Precipitation and Latent Heating Distributions from Satellite Passive Microwave Radiometry. Part I: Improved Method and Uncertainties

WILLIAM S. OLSON,[a] CHRISTIAN D. KUMMEROW,[b] SONG YANG,[c] GRANT W. PETTY,[d] WEI-KUO TAO,[e] THOMAS L. BELL,[e] SCOTT A. BRAUN,[e] YANSEN WANG,[a,*] STEPHEN E. LANG,[f] DANIEL E. JOHNSON,[g,#] AND CHRISTINE CHIU[a]

[a]Joint Center for Earth Systems Technology, University of Maryland, Baltimore County, Baltimore, Maryland
[b]Department of Atmospheric Sciences, Colorado State University, Fort Collins, Colorado
[c]School of Computational Science, George Mason University, Fairfax, Virginia
[d]Department of Atmospheric and Oceanic Sciences, University of Wisconsin—Madison, Madison, Wisconsin
[e]Laboratory for Atmospheres, NASA Goddard Space Flight Center, Greenbelt, Maryland
[f]Science Systems and Applications, Inc., Lanham, Maryland
[g]Goddard Earth Sciences and Technology Center, Greenbelt, Maryland

ABSTRACT

A revised Bayesian algorithm for estimating surface rain rate, convective rain proportion, and latent heating profiles from satellite-borne passive microwave radiometer observations over ocean backgrounds is described. The algorithm searches a large database of cloud-radiative model simulations to find cloud profiles that are radiatively consistent with a given set of microwave radiance measurements. The properties of these radiatively consistent profiles are then composited to obtain best estimates of the observed properties. The revised algorithm is supported by an expanded and more physically consistent database of cloud-radiative model simulations. The algorithm also features a better quantification of the convective and nonconvective contributions to total rainfall, a new geographic database, and an improved representation of background radiances in rain-free regions. Bias and random error estimates are derived from applications of the algorithm to synthetic radiance data, based upon a subset of cloud-resolving model simulations, and from the Bayesian formulation itself. Synthetic rain-rate and latent heating estimates exhibit a trend of high (low) bias for low (high) retrieved values. The Bayesian estimates of random error are propagated to represent errors at coarser time and space resolutions, based upon applications of the algorithm to TRMM Microwave Imager (TMI) data. Errors in TMI instantaneous rain-rate estimates at 0.5°-resolution range from approximately 50% at 1 mm h^{-1} to 20% at 14 mm h^{-1}. Errors in collocated spaceborne radar rain-rate estimates are roughly 50%–80% of the TMI errors at this resolution. The estimated algorithm random error in TMI rain rates at monthly, 2.5° resolution is relatively small (less than 6% at 5 mm day^{-1}) in comparison with the random error resulting from infrequent satellite temporal sampling (8%–35% at the same rain rate). Percentage errors resulting from sampling decrease with increasing rain rate, and sampling errors in latent heating rates follow the same trend. Averaging over 3 months reduces sampling errors in rain rates to 6%–15% at 5 mm day^{-1}, with proportionate reductions in latent heating sampling errors.

1. Introduction

Over the last decade, diagnostics of time-/space-averaged satellite rainfall estimates have helped to create a better picture of the earth's climate and its variability (e.g., Rasmussen and Arkin 1993; Xie and Arkin 1997; Curtis and Adler 2000; Adler et al. 2003). These studies have relied upon remote sensing of precipitation from infrared, passive microwave, and spaceborne radar measurements, culminating in the Tropical Rainfall Measuring Mission (TRMM; 1997–present). Moreover, it has been amply demonstrated that precipitation measurements from space have had a beneficial impact on general circulation model assimilations and numerical weather prediction model forecasts using data assimilation methods

(e.g., Hou et al. 2000, 2001, 2004; Krishnamurti et al. 2001).

Toward a better understanding of how precipitation processes affect the atmosphere, the next logical step is to consider the convective and stratiform partitioning of total precipitation, because this partitioning is linked to the organization of convective systems and the distributions of vertical motion and latent heat release in the atmosphere (Houze 1989). Latent heating is a driver of atmospheric circulations, from the scale of individual convective elements to the scales of the Hadley and Walker circulations; therefore, knowledge of the 4D distribution of latent heating gained from satellite observations can be used to study these circulations and to help quantify the diabatic heating component of the atmospheric energy budget. In a series of studies, Yang and Smith (1999a,b, 2000) demonstrated that it is possible to estimate latent heating profiles from Special Sensor Microwave Imager (SSM/I) satellite observations. Only a limited number of studies have utilized satellite estimates of convective/stratiform proportion and latent heating in numerical model assimilation experiments. However, the significant improvement of numerical model assimilations and forecasts utilizing even approximate satellite rain-rate estimates has spurred interest in new experiments in which satellite convective/stratiform rain proportion and latent heating rates would be assimilated.

