. Next, columnar vapor (V) and columnar cloud liquid L are retrieved by using only 19–37-GHz channels all resampled to the 19-GHz footprint resolution. The already retrieved SST and W values are used to specify the surface conditions (the subregimes in SST and W) in which V and L are to be retrieved. In case of detected rainy conditions the retrieval proceeds in a similar manner with the following exceptions: 1) the appropriate rain-specific dataset is selected; 2) SST, W, and V are not retrieved—instead, the values from the neighboring clear air cells are interpolated to provide the environmental conditions in which the retrieval should proceed; and 3) the retrieval of the columnar liquid L is followed by cloud/rain partitioning and by the retrieval of column-averaged rain rate.

Equation (1) indicates that the smaller the PD, the farther a particular observation is from the highly polarized radiometric signatures of clear, nonrainy areas. The emission signal we use is a multichannel combination of PD signatures defined as

$$\mathrm{RI}_{\mathrm{emission}} = 1 - \frac{1}{\sum a_i} \sum a_i \times \mathrm{PD}_i,$$

where a_i is given in Table 1, and which incorporates PD signatures not only at 19 GHz but also at 37 and 89 GHz to take advantage of the higher sensitivity and better spatial resolution of the higher-frequency channels. Note that all brightness temperatures are used at their native resolution to compute the rain indicator (and its components). This is in contrast to the processes of retrieval of the geophysical variables in which the brightness temperatures are first resampled (e.g., Ashcroft and Wentz 2000) to the footprint of the lowest-resolution observations that are used—either 10 or 19 GHz in our case. Figure 2 provides for more details on the flow of the algorithm.

3) SCATTERING SIGNAL

The emission signal is complemented with a scattering signal built around the ideas of Grody (1991) and Spencer

TABLE 1. Weights with which the polarization difference signals from the different channels are combined to obtain the emission part of the rain indicator.

	19 GHz	23 GHz	37 GHz	89 GHz
Weights a_i	15	0	5	1

et al. (1989), who developed the polarization-corrected temperature (PCT) approach to enhance the precipitation signal by minimizing the effects of surface emissivity on microwave measurements. As Kidd (1998) pointed out, under this paradigm any rain will appear as lower PCT values because of the scattering of the upwelling radiation stream. Indeed, they illustrated that, when properly tuned, the PCT algorithm can provide the delineation of rain area.

We expand on these ideas and define our scattering signal as

$$\mathrm{RI}_{\mathrm{scattering}} = 1 - \frac{(1 + 0.818)\mathrm{TB}_{\mathrm{V89}} - 0.818 \times \mathrm{TB}_{\mathrm{H89}}}{(1 + 0.818)\mathrm{TB}_{\mathrm{V89BG}} - 0.818 \times \mathrm{TB}_{\mathrm{H89BG}}}, \tag{2}$$

where the nominator and the denominator are, respectively, the PCT of the point versus that of the background and the constant 0.818 was empirically determined by Spencer et al. (1989). Here subscript 89 refers to the 89-GHz brightness temperatures (85 GHz for TMI), subscript H stands for the horizontal polarization observations, and V stands for the vertical polarization measurements. Subscript BG denotes the background observations.

4) DETERMINING THE BACKGROUND CONDITIONS

As is apparent from Eqs. (1) and (2), the quality of the emission and scattering signals is dependent upon a good knowledge of the background brightness temperatures. To provide this knowledge we originally used the National Oceanic and Atmospheric Administration (NOAA)/National Environmental Satellite, Data, and Information Service (NESDIS) algorithm (Weng and Grody 1994) to estimate the liquid water path (LWP) in the atmospheric column:

$$\mathrm{LWP}_{19} = -2.7 \times [\ln(290 - \mathrm{TB}_{19\mathrm{V}}) - 2.96 - 0.39 \times \ln(290 - \mathrm{TB}_{22\mathrm{V}})]$$

or

$$\mathrm{LWP}_{37} = -1.15 \times [\ln(290 - \mathrm{TB}_{37\mathrm{V}}) - 2.89 - 0.34 \times \ln(290 - \mathrm{TB}_{22\mathrm{V}})].$$

The likelihood of drizzle or rain is high when $\mathrm{LWP}_{19} > 0.6$ mm or $\mathrm{LWP}_{37} > 0.2$ mm. Because of that, $\mathrm{LWP}_{37} < 0.075$ mm was used as a conservative indicator of clear or cloudy but nonraining conditions. The brightness temperatures observed at each channel under nonraining conditions were then used to define the background conditions that are needed in order to compute the polarization differences and the scattering signal.

In an effort to improve the definition of the background, we developed a new formula for the LWP estimation, based on our database. Equation (3) was used in our algorithm to provide the initial estimate of the LWP. Areas with LWP < 0.075 mm are considered to be cloud free and their brightness temperatures are used to define the background conditions:

$$\mathrm{LWP} = 0.035 + 1.328 \times [-\ln(290 - \mathrm{TB}_{37\mathrm{V}}) + 4.211] - 0.472 \times [-\ln(290 - \mathrm{TB}_{23\mathrm{V}}) + 4.047]. \tag{3}$$

5) RAIN INDICATOR

Finally, the emission and the scattering signals are combined to form our rain indicator:

$$\mathrm{RainIndicator} = a\mathrm{RI}_{\mathrm{emission}} + b\mathrm{RI}_{\mathrm{scattering}} + c\mathrm{RI}^{2.}_{\mathrm{scattering}}, \tag{4}$$

where $a = 4.0$, $b = 0.0$, and $c = 18.0$. These values, as well as the a_i coefficients of $\mathrm{RI}_{\mathrm{emission}}$ (Table 1), were determined by analyzing model-produced hydrometeors and relating them to forward simulated satellite passive microwave observations with the goal to optimize the relationship between RI (function of the observable brightness temperatures) and the underlying hydrometeor structures (total liquid, rain rate, maximum reflectivity in the column).

Figure 3 illustrates the capability of the rain indicator to distinguish between rainy and nonrainy areas. The values of RI generally range from -6 to $+6$. A value smaller than 0.5 indicates nonrainy conditions, as determined from our model analysis and validated through TRMM comparisons described below.

The rain indicator has the potential not only to separate rainy from nonrainy conditions but also to identify the rain intensity. Figure 3 clearly shows that RI is quite successful in depicting finer structures in the spatial variability of the reflectivity and columnar liquid fields. A value of $0.5 < \mathrm{RI} < 4.2$ has been found to indicate the presence of light to moderate rainy conditions, while a value of $\mathrm{RI} > 4.2$ is indicative of the presence of moderate to heavy rainy conditions. Our analysis of the

