Parametric Rainfall Retrieval Algorithms for Passive Microwave Radiometers

DONG-BIN SHIN
Center for Earth Observing and Space Research/School of Computational Sciences, George Mason University, Fairfax, Virginia

CHRISTIAN KUMMEROW
Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado

(Manuscript received 19 July 2002; in final form 4 May 2003)

ABSTRACT

A methodology is described to construct fully parametric rainfall retrieval algorithms for a variety of passive microwave sensors that exist today and are planned for the future. The Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI) is used to retrieve nonraining geophysical parameters. The method then blends these background geophysical parameters with three-dimensional precipitation fields obtained by matching the TRMM precipitation radar (PR) reflectivity profiles with cloud-resolving model simulations to produce a consistent three-dimensional atmospheric description. Based upon this common description, radiative transfer simulations corresponding to specific microwave sensors are then employed to compute radiances from clear and rainy scenes, as might be seen by any specified microwave radiometer. Last, a Bayesian retrieval methodology is used in conjunction with this database to derive the most likely surface rainfall as well as its vertical structure. The algorithm performance is tested for a variety of sensor designs using synthetic retrievals to demonstrate its capability for consistent rainfall estimates. Whereas actual retrievals would be sensitive to the details of the a priori database construction, results from this study indicate that even modest radiometers can retrieve unbiased rainfall rates when constrained by an a priori database constructed from the TRMM satellite. Random errors are correlated to unobserved variations in the vertical and horizontal structure of the precipitation and, thus, depend upon sensor design specifications. The fidelity of these synthetic retrievals is briefly examined by comparing the simulated brightness temperature ($T_b$) generated in this study with direct observations by the TRMM TMI. Good physical consistency between the simulated and TRMM observed $T_b$s is found in precipitating regions for frequencies at which emission processes dominate the radiometric signal. The consistency is poor for higher-frequency microwave channels for which ice scattering is important. Greater consistency between the computed and observed $T_b$s should be sought before replacing current operational algorithms with the parametric equivalent.

1. Introduction

The evolution of rainfall retrievals from passive microwave sensors closely followed the development and improvement of satellite sensors, starting with the Electronically Scanning Microwave Radiometer (ESMR) launched on Nimbus-5 in December 1972. It was a single channel, 19-GHz radiometer. Allison et al. (1974) used data from the Nimbus-5 ESMR to map rainfall areas in a variety of tropical disturbances. Nimbus-5 was followed by Nimbus-6 in June of 1975, carrying an ESMR with a 37-GHz channel. Concurrently, more quantitative approaches were developed by Wilheit et al. (1977), Weinman and Guetter (1977), Rodgers et al. (1979), and Jung (1980). These were designed to estimate a single rainfall parameter from a single spectral measurement through idealized brightness temperature–rain-rate relationships. Once these algorithms achieved a certain maturity, however, it was clear that more than one frequency was needed in order to properly retrieve the column water amounts. The first multichannel sensor, the Scanning Multichannel Microwave Radiometer (SMMR) on Nimbus-7, was launched in October 1978 and included frequencies ranging from 6.6 to 37 GHz, with spatial resolutions ranging from 136 km $\times$ 89 km at 6.6 GHz to 22 km $\times$ 18 km at 37 GHz. SMMR measured both the horizontal and vertical polarizations. Spencer et al. (1983a) regressed SMMR multichannel radiances against radar data over the Gulf of Mexico with some success. He was also able to use the brightness temperature ($T_b$) depression caused by ice scattering at 37 GHz to develop an empirical algorithm over land surfaces (Spencer et al. 1983b). A physically based rainfall retrieval, based upon detailed modeling of the
sensor response to precipitation profiles, was developed by Olson (1989).

Following SMMR, the Defense Meteorological Satellite Program (DMSP) carried the Special Sensor Microwave Imager (SSM/I), which was first flown on DMSP-8 in August 1987. Much like the SMMR, the SSM/I is a conically scanning radiometer with seven channels, ranging from 19.3 to 85 GHz, and spatial resolutions ranging from 19.3 km × 43 km to 15 km × 13 km, respectively. The excellent calibration of the SSM/I, coupled with the data continuity provided by DMSP from 1987 to the present, lead to much attention paid to the improvement of microwave rainfall estimates. Three classes of algorithms emerged: (a) the “emission-type” algorithms, following Wilheit’s early work with SMMR (e.g., Wilheit et al. 1991; Berg and Chase 1992), (b) the “scattering” algorithms, following Spencer’s early work with SMMR over land (e.g., Grody 1991; Ferraro and Marks 1995), and (c) the “multichannel inversion”-type algorithms that expanded upon Olson’s earlier work (e.g., Mugnai et al. 1993; Kummerow and Giglio 1994; Smith et al. 1994; Petty 1994).

The Tropical Rainfall Measuring Mission (TRMM) was launched in November 1997. It carries the TRMM Microwave Imager (TMI), which is based upon SSM/I technology, but with additional 10.7-GHz horizontal polarization (H) and vertical polarization (V) channels to expand the dynamic range of emission signals in the Tropics. In addition to these channels, the TRMM satellite flies at a much lower altitude than the DMSP satellites (350 versus 833 km). This provides significantly improved spatial resolution for TMI (30 km × 18 km at 19.3 GHz) in comparison with what was previously available. TRMM rainfall retrieval algorithms further expanded upon the emission methods (Chang et al. 1999), as well as the multichannel inversion-type algorithm (Kummerow et al. 2001). The Advanced Microwave Sensing Radiometer (AMSR) was developed by the National Space Development Agency (NASDA) of Japan and launched aboard the National Aeronautics and Space Administration (NASA) Aqua satellite in May 2002. A second AMSR instrument is scheduled for launch aboard NASA’s Advanced Earth-Observing Satellite (ADEOS)-II mission in late 2002. The AMSR instruments have additional 6.6-GHz H and V channels and a much larger antenna to obtain TMI-like resolution, but from ~800-km orbit. The operational algorithms selected by NASA are modifications of those used with TRMM, while the Japanese AMSR will use a multichannel inversion algorithm that is based upon the algorithm described by Petty (1994).

Because of the continuous radiometer improvements with respect to the number of channels, their spatial resolution, and their absolute calibration, algorithm development has naturally focused only upon the latest sensors. Relatively little effort has been devoted to finding generalized algorithms, and it is, perhaps, not surprising that most of the current passive microwave algorithms contain empirically adjusted parameters or procedures that depend on specific channels or channel combinations to optimize the retrievals. Examples of empirical parameters include rain/no-rain discrimination, particularly over land, and various forms of rain-type classifications that invariably make use of specific channel combinations. Gradients of $T_\delta$ in the high-resolution channels, for instance, are useful to infer the type for rainfall observed by the satellite. Unfortunately, gradients are very sensitive to the spatial resolution of the sensor and, thus, are not easily transferable from one sensor to another. Channel-dependent parameters include the determination of the freezing level (Wilheit et al. 1977), which requires a 19/22-GHz combination, or determination of the background state over ocean that typically requires a 19/21/37-GHz combination of channels.

Efforts to produce homogeneous rainfall products for a climatological time series are hampered by satellite-specific assumptions and procedures. In the future, NASA and NASDA, in collaboration with a number of other space agencies, plan to launch the Global Precipitation Mission (GPM). This mission consists of a core satellite carrying a state-of-the-art dual-frequency precipitation radar, as well as a passive microwave radiometer. In addition, the GPM concept uses a constellation of satellites carrying passive microwave radiometers in order to achieve three hourly rainfall samplings. This constellation consists of radiometers on operational satellites such as the current SSM/I series, as well as some that are planned specifically for the GPM. As such, GPM is both a satellite “mission,” as well as a concept designed to combine the many planned international radiometers into a coherent framework. In order to achieve this conceptual benefit, however, it is imperative that algorithms be developed that allow a coherent description of rainfall to emerge from a wide-ranging set of sensor capabilities. Because these sensors will have different characteristics, a reference algorithm is necessary to address issues of consistency in the rainfall products among this diverse suite of sensors.

Such an algorithm is described in section 2. It is based upon the evolving Goddard profiling algorithm (GPROF) (Kummerow et al. 2001) concept but takes a step backward from the latest operational algorithm to reformulate all procedures that are sensor dependent. In particular, it uses the concept of a “best” satellite, TRMM in this case, to construct a database of three-dimensional a priori profiles of raining and nonraining conditions. The algorithm is constructed in such a way that all assumptions are contained in this single a priori database. While this does not necessarily represent an improvement over any particular retrieval method, it does ensure that the assumptions are independent of the specific sensor eventually used by the retrieval. Unlike current algorithms, one can have confidence that systematic differences between platforms must be related to the shortcoming in the a priori database instead of
2. The algorithm

The algorithm described here is designed to create three-dimensional geophysical parameter fields that are consistent with the TMI when no rain is present, and with both the TMI and the TRMM precipitation radar (PR) observations where the PR detects rainfall. The latter condition is not always satisfied with the available measurements, but this is an area of active research and no attempt is made here to reconcile TMI and PR observations in a physical manner. Instead, the differences are quantified, and a solution is outlined to handle less-than-ideal input data. One can further assume that the GPM, with its dual-frequency precipitation radar, will provide additional radiometric information that will lead to improved consistency between sensors in the future.

Simulations of the three-dimensional precipitation field begin with the TMI to retrieve geophysical parameters in the nonraining portion of the scene. The nonraining retrieval is needed to provide a coherent scene description independent of sensor resolution or ability to discriminate raining from nonraining conditions. Over raining fields of view (FOV), PR-observed reflectivities are matched to reflectivities computed from the cloud-resolving model (CRM) simulations. The present study uses simulations from both the Goddard Cumulus Ensemble (GCE) model (Tao and Simpson 1993), as well as the University of Wisconsin (UW) Nonhydrostatic Modeling System (NMS) (Tripoli et al. 1992b). These are described later. The CRMs provide a physically consistent set of fully three-dimensional hydrometeor and latent heating profiles. The simulated precipitation and background fields are then used to calculate the brightness temperatures for any passive microwave sensor design, and Bayesian methodology is used to retrieve rainfall for different sensors. One month of TRMM data over two different regions in the east Pacific Ocean and the Amazon basin are used here as ocean and land backgrounds to construct the database of rainfall profiles and the corresponding $T_b$ for the different sensors considered in this study. Figure 1 shows a schematic representation of the database construction.

a. The nonraining simulations

Where it is not raining, passive microwave $T_b$ depends upon column-integrated water vapor (WV), cloud liquid water (CLW), and the surface temperature and emissivity. Over ocean, the emissivity depends primarily upon the surface wind speed; over land, the emissivity depends upon a variety of factors, including soil type, soil moisture, and vegetation cover. A number of techniques to retrieve these parameters over ocean have been described in the literature. Wilheit and Chang (1980) and Wentz (1997) are examples of studies over the ocean, while Wang and Schmugge (1980) and Njoku (1999) have made some progress over land. Although these results could be used, they are not uniformly available for the TMI sensor.