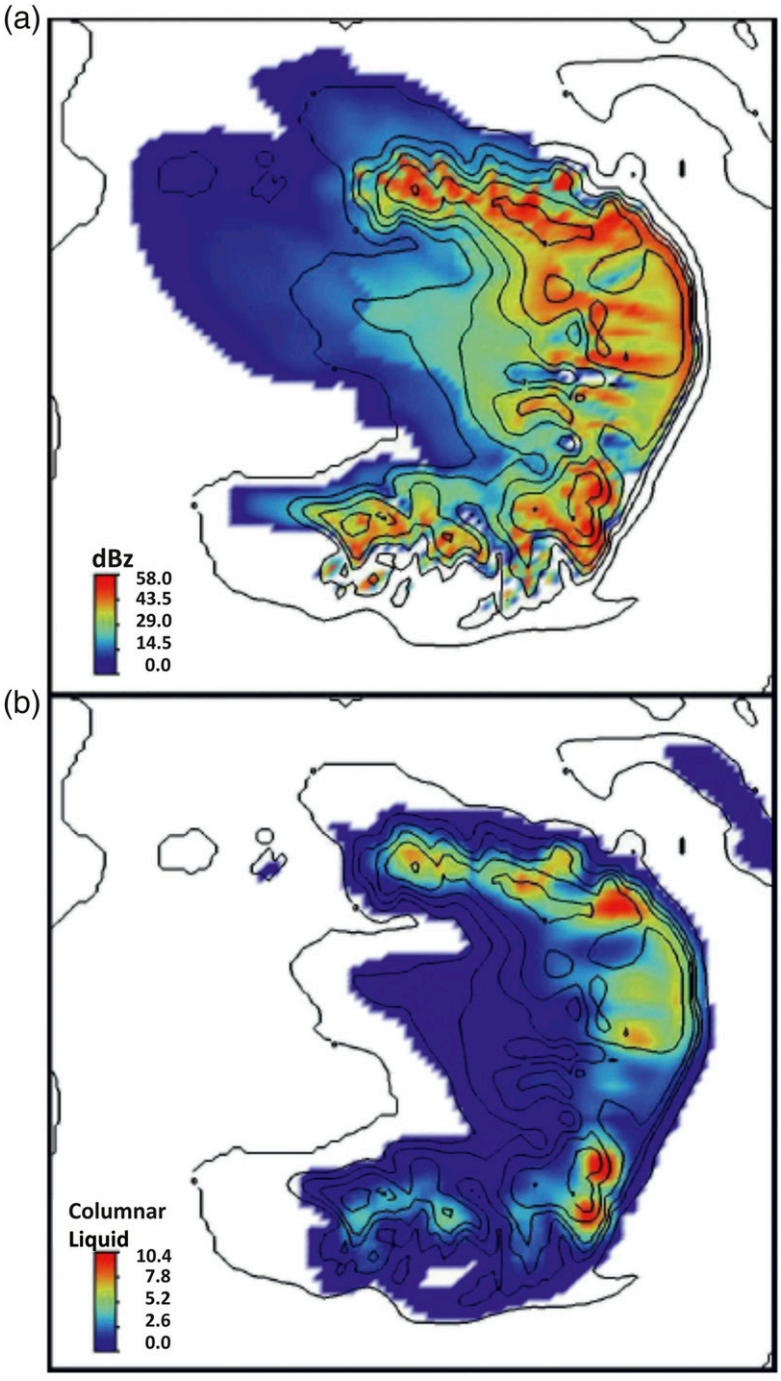

FIG. 3. Presented is the spatial correspondence between the rain indicator (contoured levels 1–6) and (a) the maximum reflectivity in the column (dBZ) and (b) columnar liquid (mm). The columnar liquid has $15 \times 15\,\text{km}^2$ spatial resolution. It is a running average of the higher-resolution ($3 \times 3\,\text{km}^2$) cloud model data. The hydrometeor profiles came from a cloud-resolving simulation of a midlatitude squall line (Hristova-Veleva 2000). These profiles were used as input to a full radiative transfer model and a surface emissivity model to compute TOA brightness temperatures at the AMSR frequencies and polarizations. Mie scattering was used to compute the microwave signatures of the rain. The simulated brightness temperatures were then used to compute the rain indicator and to relate it to the simulated columnar liquid and radar reflectivity.

TABLE 2. Statistics of the rain detection capabilities of the rain indicator in comparison with the TRMM estimates of rain.

	Rain by TRMM	No rain by TRMM
Rain by current RI	86 595 ÷ 90 438 = 95.75%	38 944 ÷ 2 467 728 = 1.58%
No rain by current RI	3843 ÷ 90 438 = 4.25%	2 428 784 ÷ 2 467 728 = 98.42%

precipitation effect on the measured scatterometer signal (not shown; see S. Hristova-Veleva et al. 2013b, unpublished manuscript) indicates that these two rain regimes differ in the dominant mode of this effect: the attenuation-dominated effect for moderate rain conditions versus the enhancement effect for heavy rain due the domination of the rain-associated volume backscatter signal.

The rain indicator algorithm was applied to a limited number of TMI observations to investigate and validate its performance. The rain indicator values were computed using TMI observations. We then performed rain/no rain classification using the RI values and thresholding on RI > 0.5. We compared this rain flagging to the flagging we would obtain by using the rain estimates from the TMI GPROF algorithm (Kummerow et al. 2001). The RI showed a very good ability to flag as rainy (RI > 0.5) any TMI retrievals that had even minimal amounts of liquid precipitation (in excess of 0.05 mm h^{-1} at any given level): 95.75% agreement in identifying rainy conditions and 98.42% agreement in identifying nonrainy conditions, thus bringing the overall rate of having the two independent rain detections in agreement to 98.33% (see Table 2).

c. *Creation of retrieval databases with modeled inhomogeneity*

One ''clear air'' and three rainy databases (used for development of the retrieval coefficients) are constructed, with each of the rainy databases having different degrees of inhomogeneity. The steps are as shown below.

1) CREATING THE CLEAR AND RAINY SCENES; DIAGNOSING THE COLUMNAR VAPOR AND COLUMNAR LIQUID FROM THE SOUNDINGS; PARTITIONING BETWEEN CLOUD AND RAIN

As already described, a large set of soundings was used together with randomly varying surface conditions (SST and W) to construct a large number of scenes. The radiosonde data were quality controlled to assure that they contained enough vertical levels, that they contained data from the lower part of the atmosphere, and that there were no significant jumps between the vertical levels. The radiosondes provided atmospheric profiles of temperature, pressure, and dewpoint depression. Columnar vapor and columnar condensed water were diagnosed from the soundings following the same approach used by Keihm et al. (1995). The columnar condensed water is diagnosed from the soundings by assuming that condensation is deemed present in atmospheric layers for which the radiosondes indicate relative humidity $> 94\%$, and then assuming that half the amount of water corresponding to the difference in absolute humidity between the cloud base and the cloud altitude in question was in the form of condensate (liquid or frozen). The factor 0.5 for computing the density of the condensed water is based on aircraft measurements of liquid density within stratus and cumulus clouds (Keihm et al. 1995).

Figure 4 shows the histograms for the distributions of the diagnosed columnar vapor and liquid. It is obvious that the sounding database has quite uniform distribution of columnar water vapor across the range of expected values (5 to 65 mm). Figure 4b shows a rather different distribution for the columnar condensed water. The number of cases almost exponentially decreases with the increasing value of columnar liquid. Among our soundings it was found that about 82% contained no precipitation (columnar liquid value $L < 0.2$ mm), 15% contained light to moderate precipitation (columnar value in the range 0.2–3 mm), and about 3% contained heavy precipitation. Indeed, this is the type of distribution noted in the literature (e.g., Wentz and Spencer 1998). All this indicates two very important things: the soundings database is very comprehensive, and the method developed by Keihm et al. (1995) and used here to diagnose the columnar liquid in the atmosphere is indeed appropriate.

We used the soundings with $L < 0.2$ mm to create our clear-air retrieval database and used all soundings with $L > 0.2$ mm to create a number of rainy databases. The next step in the diagnosing of the precipitation from these sounding data involves the partitioning of the diagnosed columnar condensed water into cloud and rain. A threshold value for the cloud mixing ratio [0.7 g kg^{-1}, suggested by Rutledge and Hobbs (1983)] was set to determine the partitioning: at each level, condensate mixing ratio amounts that exceeded this cloud mixing ratio threshold were considered to be precipitation. Figure 5 shows how our algorithm's partitioning between cloud and rain performs when applied to ''soundings'' from a cloud-resolving model (CRM) run with parameterized microphysics and compares it to the

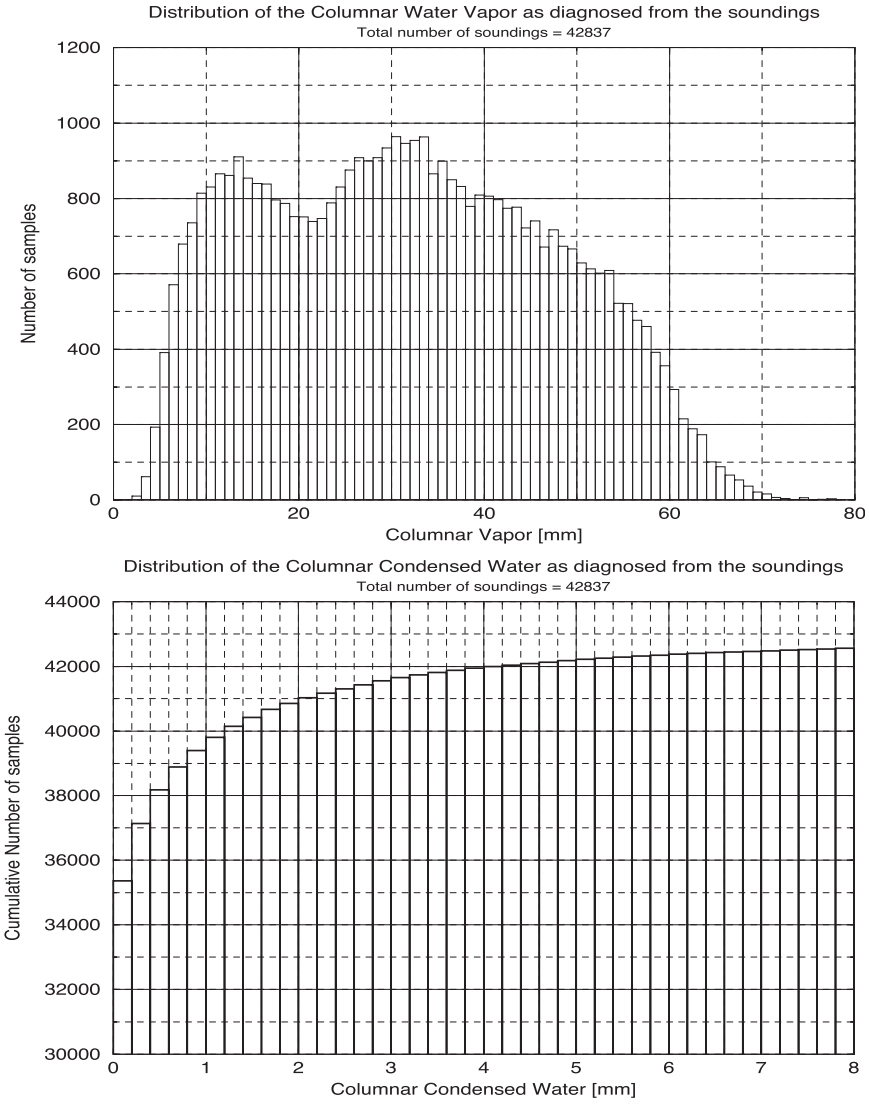

FIG. 4. Distribution of the atmospheric parameters (columnar vapor and liquid) that were diagnosed from the soundings and were used in developing the retrieval databases. (top) The distribution of columnar vapor in 1-mm bins. (bottom) The cumulative distribution of the columnar liquid in bins of 0.2 mm.

partitioning produced by the CRM run. An inspection of the scatterplots (black for our algorithm and red for the CRM partitioning) shows that our algorithm produces a more narrow distribution whereas the relationship in the CRM is less well defined, with higher scatter of the points. However, it is noteworthy that the cloud–precipitation partitioning from the two sources results in relationships with reasonably similar, on average, structures. This is illustrated by comparing the green (from algorithm) and blue (from CRM run) curves that represent the ''mean'' cloud–precipitation relationship as depicted by the median of 150-point running samples. Considering the complexity of the microphysical processes, we conclude that the two relationships compare reasonably well in an average sense.