The nonraining retrieval is, therefore, developed here based upon a Bayesian approach to ensure consistency with the subsequent rainfall retrieval. For the nonraining database, CLW and WV profiles are collected from the cloud-resolving models when the associated surface rain rate is less than 0.1 mm h$^{-1}$. A principal component analysis is used to create various realizations of non-

---

**Figure 1.** A schematic flow diagram to construct the CRM-based database.
raining profiles, as well as to reduce redundant profiles. The spatial (vertical) principal components or empirical orthogonal functions (EOFs) can be obtained by diagonalizing the covariance matrix calculated by subtracting out the spatial mean from the data. The vertical profile \( P(z) \) of CLW or WV can be written as

\[
P(z) = \sum_{n=1}^{n_{\text{layer}}} \alpha_n \phi_n(z),
\]

where the coefficients \( \alpha_n \) are random numbers that are uncorrelated with each other, \( \phi_n(z) \) is the eigenvector associated with mode \( n \), and \( n_{\text{layer}} \) is the number of vertical layers in the model. The eigenvalue of mode \( n \), \( \lambda_n \), is considered as the variance of random variable \( \alpha_n \):

\[
\langle \alpha_n \alpha_m \rangle = \lambda_n \delta_{nm},
\]

where the angular brackets denote the ensemble average or expected value. Moreover, the statistics of \( P(z) \) require zero ensemble average \( \langle \alpha_n \rangle = 0 \). The random numbers \( \alpha_n \) may be generated using various probability distribution functions as long as they satisfy the original statistics. A Gaussian distribution was chosen to represent a variety of the nonraining profiles that deviated smoothly from the original profiles. In this case, \( \alpha_n \) can be generated through a normally distributed probability function with zero mean and variance \( \lambda_n \).

Based on the statistics above, the nonraining profiles are reconstructed at 10 different sea surface temperatures (SST) and 10 different surface winds (W), satisfying the following statistics of the original profiles:

\[
\langle P(z) \rangle = 0 \quad \text{and} \quad \langle P(z)^2 \rangle = \sum_{n=1}^{n_{\text{layer}}} \lambda_n \langle \phi_n(z)^2 \rangle.
\]

Brightness temperatures for all combinations of SST, CLW, WV, and W are calculated using a one-dimensional plane-parallel Eddington model (Kummerow 1993). The profiles and their associated brightness temperatures establish the nonraining a priori database in the Bayesian retrieval. The Bayesian retrieval methodology is then applied to retrieve the above parameters for all TMI pixels determined as nonraining by the PR. A qualitative comparison with the parameters derived by Wentz (1997) indicated that the current method produces reasonable results needed in this study. Better results may be obtained in the future from a more careful treatment of this problem. Where the PR sees precipitation, the procedure uses the PR reflectivities to determine the rainfall parameters. The only complication results from the fact that the PR footprints are smaller than those of the TMI. If a portion of the TMI footprint is covered by rainfall, then nonraining retrievals cannot be performed over that pixel. The PR, however, will not completely fill the pixel with raining structures. This results in a small area of missing data that must be obtained by interpolating the nonraining TMI geophysical parameters into the partially raining pixels.

Figure 2 shows this procedure graphically for the water vapor field. Other fields are treated in the same manner. Figure 2a represents the TMI retrieval, Fig. 2b shows the interpolated field, and Fig. 2c is the final background field in which geophysical parameters have been mapped to PR pixel locations and pixels determined to be raining by the PR have been set to missing. Uncertainties introduced by the nonraining retrieval and the interpolation scheme can be quantified by a few high-resolution measurements of oceanic precipitation.

Over land the procedure is implemented in the same fashion, but emissivity and surface temperature are allowed to vary over reasonable values, while the water vapor and cloud water are kept constant because of the fact that they offer little contrast to the underlying surface. While this procedure is simple, it does not take into account any changes in the surface emissivity due to varying soil characteristics or vegetation cover. Nor does the interpolation scheme account for the possibility that surface emissivities can change quickly, depending upon the antecedent rainfall. Because these topics are all the subject of current research, this study is limited to a region in the Amazon where the background conditions are relatively homogeneous.

b. The raining scene

Precipitating areas are easily identified by the PR. If a rain signal is detected, the rain profile that best fits the PR reflectivity profile is selected from a set of precomputed CRM simulations. Only those profiles that correspond to CRM simulations with the observed surface temperature are considered. The reflectivity of the cloud model profiles was obtained by computing single particle backscattering properties based upon Mie theory and assuming the same drop size distribution (DSD) used in the CRMs. The models used in this study are the GCE and the UW NMS. A description of the GCE model can be found in Tao and Simpson (1993). The cloud microphysics includes a two-category liquid water scheme (cloud water and rain), and a three-category ice phase scheme (cloud ice, snow, and graupel). The distributions of rain, snow, and graupel are taken to be inverse exponentials of the form prescribed by Marshall and Palmer (1948), but with an intercept dictated by the physical parameterizations. The horizontal resolution of these models varies from 1 to 2 km, while the vertical coordinate is 500 m in the lower troposphere.

The second model used is the UW NMS described by Tripoli (1992b). Aside from differences in the dynamical assumptions in the model, the UW NMS considers four classes of ice: graupel, pristine crystals, snow crystals, and aggregates that need to be treated differently from the GCE profiles. A detailed description of these ice categories and their interaction may be found in Tripoli (1992a). Horizontal and vertical coordinates are similar to those of the GCE simulations.

Given the hydrometeor profiles from the cloud-resolving models, the radar reflectivity factor at each level
for a nadir-viewing radar at the PR frequency (13.8 GHz) is computed from (Meneghini and Kozu 1990)

\[
Z = \frac{\lambda^4}{\pi^3 |K|^2} \int_0^D \sigma_r(D) N(D) \ dD,
\]

where \(\lambda\) is wavelength, \(K = (m^2 - 1)/(m^2 + 1)\) is the dielectric factor of the scattering particle, \(m\) is the complex index of refraction, \(\sigma_r(D)\) is the backscattering cross section, as a function of the particle’s effective diameter, and \(N(D)\) is the number density of particles. These reflectivities were compared with the attenuation-corrected reflectivities available from the PR. In this idealized study, the profiles with the smallest root-mean-square difference between observed and CRM-computed reflectivity is selected to represent that PR pixel. Representative examples of the matched reflectivity profiles and associated hydrometeor profiles from the cloud model simulations are shown in Fig. 3. Differences in the observed and matched reflectivity profiles are due to the discrete number of CRM profiles. Because the horizontal scale is in dBZ rather than \(Z\), however, mismatches for small dBZ values are of little consequence in the hydrometeor profiles.