Once the cloud–precipitation partitioning was done for the rainy soundings, TOA brightness temperatures were computed for each clear and rainy scene at the AMSR frequencies, polarizations, and incidence angles. Mie scattering and Marshall–Palmer DSDs were used for the rain. In addition, the RI was computed for each scene and included as part of the database.

2) CREATING THE RAINY DATABASES

To quantify the inhomogeneity of rain within spaceborne radiometer fields of view, we started by analyzing

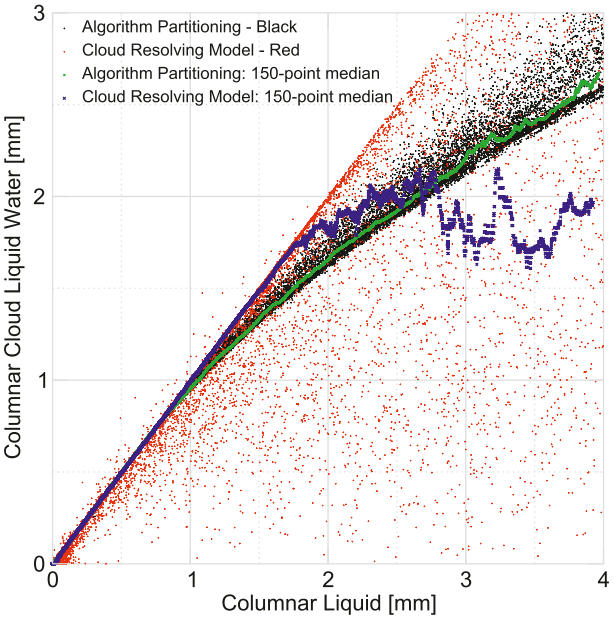

FIG. 5. Relationship between the columnar liquid (x axis) and its cloud liquid portion (y axis). The figure shows how our algorithm's cloud–rain partitioning (black) performs when applied to "soundings" from a CRM run with parameterized microphysics and compares that partitioning to the one produced by the CRM run (red). For the purpose of the comparison, the cloud mixing ratio threshold value was set to $1.25\,g\,kg^{-1}$—the value that was used as a threshold for autoconversion from cloud to rain during this particular CRM run. However, when creating the retrieval database, we used a threshold value of $0.7\,g\,kg^{-1}$ as suggested by Rutledge and Hobbs (1983). The green and blue curves show the median from each 150-point running sample; green represents the relationships from our algorithm while blue reflects the same relationship but from the CRM run.

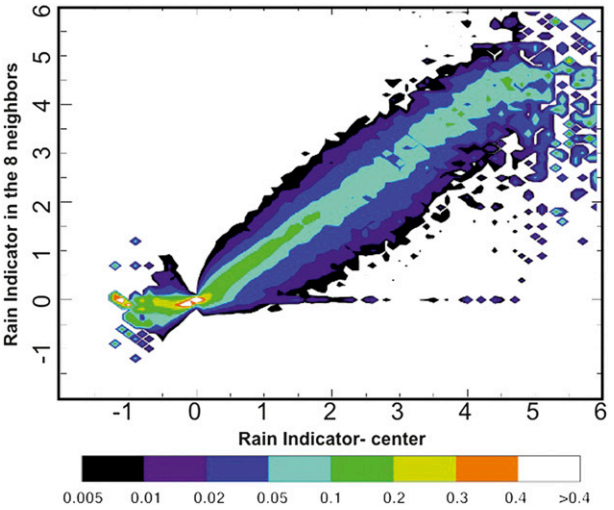

FIG. 6. TMI-based frequency of the distribution of the rain indicator for the immediate eight neighbors for a given value of the central observation. As indicated by the color bar, the percent of the data is contoured at levels of: 0.005, 0.010, 0.020, 0.050, 0.100, 0.200, 0.300, and 0.400.

TMI observations (February1998), systematically computing RI as in Eqs. (1)–(4) and comparing the value for a given beam with the values in its eight immediate neighbors (Fig. 6). Since consecutive TMI observations are approximately 10 km apart, this sample variance effectively quantifies the spatial variability of precipitation within areas of $\sim(25 \times 25)\,km^2$. We then classified the spatial variability of the TRMM data into the three homogeneity classes: the lighter intensity, reasonably homogeneous class (LIH), for which all nine RI values were in the range 0.5–2.5; the heavier intensity, reasonably homogeneous class (HIH), for which all nine RI values were in the range 2.5–6; and the inhomogeneous class (INH), for which the only requirement was that the central RI value was greater than 0.5 (i.e., precipitating). The TMI statistics showed that the INH class was the one most often observed in areas of precipitation (~65% of all rainy cases). The next most common class was the LIH class (~30% of the cases), and a distant third was the HIH class with ~5% of all rainy cases. Figure 7 illustrates the TMI-derived relationship between RI and its values over the eight immediately contiguous beams for two of the classes (INH and LIH).

Armed with this analysis, we used our sounding-derived profiles to construct three rainy databases corresponding to the three inhomogeneity classes defined above. We did this by considering nine time-consecutive profiles at a time from every sounding site. The nine profiles had the same SST and surface wind (by design we varied the surface parameters independently of the atmospheric profiles) but variable vertical profiles of temperature, vapor, and condensed liquid, all diagnosed from the time sequence of soundings at this station. This approach assured that the background brightness temperatures were well defined and the RI was reflecting the intensity of the precipitation in the profile. In constructing the three databases we applied the following rules: whenever all nine profiles had RI values that were between 0.5 and 2.5, the profiles were averaged and the result was included in the LIH database; whenever all nine RI values were between 2.5 and 6, their average profile was included in the HIH database; otherwise, if at least one value was below 2.5 and one exceeded 2.5, and provided the fifth (i.e., "center") profile had an RI greater than 0.5, the average profile was included in the INH database. As such, the average profile in the INH database is computed from nine profiles some of which could, and do, have no, light-to-moderate, or heavy precipitation. In the case of the geophysical variables (column liquid and average rain rate), the averages were

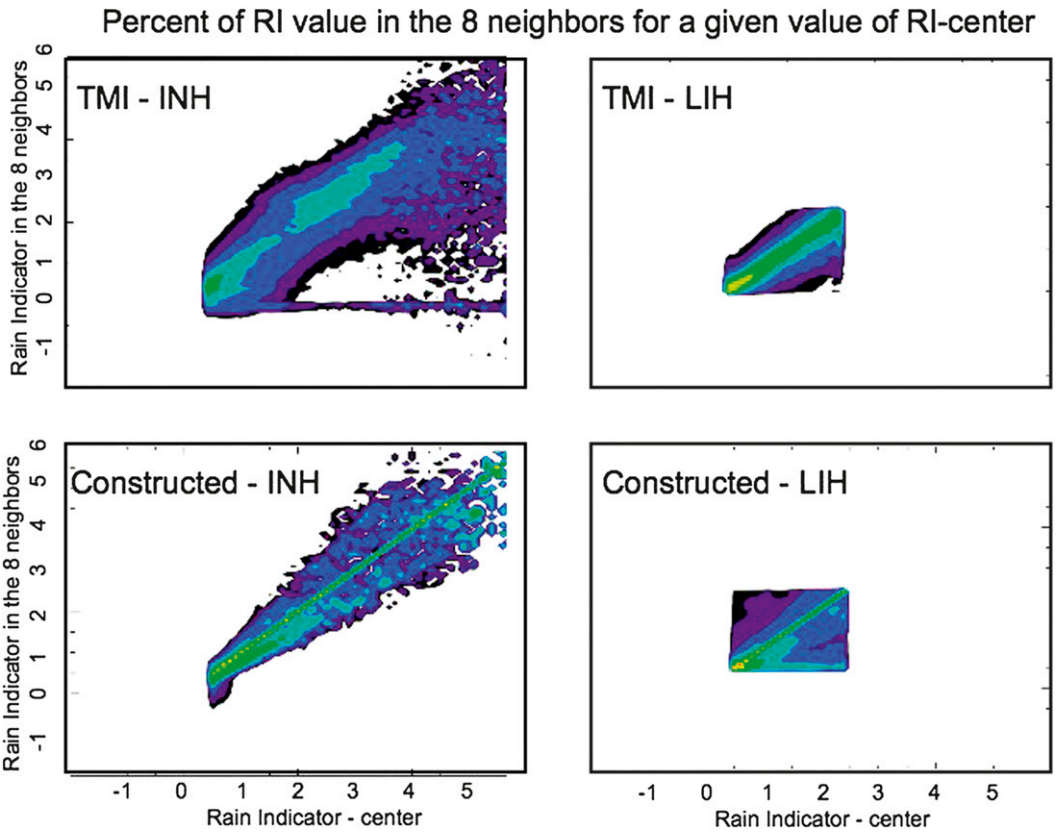

FIG. 7. (top) As in Fig. 6, but for the TMI-based relationships for two of the homogeneity classes: (left) INH and (right) LIH. (bottom) These relationships are compared with the ones that were built into our constructed databases.

calculated linearly; however, to calculate the corresponding ''average'' brightness temperatures we applied the weighting that would correspond to the 19-GHz antenna pattern if the nine profiles were arranged as a 3×3 square centered on the fifth profile.