Figure 4 also describes the overall agreement between the observed and matched (simulated) reflectivity profiles for a database over the east Pacific. The details of the database will be presented in the following section. In Fig. 4, the observed reflectivity profiles are separated at 5-dBZ bins based on surface reflectivity and then convective and stratiform precipitation. For each interval, the averages for the observed and simulated reflectivity profiles are obtained (thick dotted and solid lines, respectively). The mean deviation between the observed and simulated profiles is computed at each layer (thin dotted line). The mean observed deviation from the average observed profile, which may show the natural variability of the profiles, is also presented for the reference of the matching skill (thin solid line). For
every interval of stratiform precipitation, the computed reflectivity profiles seem to match well with the observed profile below the melting layer between 4 and \( 4.5 \) km. Despite the fact that the mean profiles are quite similar to each other, the mean deviations between the computed and observed profiles are smaller than the natural variability of observations below the layer. Above the layer, the matching accuracy tends to be degraded. The matching for convective precipitation shows similar results. This shortcoming in the matching procedure will be discussed in section 4.

The above-mentioned procedure introduces a potential error in that the PR attenuation-corrected reflectivity makes certain DSD assumptions that are potentially inconsistent with the DSD assumptions in the CRM. These differences are minor and are thought to be acceptable for a simulation study. For an operational algorithm, the procedure should probably be replaced by one in which the actual PR observations, including the total path attenuation (when deemed reliable) are matched. For this method to incorporate a quantitative error estimate, the deterministic procedure employed here must further be replaced by a probabilistic one, where a family of profiles that are consistent with the radar observations is derived. In addition, the radiometer itself could be used to eliminate the potential profiles that are inconsistent with the radiometric measurements, thus, ensuring that the radar-only, radiometer-only, and radar–radiometer algorithms all start with the same a priori database. For this simplified study, the profiles selected from the cloud-resolving model, and their associated surface rainfall rate, are assumed to be the truth. Brightness temperatures calculated from the raining scene turn out to be reasonably close to the actual observations, providing confidence in the procedure if all the uncertainties in the matching procedure are properly accounted for.

c. The databases

The above procedure, once repeated for sufficient satellite orbits [small sections of the east Pacific (\( 5^\circ - 15^\circ \)N,
Fig. 4. Mean observed and simulated reflectivity profiles at the different surface reflectivity intervals and precipitation types [(top) stratiform, (bottom) convective] for a database over the east Pacific, indicated by thick dotted and solid lines, respectively. The mean deviations between the observed and simulated profiles are represented by thin dotted lines. The mean observed deviations from the mean observed profiles are also indicated by thin solid lines.

150°–120°W) for ocean background, and the Amazon (20°–5°S, 70°–40°W) for land background during December 1999 were used in this study], creates a three-dimensional distribution of hydrometeor profiles with 4-km resolution that are consistent with TMI-observed \( T_b \) in the nonraining areas, and consistent with the PR reflectivities when rain is present. Radiative transfer computations were applied at this 4-km resolution using a plane-parallel Eddington approximation (Kummerow 1993). While three-dimensional radiative transfer solutions are available, it was felt that the extra computational resources were not warranted for this initial study, given that only synthetic radiances would eventually be considered.

The a priori database for each of the sensors considered in this study is constructed from the above fields. Hydrometeors were averaged for 3 km × 3 km radar fields of view corresponding to 12 km × 12 km hydrometeor profiles. This dimension determines the resolution of the retrieval. Because radiometer sensor ground resolutions increase linearly with frequency, there is no unique resolution at which hydrometeors should be retrieved for each sensor. A fixed 12-km hydrometeor resolution was selected in this study in order to directly compare results from each of the sensors. In practice, different resolutions can be employed for each sensor based upon its characteristics. The corresponding \( T_b \) were computed by assuming a circular aperture and a Gaussian antenna gain function for each of the sensors considered in this study. This produces FOV’s that are within about 1 km of the actual resolutions cited for the existing sensors. The weighting functions were centered on the 12 km × 12 km hydrometeor profile.

The construction of the a priori database is straightforward, particularly in the current study that deals only with synthetic radiances. Real applications must consider the uncertainties discussed in section 3. In particular, we note that the current method for generating the a priori database assumes that the combination of TMI and PR can retrieve the relevant physics needed to compute \( T_b \) at other frequencies. This is not the case for frequencies such as the 60-GHz oxygen absorption band. These frequencies should not be included because they respond primarily to the atmospheric temperature...
structure that is not observed by the TMI and PR combination.

d. The retrieval

Once the database is constructed, a Bayesian retrieval methodology is used to select those profiles that are consistent with the observations. Bayesian methodologies for microwave rain estimation are well covered in the literature (e.g., Kummerow et al. 1996; Olson et al. 1996; Marzano et al. 1999). The Bayesian methodology dictates that channel weights be proportional to a Gaussian function whose width is determined by the uncertainty in the forward model and the radiometric noise. Because some of the sensors are hypothetical, as well as the difficulties associated with quantifying retrieval performance without adequate ground-based observations, only synthetic retrievals are considered here. It is, therefore, assumed that the PR–CRM profile represents the truth, and the only uncertainty is the noise (NEAT) of the radiometer itself. The present retrieval is, thus, particularly simple. Sensor noise is taken to be independent of channel and 0.5 K in the present study. For normally distributed errors, a Gaussian weight is assigned to each profile based upon root-mean-squared difference between observed $T_b$ and that computed for each profile in the a priori database. The width of the Gaussian function is determined from the mean observational error, which is simply the sum of independent 0.5-K channel noise figures.

The only free parameters, and, thus, the nomenclature of parametric retrieval algorithm, are the sensor specifications themselves. The sensor dependence is fully specified by (a) the frequency and resolutions needed to construct the sensor-specific database from the geophysical parameters derived by the core satellite, and (b) the sensor noise characteristics needed by the retrieval itself.

e. Retrieval results

Because of the simplifications needed to remove all instrument dependencies in the current scheme, one might expect a radical departure from the current methodologies that are optimized to specific sensors. However, as indicated in Fig. 5, a comparison of the rainfall estimated from this retrieval and the TRMM operational level-2 algorithms (2A25 for PR and 2A12 for TMI) shows that the simple parametric algorithm developed here has adequate realism relative to the operational algorithms that have been optimized over time for each individual sensor.

The performance of various radiometers can now be tested. In this study, we consider a TMI with a 61-cm dish in a 350-km orbit (hereinafter TMI-350-61), a TMI with a 35-cm dish in a 600-km orbit (hereinafter TMI-600-35), an SSM/I with a 61-cm dish in a 383-km orbit (hereinafter SSM/I-383-61), and the cross-track-scanning constellation microwave radiometer (CMR) being developed for GPM, with a combination of channels at 10, 19, 37, and 85 GHz. The CMR is assumed to have a 100-cm antenna and a 600-km orbit. Because microwave radiometers are diffraction limited, the ground resolution is inversely proportional to the orbit altitude.

1) Ocean background

The performance of the radiometers is evaluated for $T_b$, generated over the east Pacific. As outlined previously, an independent set of orbits is used to generate the synthetic $T_b$ corresponding to each assumed sensor. The PR–CRM-derived rainfall represents the truth in this study. Rain detection statistics, as well as rain bias,
Fig. 6. Two-dimensional histogram (true vs retrieved rain rates) from synthetic rainfall retrievals for the different radiometer specifications with the nadir angle of 45° over ocean background. Color in square box indicates frequency in log scale, as denoted at bottom.

correlation, and rms statistics for each sensor specification are computed. Figure 6 and Table 1 show retrieval performance of TMI, SSM/I, and CMR at different altitudes with the nadir angle of 45°, which resembles the viewing geometry of TMI and SSM/I. Both the table and figure show that high spatial resolutions produce higher-quality retrievals. Higher spatial resolutions reduce the inhomogeneity of rain within a FOV, and thus improve the physical relationship between rainwater content and $T_b$. The effect of using a lower orbit can be clearly seen in the results from two TMI sensors (Figs. 6a,b) or SSM/I sensors (Figs. 6c,e) at different altitudes. The third experimental sensor (SSM/I-600-61) is designed to be identical to the second sensor (TMI-600-61), but without the 10-GHz channels (Fig. 6c). Retrieval statistics of this design do not seem to differ significantly from those of the TMI-600-61 with all channels, except for a slightly larger bias and rms. This result suggests that the ignored low-frequency channels (10 GHz) are not important at this relatively coarse resolution. The reason may be that these low-frequency channels become important primarily when the 19-GHz
channel begins to saturate, which is less likely with the larger FOV. The last four designs are devoted to quantifying the performance of CMR with various combinations of channels. The absorption-oriented design turns out to perform only slightly better than CMR at 10, 19, 37, and 85 GHz (hereinafter CMR-10-19-37-85, Fig. 6f), and actually shows slightly better performance than CMR at 10, 37, and 85 GHz (hereinafter CMR-10-37-38-85) (Fig. 6g). Figure 6i examines a sensor specification by adding high-frequency channels (85 GHz) to CMR-19-37. This sensor design turns out to perform only slightly better than the CMR-19-37. Bias statistics that may be an important factor from a climatological perspective appear all to be very small (~1%) and appear to be in the noise.

2) LAND BACKGROUND

Similar retrievals and results are presented for the land background cases in Fig. 7 and Table 2. As can be seen, correlations are significantly lower, but biases are still relatively small. The decrease in correlation is a reflection of the poor radiometric information that is available over land, where only scattering signals are readily discernable over the radiometrically warm surface. Closer inspection of Fig. 7 also reveals the tendency of the radiometers to overestimate light rain, while underestimating the heavier precipitation. This behavior can also be explained by the loss of radiometric signal. In the limit that no radiometric signal is available, the Bayesian scheme would always retrieve the mean of the a priori database. As such, it is not surprising that the biases are still small, but that the quality of the retrieval is decreased significantly.

The relative merits of different sensors over land may also be less reliable than over oceans because of the simplified emissivity model employed in this study. The model does not correctly represent the true variability, which in practice all but destroys any correlation between rainfall and the lower-frequency channels that are sensitive to land surface properties. The relatively good result of the CMR-19-37 (Fig. 7h) may be an artifact of this simplification that generates an artificially homogeneous surface against which even small scattering depressions can be observed. Results are more reliable at high-frequency channels, which are less sensitive to the surface properties.

3. Errors introduced by the cloud-resolving model

In order to construct the fully three-dimensional hydrometeor profiles needed by the parametric retrievals, it is necessary to use a CRM to convert PR reflectivities into a geophysical set of parameters because the single observed quantity, Z, is not sufficient to predict the extinction, single scattering albedo, and asymmetry factor

---

**Table 1.** Statistics corresponding to synthetic rainfall retrievals for various radiometers with the nadir angle of 45° for oceanic background. Numbers in parentheses indicate percent value of true mean. Statistics represent 67 000–70 000 retrievals (depending upon the sensor) with 12-km spatial resolution.