It should be pointed that in our approach we represent the spatial inhomogeneity making use of time-varying vertical atmospheric profiles that were collected every 6 to 12 h at a given station. We argue that this is a reasonable approach on the following grounds. 1) Since the precipitation varies on shorter time scales than columnar vapor, the nine selected profiles vary mostly in their liquid profiles, followed by variability in the humidity profiles. These two characteristics are also typical of the spatial variability of the geophysical variables since the horizontal scales of vapor, SST, and surface wind are larger compared to the scale of clouds and precipitation. Also, 2) the degree of spatial inhomogeneity we design reflects the observed, as depicted by our analysis of TRMM data.

d. Retrieving the columnar liquid; selecting the appropriate retrieval database

At each AMSR observation, the rain indictor at that point is considered together with the RI of the eight surrounding observations to assess the inhomogeneity and rain intensity in the observed scene. Based on that, the appropriate retrieval database is selected. In each case, the corresponding retrieval coefficients are used together with the set of resampled brightness temperatures for the two-step piecewise multiple linear regression estimation of the columnar liquid as described in section 2a. The algorithm flow is shown in Fig. 2 and its application is illustrated in Fig. 8.

e. Retrievals of rain rate

Estimating the rain (in addition to the total condensed liquid discussed in sections 2a and 2d) is important because the rain contributes to all three components of the atmospheric signature in the scatterometer observations: attenuation, backscatter from the rain, and surface roughening that is proportional to the rain rate. This is in contrast to the cloud liquid, whose reflectivity is so low at Ku band that it is a negligible in our problem.

To estimate column average rain rate, we could apply the same approach that we used for the total column liquid (i.e., derive offline the piecewise-linear regressions that are appropriate for each database). The retrieval would then consist in using RI to identify the

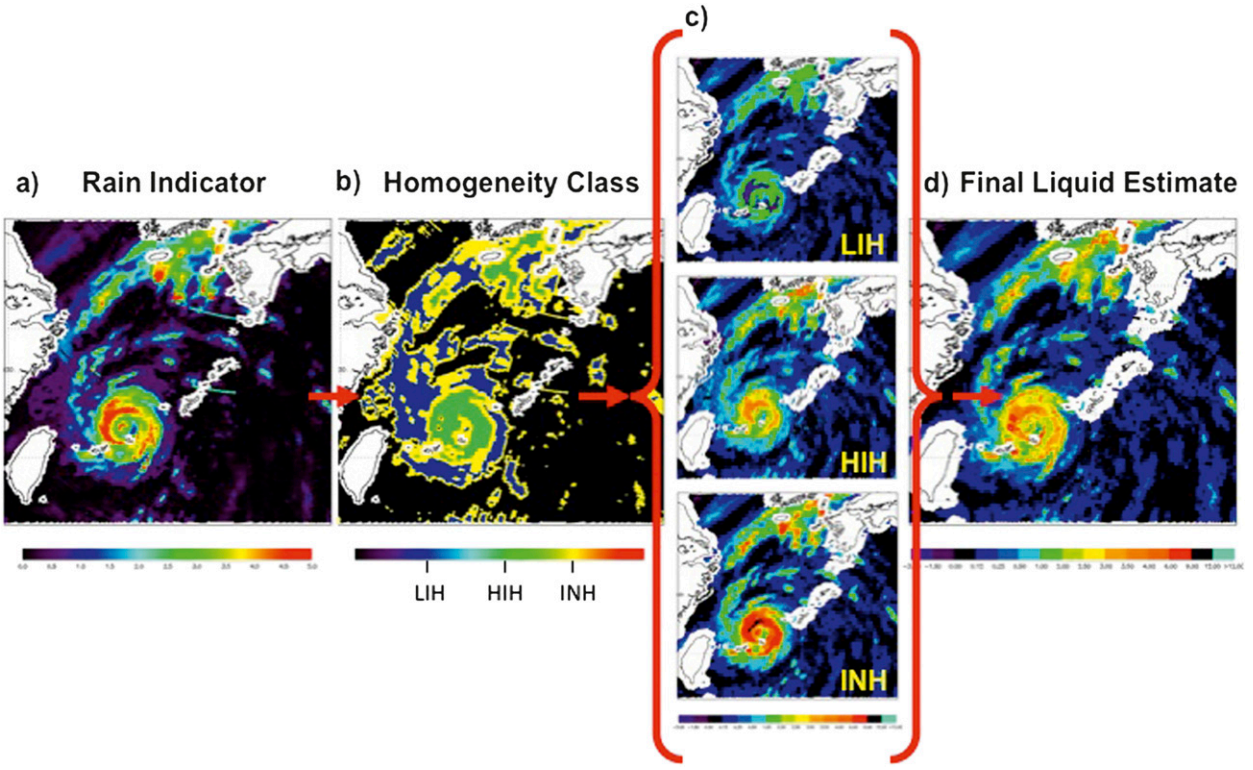

FIG. 8. Illustration of how the algorithm works (using data from Supertyphoon Maemi, 11 Sep 2003). (a) The rain indicator. (b) The inhomogeneity classification based on the values and the spatial variability of RI. (c) With this information in hand, the appropriate retrieval database is used at each observation [retrieval shown if (top) LIH, (middle) HIH, and (bottom) INH is used at all points that were identified by the classification algorithm as rainy]. (d) The final result, combining all three retrievals so that at each point the appropriate database is used based on the RI classification. The high spatial consistency of the final retrieval, the increased detail, and the spatial gradients of the retrieved precipitation all testify to the value of the approach.

database that is appropriate given the measured brightness temperatures at hand, then applying the regressions that are appropriate for the particular database and the particular subcube containing the measurement. However, this approach would not guarantee any consistency between the retrieved column-integrated liquid and the retrieved column-averaged rain rate.

To ensure consistency, we consider the main difference between the signature of the precipitation liquid and that of the cloud liquid in our application, namely the attenuation of the scatterometer signal. Assuming monodisperse clouds (with 15-μm-diameter droplets), we first relate the total 13.4-GHz path attenuation A_c due to a total column cloud mass C using a linear fit to the calculated attenuation corresponding to the profiles in our clear-air database:

$$A_C = a_C C. \tag{5}$$

Similarly, for each of our rainy databases we obtain a relationship between the column liquid mass (cloud plus rain) and the calculated for each profile 13.4-GHz attenuation A_L. We obtain the database-specific coefficients, using this time a quadratic fit:

$$A_L = a_{1L} L + a_{2L} L^2. \tag{6}$$

Note that the coefficient $a_C = 0.0326[1 - 0.018(T_C - 283)]$ is a function of the SST since T_C, the representative cloud temperature, is taken to be the mean temperature between the surface and the freezing level [0.5(SST + 273.16)]. On the other hand, the coefficients a_{1L} and a_{2L} are not functions of any other geophysical variables, but are defined separately for each of the three homogeneity classes (see Table 4).

Next, we obtain an expression for the columnar attenuation by rain A_R that would correspond to a columnar rain mass M_R. To that end, we start by using a relationship between the rain rate RR (mm h^{-1}) and the associated 13.4-GHz DSD-dependent attenuation A_{RR} (dB km^{-1}) as given by Haddad et al. (1997):

$$A_{RR} = \alpha(RR)^{\beta}. \tag{7}$$

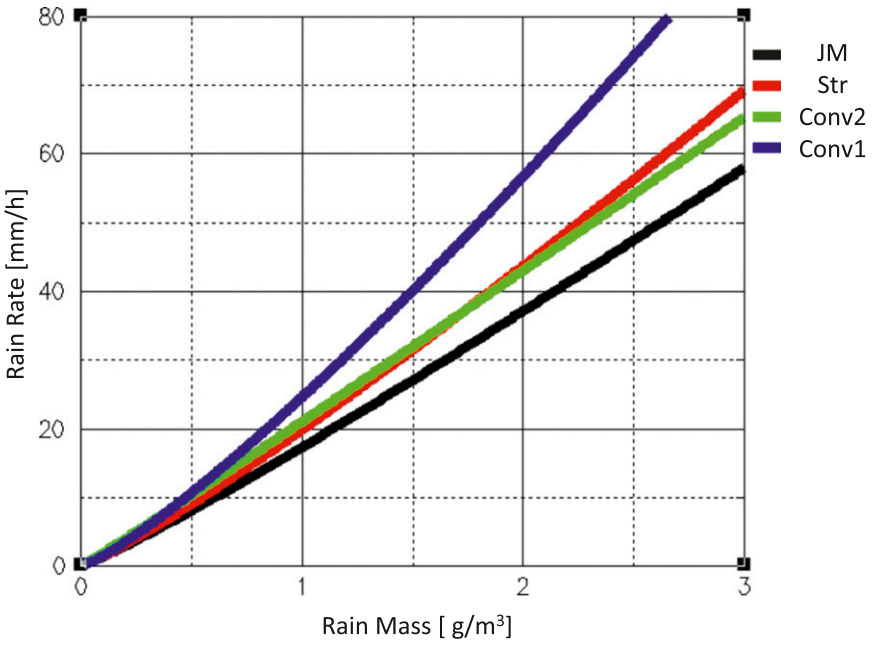

FIG. 9. Rain rate as a DSD-dependent function of the rain density. Three of the cases (conv1 in blue, conv2 in green, and str in red) came from observations in convective and stratiform rain (Battan 1973) while the fourth (JM in black) was developed by J. Meagher (2003, personal communication) using DSDs observed during Tropical Ocean and Global Atmosphere Coupled Ocean–Atmosphere Response Experiment (TOGA COARE).