<table>
<thead>
<tr>
<th>Sensor name</th>
<th>Mean True</th>
<th>Retrieved</th>
<th>Std dev True</th>
<th>Retrieved</th>
<th>Bias</th>
<th>Rms</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMI-350-61</td>
<td>2.375</td>
<td>2.180</td>
<td>3.189</td>
<td>2.408</td>
<td>-0.195 (8.20)</td>
<td>1.782 (75.0)</td>
<td>0.84</td>
</tr>
<tr>
<td>TMI-600-61</td>
<td>2.380</td>
<td>2.281</td>
<td>3.119</td>
<td>2.261</td>
<td>-0.100 (4.18)</td>
<td>1.934 (81.3)</td>
<td>0.79</td>
</tr>
<tr>
<td>SSM/I-600-61</td>
<td>2.380</td>
<td>2.295</td>
<td>3.119</td>
<td>2.172</td>
<td>-0.085 (3.58)</td>
<td>2.061 (86.6)</td>
<td>0.75</td>
</tr>
<tr>
<td>TMI-400-100</td>
<td>2.343</td>
<td>2.142</td>
<td>3.207</td>
<td>2.382</td>
<td>-0.202 (8.61)</td>
<td>1.782 (76.1)</td>
<td>0.84</td>
</tr>
<tr>
<td>SSM/I-833-61</td>
<td>2.343</td>
<td>2.224</td>
<td>3.156</td>
<td>2.041</td>
<td>-0.119 (5.09)</td>
<td>2.147 (91.6)</td>
<td>0.74</td>
</tr>
<tr>
<td>CMR-10-19-37-85</td>
<td>2.375</td>
<td>2.171</td>
<td>3.189</td>
<td>2.455</td>
<td>-0.204 (8.60)</td>
<td>1.767 (74.4)</td>
<td>0.84</td>
</tr>
<tr>
<td>CMR-10-37-85</td>
<td>2.375</td>
<td>2.188</td>
<td>3.189</td>
<td>2.447</td>
<td>-0.187 (7.88)</td>
<td>1.808 (76.1)</td>
<td>0.83</td>
</tr>
<tr>
<td>CMR-19-37</td>
<td>2.375</td>
<td>2.252</td>
<td>3.189</td>
<td>2.418</td>
<td>-0.123 (5.20)</td>
<td>1.882 (79.2)</td>
<td>0.81</td>
</tr>
<tr>
<td>CMR-19-37-85</td>
<td>2.375</td>
<td>2.206</td>
<td>3.189</td>
<td>2.369</td>
<td>-0.170 (7.14)</td>
<td>1.849 (77.9)</td>
<td>0.82</td>
</tr>
</tbody>
</table>

**Table 2.** Same as in Table 1, but for land background.

<table>
<thead>
<tr>
<th>Sensor name</th>
<th>Mean True</th>
<th>Retrieved</th>
<th>Std dev True</th>
<th>Retrieved</th>
<th>Bias</th>
<th>Rms</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMI-350-61</td>
<td>1.307</td>
<td>1.278</td>
<td>1.973</td>
<td>1.751</td>
<td>-0.029 (2.23)</td>
<td>0.783 (59.9)</td>
<td>0.92</td>
</tr>
<tr>
<td>TMI-600-61</td>
<td>1.381</td>
<td>1.393</td>
<td>2.044</td>
<td>1.717</td>
<td>+0.012 (0.88)</td>
<td>1.016 (73.6)</td>
<td>0.87</td>
</tr>
<tr>
<td>SSM/I-600-61</td>
<td>1.381</td>
<td>1.397</td>
<td>2.044</td>
<td>1.713</td>
<td>+0.016 (1.16)</td>
<td>1.023 (74.1)</td>
<td>0.87</td>
</tr>
<tr>
<td>TMI-400-100</td>
<td>1.307</td>
<td>1.274</td>
<td>1.973</td>
<td>1.831</td>
<td>-0.033 (2.50)</td>
<td>0.728 (55.7)</td>
<td>0.93</td>
</tr>
<tr>
<td>SSM/I-833-61</td>
<td>1.385</td>
<td>1.401</td>
<td>2.101</td>
<td>1.720</td>
<td>+0.015 (1.11)</td>
<td>1.133 (81.8)</td>
<td>0.84</td>
</tr>
<tr>
<td>CMR-10-19-37-85</td>
<td>1.307</td>
<td>1.282</td>
<td>1.973</td>
<td>1.765</td>
<td>-0.025 (1.94)</td>
<td>0.777 (59.5)</td>
<td>0.92</td>
</tr>
<tr>
<td>CMR-10-37-85</td>
<td>1.307</td>
<td>1.281</td>
<td>1.973</td>
<td>1.671</td>
<td>-0.026 (2.01)</td>
<td>0.891 (68.2)</td>
<td>0.89</td>
</tr>
<tr>
<td>CMR-19-37</td>
<td>1.307</td>
<td>1.289</td>
<td>1.973</td>
<td>1.741</td>
<td>-0.018 (1.37)</td>
<td>0.844 (64.5)</td>
<td>0.90</td>
</tr>
<tr>
<td>CMR-19-37-85</td>
<td>1.307</td>
<td>1.287</td>
<td>1.973</td>
<td>1.760</td>
<td>-0.020 (1.52)</td>
<td>0.805 (61.6)</td>
<td>0.91</td>
</tr>
</tbody>
</table>
of the raining volume needed by the passive microwave radiative transfer code.