We next acknowledge that the rain rate is often expressed as a DSD-dependent function of the rain density m_R ($\mathrm{g\,m^{-3}}$) obtained by dividing the columnar rain mass M_R ($\mathrm{kg\,m^{-2}}$) by the height of the rainy layer H (km):

$$\mathrm{RR} = a m_R^b = a(M_R/H)^b, \tag{8}$$

where a and b are DSD-dependent coefficients. The DSD assumptions play a significant role in determining the relationship, especially at the higher rain rates as illustrated in Fig. 9. Table 3 provides the coefficients a and b for the four functions in Fig. 9. Table 4 summarizes the set of assumptions (coefficients) we used as a function of the rain indicator–based classification of the rain intensity and homogeneity of the observed scene.

As (8) indicates, in order to obtain the rain rate from the columnar rain M_R, we need a reasonably accurate estimate of the rainy layer depth H. Following Wentz and Spencer (1998) we estimate H (km) as a function of the AMSR-derived sea surface temperature in the following manner:

$$\begin{aligned} H &= 0.9286 + 0.1374\mathrm{SST} + 0.003\,64\mathrm{SST}^2 \\ &\quad - 0.000\,126\,8\mathrm{SST}^3 \quad \text{for} \quad \mathrm{SST} < 29.84^\circ\mathrm{C}; \\ H &= 5 \quad \text{for} \quad \mathrm{SST} > 29.84^\circ\mathrm{C}. \end{aligned} \tag{9}$$

We can finally combine (7)–(9) to obtain an expression for the vertically integrated nadir rain attenuation A_R [note that A_R (dB) stands for the attenuation by the columnar rain, while A_{RR} ($\mathrm{dB\,km^{-1}}$) is the attenuation associated with a particular rain rate ($\mathrm{mm\,h^{-1}}$) and the corresponding rain density ($\mathrm{g\,m^{-3}}$)]:

$$A_R = HA_{\mathrm{RR}} = H\alpha \times \mathrm{RR}^\beta = H\alpha[a(M_R/H)^b]^\beta. \tag{10}$$

We are now ready to combine all pieces of information, acknowledging that the AMSR-derived columnar liquid L ($\mathrm{kg\,m^{-2}}$) is equal to the sum of the unknown columnar cloud liquid C ($\mathrm{kg\,m^{-2}}$) and the unknown columnar rain mass M_R ($\mathrm{kg\,m^{-2}}$):

$$L = C + M_R, \tag{11}$$

TABLE 3. Coefficients a and b in the relationship $\mathrm{RR} = am_R^b$ in Eq. (8) and Fig. 9. The rain rate (RR) is in millimeters per hour, and the rain mass (m_R) is in grams per meter cubed.

	JM	Str	Conv2	Conv1
a	17.348	19.885	21.072	24.628
b	1.099	1.1364	1.0309	1.2048

TABLE 4. Class-specific selections for the attenuation–RR relationships [as in Haddad et al. (1997), from their Tables XI and XII for the 13.8-GHz frequency and the choices of D'' given below, where D'' (mm) is the mass-weighted mean drop diameter that is uncorrelated with the rain rate], the RR–M_R relationships listed in Table 3, and the attenuation–liquid relationship.

	LIH	HIH	INH
Attenuation–RR [Eq. (7)]	$D'' = 1.2$	$D'' = 1.1$	$D'' = 1.1$
	$\alpha = 0.0283$	$\alpha = 0.0253$	$\alpha = 0.0253$
	$\beta = 1.134$	$\beta = 1.145$	$\beta = 1.145$
RR–M_R [Eq. (8)]	JM	Conv1	Conv1
	$a = 17.348$	$a = 24.628$	$a = 24.628$
	$b = 1.0990$	$b = 1.2048$	$b = 1.2048$
Attenuation–liquid [Eq. (6)]	$a_{1L} = 0.042\,373\,7$	$a_{1L} = 0.081\,616\,5$	$a_{1L} = 0.035\,119\,4$
	$a_{2L} = 0.011\,743\,1$	$a_{2L} = 0.017\,016\,3$	$a_{2L} = 0.027\,003\,6$

and, consequently, the columnar liquid attenuation A_L is equal to the sum of the columnar cloud attenuation A_C and the columnar rain attenuation A_R:

$$A_L = A_C + A_R. \tag{12}$$

Using (5), (6), and (10) and substituting in (12) we obtain an equation for the unknown M_R, having used (11) to substitute $L - M_R$ for C in the latter equation. The result is

$$a_{1L}L + a_{2L}L^2 - a_C L = (H^{1-\beta b}\alpha a^{\beta} M_R^{\beta b-1} + a_C)M_R, \tag{13}$$

which we solve for M_R (given L) in five iterations. On the first iteration we assume that $\beta b - 1 = 0$ and thus $M_R^{\beta b-1} = 1$. This assumption is workable since usually $\beta \approx 1.14$ and $b \approx 1.15$, leading to $\beta b \approx 1.3$. After obtaining the first estimate of M_R we compute $M_R^{\beta b-1}$ and use this new estimate in the following iterations. Another important note to make is that during the iterations we use Eq. (13) if the denominator P,

$$P = \left(\frac{H^{1-\beta b}\alpha a^{\beta}}{10\log e} M_R^{\beta b-1} - a_{0C}\right), \tag{14}$$

$\neq 0$; otherwise, we determine M_R from $P = 0$.

Finally, an additional constraint is put on the estimated rain M_R:

if $L < 0.3$ or > 1.5, then $M_R = L - (0.008\,324 + 0.7257L - 0.1112L^2 + 0.007\,896L^3 - 0.000\,190\,9L^4)$.

The constraint uses a parameterized relationship for the cloud versus rain partitioning that was developed based on diagnoses from the sounding quantities. It assures that 1) the cloud liquid water will not start decreasing for higher values of M_R after a max of $C = 0.75$ mm at $L \approx 1.5$ mm, and 2) the estimated cloud liquid will not exceed 1.75 mm.

Figure 10 illustrates the performance of the cloud–rain partitioning and shows how our approach allows for scatter in the L versus M_R relationship in the range $0.3 < L < 1.5$ mm. While the upper bound of $L = 1.5$ mm might appear restrictive, the statistics of the AMSR-retrieved columnar liquid L show that ~94.5% of the rainy data fall below that limit. We believe that the scatter in Fig. 10 reflects the nonuniqueness of that relationship as observed in real precipitation. Indeed, this conclusion is supported by the results in Fig. 11, which compares the cloud–liquid relationship reconstructed from the AMSR-retrieved columnar liquid to the same relationship as diagnosed from the sounding database. As Fig. 11 illustrates, our cloud–rain partitioning allows reconstructing of much of the naturally observed variability.

Once we have obtained the columnar mass of rain M_R (kg m^{-2}), we use Eq. (8) to solve for the desired rain rate

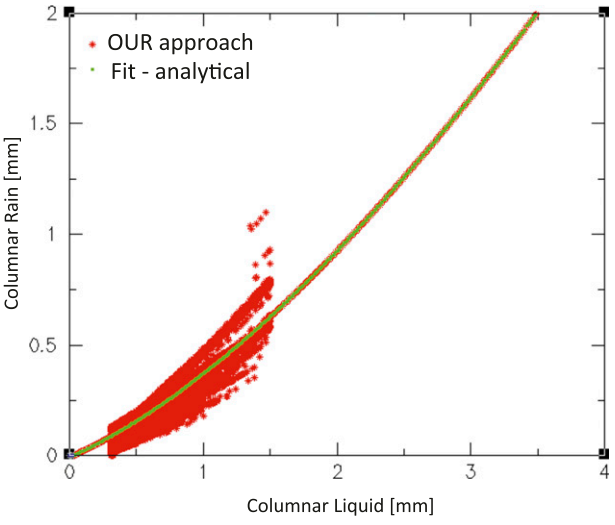

FIG. 10. Relationship between the columnar liquid retrieved from AMSR and the columnar rain mass estimated by our cloud–rain partitioning.

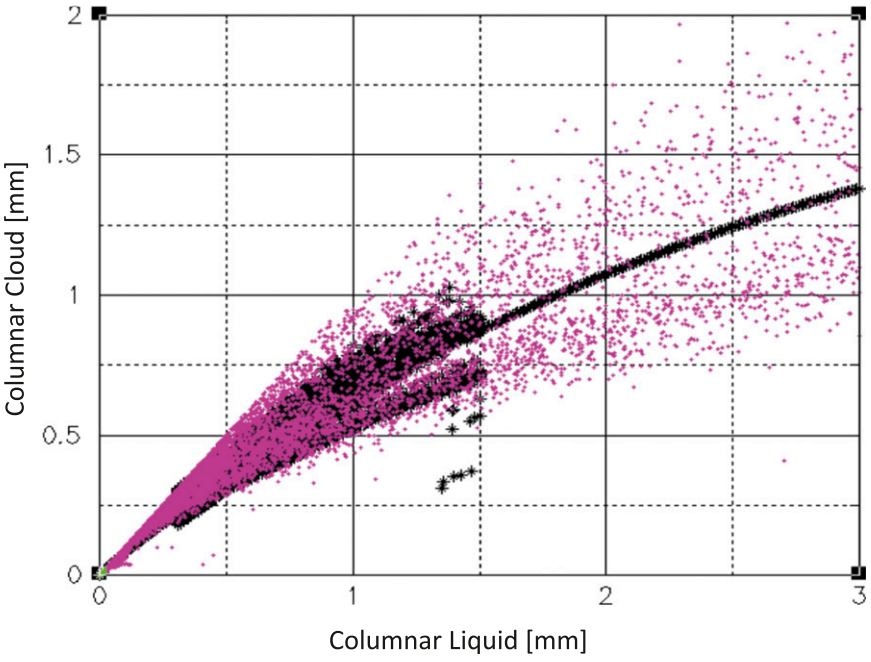

FIG. 11. The relationship of the columnar cloud to the columnar liquid from two different sources: our cloud–rain partitioning algorithm applied to our AMSR-retrieved columnar liquid from *ADEOS-II* orbit 3557 (black), and the cloud–rain partitioning as diagnosed from the sounding database (magenta).

RR. The behavior of the cloud-rain partitioning was analyzed and the following were found: 1) rain in excess of 0.05 mm h^{-1} begins when the total liquid $L > 0.1$ mm; 2) the cloud liquid C increases with rain; and 3) the cloud versus rain relationship levels off at high rain with C reaching a maximum value between 1 and 2 mm. These three characteristics were found to be true for 38 northeast Pacific extratropical cyclones (Wentz 1990), with the exception that the onset of rain was associated with $L > 0.18$ mm. Indeed, Wentz and Spencer (1998) developed a parameterization of cloud-rain rate partitioning based on the above study.

3. Unique features of our rain retrieval approach

The proposed approach has several unique features:

a) It deals with the beam-filling problem explicitly by quantifying the observed inhomogeneity and using it to perform the retrievals with the most appropriate for the observed scene retrieval database, selected from a number of offline developed retrieval databases, each representing different rain intensity and spatial inhomogeneity scenarios, determined based on a priori knowledge of the dominant modes of spatial variability.

b) The cloud structures that we use were diagnosed from soundings. In particular, the height of the freezing level, the cloud-base heights, and the cloud depths come directly from measurements.

c) The cloud-versus-rain partitioning that is part of the rain rate retrievals is unique. It allows for using a variety of DSD assumptions—possibly a different one for each of the identified inhomogeneity classes. As illustrated, this approach allows us to reconstruct, during the retrieval, some of the natural variability in the cloud–rain partitioning that was diagnosed from the soundings.

The above issues have been addressed to some degree by the algorithms proposed by Kummerow and Giglio (1994) and Bauer and Schluessel (1993). The approach presented here builds upon their ideas and expands them further in several ways. The most important one is the development and use of the multichannel rain indicator (its value and spatial variability) as a measure of both rain intensity and inhomogeneity within the satellite's field of view. This multichannel approach is robust when used in varying environmental conditions and rain intensity regimes.

d) In the absence of substantial rain, the algorithm retrieves a number of parameters describing the atmosphere and the underlying surface. These parameters,

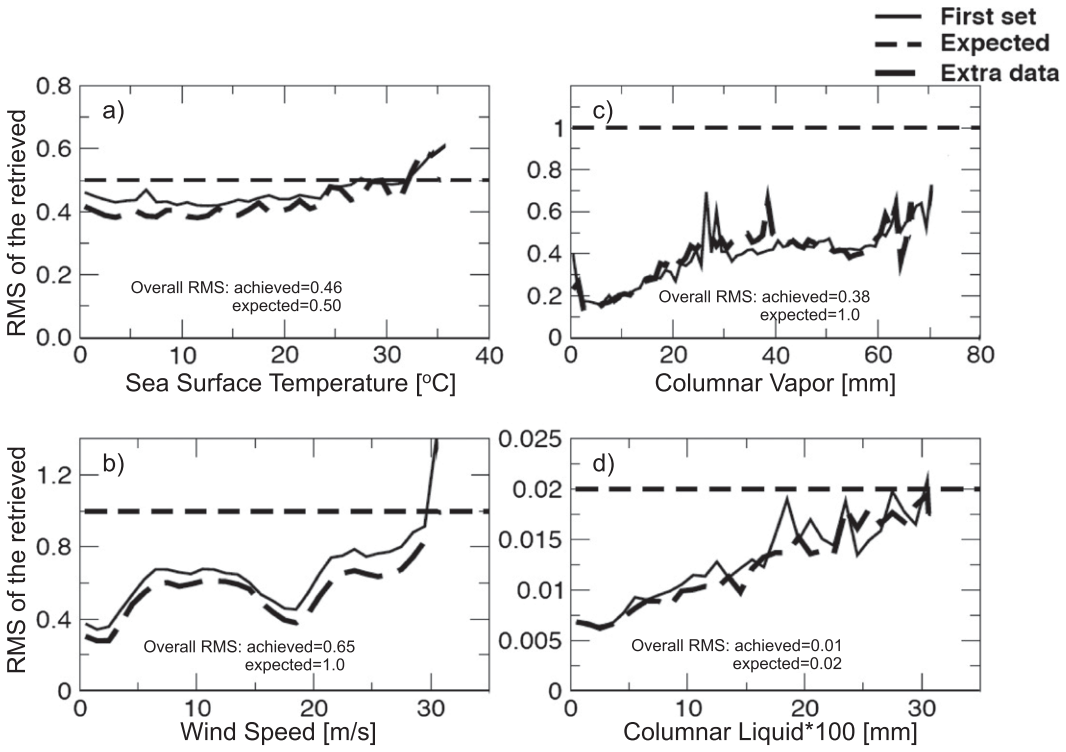

FIG. 12. RMS errors of the retrieved parameters as functions of the magnitude of the retrieved parameters. The statistics apply for nonrainy retrievals only. The statistics were computed by applying the algorithm to two different datasets: the first set, on which the coefficients were developed (solid line), and the second set, constructed from 20 000 soundings collected over a 3-yr period 20 years earlier and in 20 completely different island locations (long-dashed line). The expected RMS is shown as straight dashed lines.

while valuable on their own, carry also a benefit to the rain retrieval itself: they provide information about the environmental conditions, thus helping further decrease the uncertainty in the rain estimates.

These four unique features address some of the still outstanding problems associated with cloud and precipitation retrievals identified in a review paper by Stephens and Kummerow (2007).

4. Performance of the algorithm on simulated data

As discussed, the algorithm's retrieval databases were constructed using a large number of atmospheric soundings collected in 38 different oceanic stations between 1998 and 2001 (red marks in Fig. 1). To test the representativeness of the sounding database, the developed retrieval coefficients were used to perform retrievals on a totally independent database constructed from 20 000 soundings collected from 20 completely different oceanic sites and over a 3-yr period (1978–82) 20 years earlier (blue marks in Fig. 1). The RMS error for the independent database was essentially the same as that when the retrieval was performed on the original retrieval database. This is true for both nonrainy and rainy conditions (Figs. 12 and 13). This assures that the atmospheric structures that were used to build that retrieval database are, indeed, very representative. It also illustrates the robustness of the algorithm.

In nonrainy conditions (Fig. 12) the overall RMS error that was achieved for each of the four parameters was smaller or equal to the one stated as desired in the Algorithm Theoretical Basis Document (ATBD; Wentz and Meissner 2000) and plotted with short-dashed straight lines.