Figure 8 shows scatter diagrams of the computed and observed $T_b$s at nine channels for the east Pacific. The observed $T_b$ is obtained directly by the TMI for the same pixels, without going through the CRMs and radiative transfer equations. As such, the directly observed data are of little use to the parametric algorithms because this data cannot be readily transformed to other channels or viewing geometries without going through the profiles provided by the CRMs. Nonetheless, these data are useful to look at the errors introduced by the CRMs. The two $T_b$s agree rather well for the emission-based channels (10–37 GHz) but disagree rather substantially at 85 GHz. This may be caused by the higher reflectivities of liquid drops relative to ice particles of the same size. Because the matching procedure uses $Z$ to find the most suitable match, the reduced weight assigned to ice particles could cause some of this disagreement. Increasing the reflectivity by 7 dBZ (Battan 1973) above the freezing level in order to create a more equal weighting between liquid and ice hydrometeors, however, did
not change the general results presented in Fig. 8. In both cases, it appears that when the liquid portion of the profile is matched, the CRMs consistently specify ice particles of an incorrect size and density, which in turn leads to a lower than observed $T_b$. This divergence between CRMs and observations is consistent with the study by Viltard et al. (2000) that also showed significant disagreement between PR and TMI at the scattering frequencies.

A number of steps can be taken to alleviate this problem. The simplest is to use the lack of fit between model-computed $T_b$ and the observed $T_b$ as a measure of the uncertainty in the forward model, which can then be used directly by the Bayesian retrieval methodology. In practice, this would imply that the 85 GHz would be generally ignored by the retrieval algorithm. This solution, however, is not particularly attractive because Fig. 8 clearly shows that a useful relation between 85-GHz $T_b$ and surface does exist in nature if not in the CRMs. A better choice would be to continue the development of the CRM physics to ensure that simulations properly match, in an average sense, the observed relationship between ice scattering and the rainfall column. Alternatively, an approach could be developed to use the 85-GHz $T_b$ from TMI to add an additional constraint to the profile selection procedure. There are a number of issues related to differences in the FOV size and incidence angles that must be resolved in this ap-
Table 3. Comparison of mean statistics of the parameters between the two databases: freezing-level height (FL), the ratio of total columnar rainwater (Σ W_i) to surface rainwater (W_{sfc}); and the coefficient of variation (σ/μ) at the FOVs of two emission channels and two rain-intensity categories.

<table>
<thead>
<tr>
<th></th>
<th>FL (m)</th>
<th>Σ W_i/ W_{sfc}</th>
<th>σ/μ at 10-GHz FOV</th>
<th>σ/μ at 19-GHz FOV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>&lt;5 mm h⁻¹</td>
<td>&gt;5 mm h⁻¹</td>
</tr>
<tr>
<td>DB-East</td>
<td>4230</td>
<td>1.19</td>
<td>1.51</td>
<td>1.30</td>
</tr>
<tr>
<td>DB-West</td>
<td>4655</td>
<td>1.36</td>
<td>1.76</td>
<td>1.90</td>
</tr>
</tbody>
</table>

4. Retrieval uncertainties and regional differences

The method by which the a priori database is constructed allows one to construct regional databases that reflect differences in PR observed profiles within each region. This allows us to explore the extent to which current passive microwave algorithms, based upon a universal a priori database, can distinguish rainfall characteristics in different regimes. A second oceanic database centered in the west Pacific (2°–12°N, 130°–160°E) was constructed for this purpose. Berg et al. (2002) documented the climate regime differences between these regions. Differences can also be seen in the mean attenuation-corrected PR profiles corresponding to a surface echo of 35 ± 1 dBZ as shown in Fig. 9.

Figure 10 shows the $T_b$-to-rain-rate relations at the four different frequencies (horizontal polarization) as well as its standard deviation within the rain interval obtained from the PR and CRM simulation-based databases (DB) for the east (DB-East) and west (DB-West) Pacific Ocean. The difference in the relationship between the two regional databases is prominent at the frequencies where emission dominates. One can note that rain-layer thickness differences, seen in Fig. 9, have a large impact on the two different concave relations below about 5 mm h⁻¹ in the emission channels (10 and 19 GHz). The rationale is presented in Table 3. The east Pacific is characterized by a lower freezing level but a greater ratio of total column rainwater (Σ W_i) to surface rainwater (W_{sfc}) than those of the west Pacific. Taken together, these properties will result in warmer $T_b$s with the same rain intensity for the west Pacific than the east, and it appears to be applicable to the lower rain rates up to about 5 mm h⁻¹. However, as rain intensity increases, the coefficient of variation or inhomogeneity factor, the ratio of standard deviation to mean rain rate within a radiometer FOV, becomes much larger for the west Pacific: 1.90 and 1.43 at the rain rates greater than 5 mm h⁻¹ for 10- and 19-GHz FOVs, respectively. The larger inhomogeneity in the west Pacific significantly weakens the physical relation between increasing the liquid water content and microwave radiance. It suggests that rainfall estimates based on emission channels can have significant variability superimposed on the mean beam-filling correction discussed by Wilheit (1986). One can also note that even when the inhomogeneity factors are similar at both low and high rain rates, as observed in the west Pacific (1.42 and 1.43 at 19-GHz FOV), its effect is more prominent at the higher rain rates. This is related to the higher nonlinearity at the high rain rates in the relationship resulting in a greater averaging bias within the FOV. The relationships at 37 GHz for the east and west Pacific are rapidly saturated and look fairly similar. At the scattering frequency (85 GHz), where the rain inhomogeneity is not significant due to the high spatial resolution, the brightness temperature seems to have little relation to the surface rainfall. This, however, is due primarily to the PR-CRM shortcomings as the actual $T_b$s for the TMI sensor did show significantly greater correlations as shown in Fig. 8. While not shown here, the trend between surface
Fig. 10. Relationship between $T_b$ and rain rate ($R$) at four different horizontally polarized channels of TMI obtained from the simulated databases over the east and west Pacific during Dec 1999.