As discussed in section 2a(1) our retrieval database was constructed by placing each of the atmospheric profiles (from soundings) over an ocean surface with SSTs spanning the range between −2° and 34°C. Our decision not to incorporate expected correlations between surface and atmospheric parameters was motivated by the desire to allow for the retrieval of the usual and the unusual with no a priori preferential probability. However, such a choice poses the question of whether

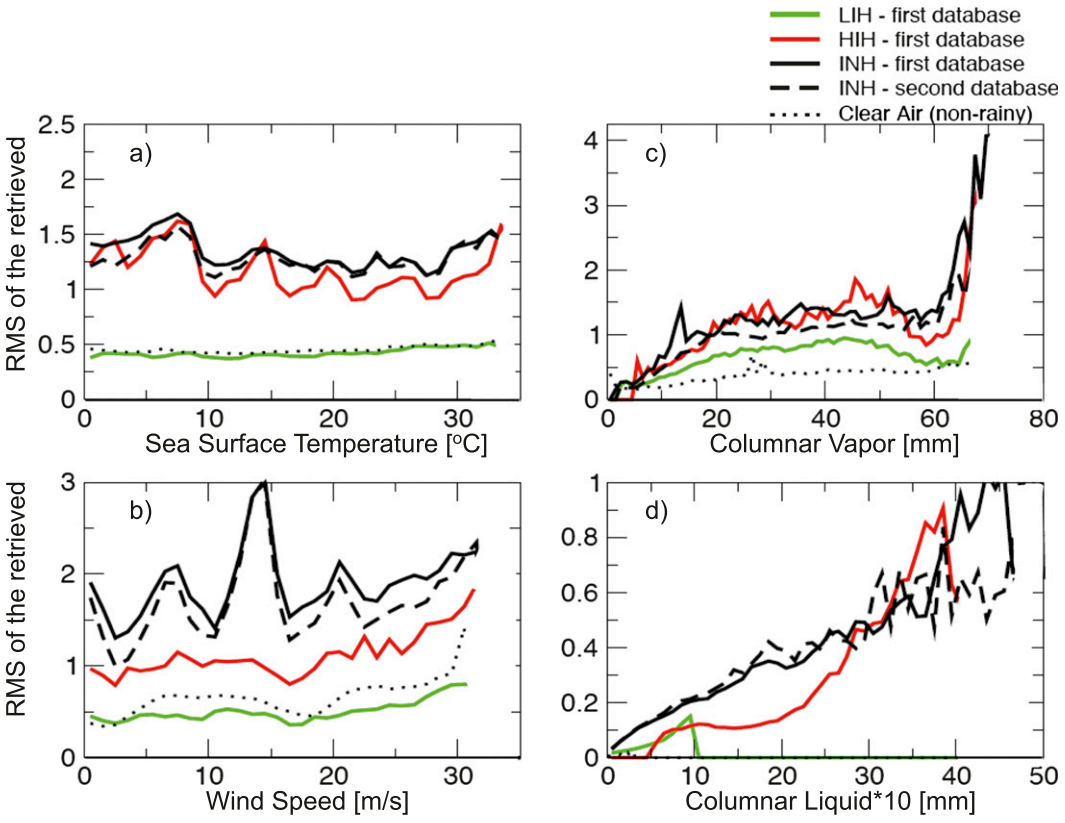

FIG. 13. As in Fig. 12, except the statistics are for rainy conditions and for the three different inhomogeneity regimes (INH–black solid; HIH–red; LIH–green). Black long-dashed lines show the statistics from the INH class retrievals when applied to the second dataset (referred to in Fig. 12). Clear air statistics (black short-dashed lines) are included for comparison. See the text for explanation about the three inhomogeneity regimes.

our retrievals might suffer from the inclusion of some unrealistic cases. To remedy the possibility for erroneous retrievals, our algorithm uses the stepwise and iterative approaches [described in sections 2a(1) and 2a (2)] in which we first retrieve SST and wind using information from the lower frequency channels that are mostly sensitive to the surface conditions. We then use this information to specify the surface subregimes (conditions) under which the vapor and rain retrievals should proceed, using regression coefficients that are derived and valid in these subregimes. Our analysis in Fig. 12 illustrates the ability of our algorithm to accurately retrieve SST and columnar vapor, across their spectrum. This performance comes regardless of the fact that our database contains such extreme conditions and points that the algorithm can effectively separate the surface impact from that of the atmosphere.

Figure 13 illustrates several points that are specific to the retrievals in rain: 1) the RMS error of the retrieved parameters is significantly different for each of the three inhomogeneity regimes; 2) the uncertainty in the LIH case retrievals is very similar to that in nonrainy conditions, something that should be expected; and 3) the uncertainty of the retrievals in the other cases (HIH and INH) is much higher. All of this points to the clear benefit of using (whenever possible) either of the two ''homogeneous'' regimes (LIH or HIH) as compared with using the general INH regime retrieval database (see the RMS error of the retrieved wind speed and columnar liquid, in particular).

5. Summary

Scatterometer wind retrievals provide very valuable information to a variety of climate studies and help improve weather forecast accuracy when assimilated into numerical models. Unfortunately, scatterometer winds are significantly degraded when rain is present in the sensor's field of view. Correcting the scatterometer wind for rain is a very challenging problem that requires very accurate estimates of not only the vertically integrated precipitation amounts but also the surface rain rate.

Even today accurate rain estimates from space remain a challenge. In spite of the significant progress made

after TRMM's launch, there still are significant differences in the estimates made from different algorithms and instruments. The source of the uncertainty lies in the accuracy of the assumptions that go into the development of the retrieval databases and algorithms.

We subscribe to the idea that progress in the accuracy of rain estimates from space will be achieved if the multitude of observations that exist today are utilized to build a number of retrieval databases, each capturing a mode of the observed natural variability of rain. The next step in improving the accuracy of the precipitation retrievals is designing algorithms that have the ability to properly identify the micro and macro characteristics of the observed scene and to select the retrieval database that is most appropriate for that scene.

We have developed a new passive microwave rain retrieval algorithm that is built around these ideas. The algorithm was developed to provide atmospheric correction for scatterometer wind retrieval from the SeaWinds scatterometer onboard the short-lived *ADEOS-II* mission. This algorithm was designed to use Advanced Microwave Scanning Radiometer (AMSR) observations but can be easily tailored to any other conically scanning microwave radiometer with similar set of frequencies (e.g., the upcoming GMI).

The uniqueness of our approach can be summarized as follows. 1) It uses a globally representative set of soundings to develop the database. The temperature and humidity structure provided by the soundings are used to diagnose the amount of water in the atmosphere that can be condensed to form precipitation. This approach helps us avoid making assumptions about the geographically varying atmospheric temperature, vapor, total liquid amounts, height of the freezing level, cloud depth, and cloud-base height, thus assuring proper representation of their natural variability. In addition, there is a good representation of the sea surface variability. 2) It treats the spatial inhomogeneity of rain in a new way, by developing multiple retrieval databases with different built-in inhomogeneity and rain intensity and developing a mechanism for selecting the most appropriate one for each observed scene based on the value and the spatial variability of the specially designed rain indicator (Fig. 6). 3) The cloud-versus-rain partitioning that is part of the rain rate retrievals is unique. It allows for the use of a variety of DSD assumptions—possibly a different one for each of the identified inhomogeneity classes. As our analyses show, this approach allows us to reconstruct, during the retrieval, some of the natural variability in the cloud/rain partitioning that was diagnosed from the soundings (Fig. 9). 4) The algorithm simultaneously retrieves a set of parameters that describe the atmosphere and the underlying surface.

These four unique features address some of the still outstanding problems associated with cloud and precipitation retrievals identified in a review paper by Stephens and Kummerow (2007), namely 1) the need to increase the realism of the atmospheric models that are used, 2) the issues related to discrimination of precipitating cloudy scenes from nonprecipitating scenes, and 3) the need to better account for the variability of the atmospheric and surface states in the construction of the databases and during the retrieval.

We evaluated the performance of the retrieval algorithm by applying it to a totally independent simulated database (constructed from 20 000 soundings that were collected from 20 completely different oceanic sites and over a 3-yr period that was 20 years earlier than the soundings used to derive the retrieval coefficients). The RMS error for the independent database was essentially the same as that when the retrieval was performed on the original retrieval database. These results assure that the atmospheric structures that were used to build that retrieval database are, indeed, very representative. Furthermore, the analyses (Fig. 12) illustrate the ability of our algorithm to accurately retrieve SST and columnar vapor, across their spectrum. This performance comes regardless of the fact that our database contains extreme conditions and points that the algorithm can effectively separate the surface impact from that of the atmosphere, illustrating its robustness.

In Part II of this study (S. Hristova-Veleva et al. 2013a, unpublished manuscript) we will evaluate and validate the performance of our algorithm when applied to AMSR observations. Comparison with NEXRAD estimates and other satellite retrievals of rain validates our approach and illustrates its strengths. Furthermore, as will be illustrated in Part II, applying the estimated atmospheric correction to the scatterometer observations has resulted in significant improvement of the scatterometer winds in rain. These results demonstrate that our approach to passive microwave rain retrievals from space has some very important benefits. However, we should point out that there are areas that need further improvement. The most notable ones include the need to improve the accuracy of the assumptions about the vertical hydrometeor structures and particle size distributions. High-resolution vertical reflectivity profiles (from TRMM in the tropics and from *CloudSat*, airborne, and upcoming GPM observations in the higher latitudes) should help define the characteristic precipitation profiles. Multifrequency wind profiler data can be used to deduce the variability in the observed DSD parameters and the cloud–rain partitioning. Information from such sources should be used to design a greater number of retrieval

databases sufficient to capture all of the main modes of natural rain variability.

Acknowledgments. The research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. We thank Dr. Ziad Haddad, of JPL, for the numerous discussions during the algorithm development and for his greatly appreciated help in modifying the paper to enhance the clarity. We also thank three anonymous reviewers whose comments and insightful suggestions greatly helped in improving this paper.

REFERENCES

Ashcroft, P., and F. J. Wentz, 2000: Algorithm Theoretical Basis Document AMSR level 2A algorithm. RSS Tech. Rep. 121599B-1, 27 pp.

Battan, L. J., 1973: *Radar Observation of the Atmosphere.* Tech-Books, 324 pp.

Bauer, P., and P. Schluessel, 1993: Rainfall, total water, ice water and water vapor over sea from polarized microwave simulations and Special Sensor Microwave/Imager data. *J. Geophys. Res.,* **98** (D11), 20 737–20 759.

Biggerstaff, M. I., E.-K. Seo, S. M. Hristova-Veleva, and K.-Y. Kim, 2006: Impact of cloud model microphysics on passive microwave retrievals of cloud properties. Part I: Model comparison using EOF analyses. *J. Appl. Meteor. Climatol.,* **45,** 930–954.

Chavas, D., and K. Emanuel, 2010: A QuikSCAT climatology of tropical cyclone size. *Geophys. Res. Lett.,* **37,** L18816, doi:10.1029/2010GL044558.

Chelton, D. B., M. G. Schlax, M. H. Freilich, and R. F. Milliff, 2004: Satellite measurements reveal persistent small-scale features in ocean winds. *Science,* **303,** 978–983.

Cunningham, S. A., and Coauthors, 2007: Temporal variability of the Atlantic meridional overturning circulation at 26.5°N. *Science,* **317,** 935–938, doi:10.1126/science.1141304.

Grody, N. C., 1991: Classification of snow cover and precipitation using the Special Sensor Microwave Imager. *J. Geophys. Res.,* **96,** 7423–7435.

Haddad, Z. S., D. A. Short, S. L. Durden, E. Im, S. Hensley, M. B. Grable, and R. A. Black, 1997: A new parameterization of the rain drop size distribution. *IEEE Trans. Geosci. Remote Sens.,* **35,** 532–539.

Hristova-Veleva, S. M., 2000: Impact of microphysical parameterizations on simulated storm evolution and remotely-sensed characteristics. Ph.D. thesis, Texas A&M University, 201 pp.

Huddleston, J. N., and B. W. Stiles, 2000: A multi-dimensional histogram rain flagging technique for SeaWinds on QuikSCAT. *Proc. IGARSS 2000,* Honolulu, HI, IEEE, Vol. 3, 1232–1234.

Jiang, C., L.-A. Thompson, and N. D. K. A. Kelly, 2008: Equatorial influence of QuikSCAT winds in an isopycnal ocean model compared to NCEP2 winds. *Ocean Modell.,* **24,** 65–71.

Keihm, S. J., M. A. Janssen, and C. Ruf, 1995: TOPEX/Poseidon Microwave Radiometer (TMR): III. Wet troposphere range correction algorithm and pre-launch error budget. *IEEE Trans. Geosci. Remote Sens.,* **33,** 147–161.

Kidd, C., 1998: On rainfall retrieval using polarization-corrected temperatures. *Int. J. Remote Sens.,* **19,** 981–996.

Kummerow, C., 1993: On the accuracy of the Eddington approximation for radiative transfer in the microwave frequencies. *J. Geophys. Res.,* **98,** 2757–2765.

——, and L. Giglio, 1994: A passive microwave technique for estimating rainfall and vertical structure information from space. Part I: Algorithm description. *J. Appl. Meteor.,* **33,** 3–18.

——, I. M. Hakkarinen, H. F. Pierce, and J. A. Weinman, 1991: Determination of precipitation profiles from airborne passive microwave radiometric measurements. *J. Atmos. Oceanic Technol.,* **8,** 148–158.

——, W. S. Olson, and L. Giglio, 1996: A simplified scheme for obtaining precipitation and vertical hydrometeor profiles from passive microwave sensors. *IEEE Trans. Geosci. Remote Sens.,* **34,** 1213–1232.

——, and Coauthors, 2001: The evolution of the Goddard profiling algorithm (GPROF) for rainfall estimation from passive microwave sensors. *J. Appl. Meteor.,* **40,** 1801–1820.

——, P. Poyner, W. Berg, and J. Thomas-Stahle, 2004: The effects of rainfall inhomogeneity on climate variability of rainfall estimated from passive microwave sensors. *J. Atmos. Oceanic Technol.,* **21,** 624–638.

Liu, G., J. Curry, and R.-S. Sheu, 1995: Classification of clouds over the western equatorial Pacific Ocean using combined infrared and microwave satellite data. *J. Geophys. Res.,* **100** (D7), 13 811–13 826.

Lovenduski, N. S., and N. Gruber, 2005: Impact of the southern annular mode on Southern Ocean circulation and biology. *Geophys. Res. Lett.,* **32,** L11603, doi:10.1029/2005GL022727.

Mie, G., 1908: Beiträge zur Optik trüber Medien speziell kolloidaler Metallösungen (Contributions to the optics of turbid media, particularly of colloidal metal solutions). *Ann. Phys.,* **330,** 377–445.

Minobe, S., A. Kuwano-Yoshida, N. Komori, S.-P. Xie, and R. J. Small, 2008: Influence of the Gulf Stream on the troposphere. *Nature,* **452,** 206–209.

Mugnai, A., and E. A. Smith, 1988: Radiative transfer to space through a precipitating cloud at multiple microwave frequencies. Part I: Model description. *J. Appl. Meteor.,* **27,** 1055–1073.

Olson, W. S., C. D. Kummerow, Y. Hong, and W.-K. Tao, 1999: Atmospheric latent heating distributions in the tropics derived from satellite passive microwave radiometer measurements. *J. Appl. Meteor.,* **38,** 633–664.

O'Neill, L. W., D. B. Chelton, S. K. Esbensen, and F. J. Wentz, 2005: High-resolution satellite measurements of the atmospheric boundary layer response to SST perturbations over the Agulhas Return Current. *J. Climate,* **18,** 2706–2723.

Panegrossi, G., and Coauthors, 1998: Use of cloud model microphysics for passive microwave-based precipitation retrieval: Significance of consistency between model and measurement manifolds. *J. Atmos. Sci.,* **55,** 1644–1673.

Petty, G. W., 1994a: Physical retrievals of over-ocean rain rate from multichannel microwave imagery. Part I: Theoretical characteristics of normalized polarization and scattering indices. *Meteor. Atmos. Phys.,* **54,** 79–99.

——, 1994b: Physical retrievals of over-ocean rain rate from multichannel microwave imagery. Part II: Algorithm implementation. *Meteor. Atmos. Phys.,* **54,** 101–121.

——, and K. B. Katsaros, 1992: *Nimbus-7* SMMR precipitation observations calibrated against surface radar during TAMEX. *J. Appl. Meteor.,* **31,** 489–505.

Rutledge, S. A., and P. V. Hobbs, 1983: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. VIII: A model for the ''seeder–feeder'' process in warm-frontal rainbands. *J. Atmos. Sci.,* **40,** 1185–1206.

Simpson, J. R., and W.-K. Tao, 1993: Goddard Cumulus Ensemble Model. Part II: Applications for studying cloud precipitating processes and for NASA TRMM. *Terr. Atmos. Ocean. Sci.,* **4,** 73–116.

Smith, E. A., and A. Mugnai, 1988: Radiative transfer to space through a precipitating cloud at multiple microwave frequencies. Part II: Results and analysis. *J. Appl. Meteor.,* **27,** 1074–1091.

——, ——, H. J. Cooper, G. J. Tripoli, and X. Xiang, 1992: Foundations for statistical–physical precipitation retrieval from passive microwave satellite measurements. Part I: Brightness-temperature properties of a time-dependent cloud-radiation model. *J. Appl. Meteor.,* **31,** 506–531.

——, and Coauthors, 1998: Results of WetNet PIP-2 project. *J. Atmos. Sci.,* **55,** 14 813–1536.

Spencer, R. W., H. M. Goodman, and R. E. Hood, 1989: Precipitation retrieval over land and ocean with SSM/I: Identification and characteristics of the scattering signal. *J. Atmos. Oceanic Technol.,* **6,** 254–273.

Stephens, G. L., and C. D. Kummerow, 2007: The remote sensing of clouds and precipitation from space: A review. *J. Atmos. Sci.,* **64,** 3742–3765.

Tao, W.-K., J. Simpson, S. Lang, M. McCumber, R. Adler, and R. Penc, 1990: An algorithm to estimate the heating budget from vertical hydrometeor profiles. *J. Appl. Meteor.,* **29,** 1232–1244.

Weinman, J. A., and R. Davies, 1978: Thermal microwave radiances from horizontally finite clouds of hydrometeors. *J. Geophys. Res.,* **83,** 3099–3107.

Weng, F., and N. C. Grody, 1994: Retrieval of cloud liquid water using the Special Sensor Microwave Imager (SSM/I). *J. Geophys. Res.,* **99,** 25 535–25 551.

Wentz, F. J., 1990: SBIR phase II report: West coast storm forecasting with SSM/I. RSS Tech. Rep. 033190, 378 pp. [Available from Remote Sensing Systems, 1101 College Avenue, Suite 220, Santa Rosa, CA 95404.]

——, and R. W. Spencer, 1998: SSM/I rain retrievals within a unified all-weather ocean algorithm. *J. Atmos. Sci.,* **55,** 1613–1627.

——, and T. Meissner, 2000: Algorithm Theoretical Basis Document: AMSR Ocean Algorithm. RSS Tech. Proposal 121599A-1, 44 pp.

Wilheit, T. T., A. T. C. Chang, M. S. V. Rao, E. B. Rodgers, and J. S. Theon, 1977: A satellite technique for quantitatively mapping rainfall rate over the oceans. *J. Appl. Meteor.,* **16,** 551–560.