Rainfall and observed 85 GHz did not vary significantly between the two regions.

In order to quantify the advantage of building regional databases instead of global ones, three experiments are performed. The first consists of using the separate databases for the east and west Pacific, and running retrievals separately for each location. The results are presented in Table 4. Note that the results for the east Pacific are those already presented in Table 1, but repeated here for clarity. As can be seen, the composite statistics for the west Pacific show a somewhat lower retrieval performance. This is mainly because the radiometric signature of the west Pacific has a smaller dynamic range due to the greater rainfall inhomogeneity (see Fig. 10).

The second experiment combines the databases from the east and west Pacific into a single entity, and retrievals for both the east and west Pacific use this common database. Results from this retrieval are presented in Table 5. Retrievals in the east Pacific are degraded, while those over the west Pacific remain largely unchanged. This suggests that retrievals for a region, such as the east Pacific, with a large dynamic response of $T_b$ to surface rainfall (Fig. 10) can be easily corrupted by some inappropriate information. Adding corrupt info-

### Table 4

<table>
<thead>
<tr>
<th>Sensor name</th>
<th>Mean</th>
<th>Std dev</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True</td>
<td>Retrieved</td>
</tr>
<tr>
<td>TMI-350-61 East</td>
<td>1.307</td>
<td>1.278</td>
</tr>
<tr>
<td>TMI-350-61 West</td>
<td>1.479</td>
<td>1.348</td>
</tr>
</tbody>
</table>
Fig. 11. Mean reflectivity profiles for the east and west Pacific (similar to Fig. 9), but now retrieved profiles with (a) correct databases (b) combined databases, and (c) reversed databases.

Table 5. Same as Table 4, but for “combined databases.”

<table>
<thead>
<tr>
<th>Sensor name</th>
<th>Mean</th>
<th>Std dev</th>
<th>Bias</th>
<th>Rms</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMI-350-61 East</td>
<td>1.307</td>
<td>1.112</td>
<td>1.973</td>
<td>1.622</td>
<td>0.195 (14.92)</td>
</tr>
<tr>
<td>TMI-350-61 West</td>
<td>1.479</td>
<td>1.349</td>
<td>2.211</td>
<td>1.789</td>
<td>0.130 (8.77)</td>
</tr>
</tbody>
</table>

The information to the west Pacific database has little impact on those retrievals because the dynamic response was never very good.

The final experiment is one in which the databases from the east and west Pacific are deliberately switched so that the $T_b$s in the east Pacific are interpreted using the database created in the west Pacific and vice versa. These results are presented in Table 6. In addition, Fig. 11 shows the mean vertical profile of computed reflectivity for the east and west Pacific for the three scenarios discussed above. The vertical profiles of reflectivity retrieved with the exact a priori information turn out to include quite similar characteristics of observations for both the regions (Fig. 11a). Using the combined databases (Fig. 11b), the difference between the east and west Pacific does not appear to be retrieved as clearly.
as in the previous case, while inappropriate databases seem to produce incorrect vertical information (Fig. 11c).

From these simple experiments, it is clear that the observed brightness temperatures, even for relatively high-quality sensors like the TRMM TMI, do not retrieve all information independent of the a priori database. It explains our previous assertion that the quality of the synthetic retrievals is partially due to the a priori database significantly restricting the profiles available to the retrieval algorithm only to those that are actually observed in a particular region. It further demonstrates that the performance of radiometers can vary regionally depending on how well the radiometric signature is related to precipitation.

Regional databases become somewhat more difficult to implement in an operational scenario. The simplest notion is to define a priori databases on a regular latitude/longitude/season grid. The disadvantage of this approach is that grid boxes over regions without much rain would have to be blended with neighbors before a sufficiently diverse a priori database emerges. The best way to implement this without corrupting regional databases with profiles from potentially different regimes needs further investigation.

5. Conclusions

In this paper, we demonstrated that the information from TRMM PR can be used in conjunction with cloud-resolving models to construct a parametric framework for passive microwave retrieval algorithms of rain from any sensor. All sensor configurations examined in this study were unbiased relative to each other. Correlations and rms errors were related to sensor capabilities as expected. While agreement between computed $T_b\text{a}$ and those observed directly by the TMI were generally good, the study showed that adding a 10-GHz channel did not improve the retrieval quality. The large FOV of the lower-frequency channels appears to be of little utility until the 19-GHz channel becomes saturated. This is more important for higher-resolution sensors than it is for the relatively coarse-resolution sensors like the current SSM/I. Advances in surface emissivity modeling may be needed before the results shown over land can be made as robust as they are over the oceans.

The regional nature of the a priori database developed here allowed us to compare retrievals using location-specific databases to those achievable with only a universal database. Results showed that location-specific databases can improve results significantly. Current research aimed at improving CRM simulations, coupled with improved observational capabilities proposed for the GPM core satellite, can only improve upon the results shown here.

Acknowledgments. NASA supported this research under Grant NCC5-288.

REFERENCES

—, and L. Giglio, 1994: A passive microwave technique for es-

<table>
<thead>
<tr>
<th>Sensor name</th>
<th>True</th>
<th>Retrieved</th>
<th>Bias</th>
<th>Rms</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMI-350-61 East</td>
<td>1.307</td>
<td>1.107</td>
<td>1.973</td>
<td>1.622</td>
<td>0.200 (15.30)</td>
</tr>
<tr>
<td>TMI-350-61 West</td>
<td>1.479</td>
<td>1.557</td>
<td>2.211</td>
<td>1.792</td>
<td>–0.078 (–5.28)</td>
</tr>
</tbody>
</table>

Table 6. Same as Table 4, but for “reversed databases.”