One of the primary objectives of TRMM is to gain a better understanding of the three-dimensional distribution and evolution of atmospheric latent heating in the Tropics; see Simpson et al. (1988). The TRMM polar-orbiting satellite observatory is fitted with passive and active microwave sensors that provide measurements of the horizontal and vertical structure of precipitation in the atmosphere with relatively high spatial sampling (minimum footprint spacing ~5 km). The aim of the present study is to examine the potential for estimating consistent precipitation and latent heating rates based upon passive microwave observations from the TRMM Microwave Imager (TMI). The heritage of this study is an investigation in which a Bayesian estimation method was applied to measurements from the Special Sensor Microwave Imager to estimate the surface rainfall rate, convective rainfall proportion, and latent heating rates (Olson et al. 1999). In the previous study it was demonstrated that in addition to measured microwave radiances, empirical estimates of the convective fraction of precipitation within the nominal satellite footprint were required to retrieve reasonably unambiguous estimates of surface rainfall rate and latent heating. One reason for the positive impact of convective fraction information is that it provides a measure of the hori-zontal inhomogeneity of the rain field within the satellite footprint, which is critical for establishing the link between footprint-averaged rain-rate and upwelling microwave radiances. The convective fraction is also an indicator of the vertical motion and latent heating profile. For example, a large convective fraction at the scale of the sensor footprint (~10 km) is correlated with stronger upward motion and positive latent heating through the depth of the cloud layer, while a large fraction of nonconvective (stratiform) rain generally indicates weak mesoscale ascent and heating in the upper troposphere with descent and evaporative cooling at lower altitudes; see Houze (1989).

The objectives of the present study are to improve and extend the Bayesian estimation method described in Kummerow et al. (2001; version 5) in order to provide consistent estimates of precipitation and latent heating, based upon passive microwave observations from the TMI over ocean. The technique is applied only over ocean surfaces because any vertical structure information contained in TMI observations is compromised by the strong microwave emission from land surfaces; that is, microwave emission/absorption by liquid precipitation cannot be easily distinguished from variations in land emission. The method is improved by (a) expanding the algorithm's cloud-radiative model database to include a greater diversity of precipitation systems, (b) making basic adjustments to the ice precipitation microphysics in the cloud-resolving model (CRM) simulations of the algorithm's database to produce more realistic graupel and snow distributions, (c) including the effects of mixed-phase precipitation in the cloud-resolving model simulations, (d) utilizing consistent definitions of convective rain and total rain area in the context of cloud-resolving model simulations and satellite observations, (e) including a new geographic database to better separate ocean, coast, and land areas, and (f) establishing a more consistent microwave radiance "background" in rain-free areas. The combination of these changes leads to passive microwave estimates of precipitation that have greater consistency with independent radar estimates and latent heating distributions that have more fidelity with climatological distributions; see Yang et al. (2006, hereinafter Part II). The resulting Bayesian estimation method is at the heart of the current TRMM facility algorithm (2A12, version 6) for estimating precipitation-related parameters from TMI observations.

In section 2, the characteristics of the TMI observations and the basic estimation method are briefly reviewed, with modifications for improved parameter estimation described in section 3. Synthetic retrieval studies, in which the algorithm is applied to microwave

radiances synthesized from cloud-resolving model simulations, are used to estimate the biases and random errors in retrieved parameters. Estimates of random errors resulting from incomplete information in the microwave data can also be derived from the Bayesian (algorithm based) method. Both synthetic retrievals and the algorithm-based method are described and applied in section 4. The general capability of the microwave algorithm is demonstrated through applications to TMI observations from July 2000 over the tropical and subtropical oceans. Monthly mean estimates of precipitation and latent heating and their uncertainties, including errors resulting from infrequent sampling, are examined. The paper concludes with a brief summary and reflections on the direction of future work (section 5).

As described, this Part I of the two-part series provides a review of the changes introduced into the version-6 TMI algorithm and the theoretical "modeling" of uncertainties based upon the algorithm. In Part II, estimates of algorithm errors at various scales are obtained directly from comparisons with independent observations, and these error estimates are compared with the modeled uncertainties derived in Part I. Remote sensing estimates of rain rate and latent heating, and knowledge of their uncertainties, are important for current and future applications in data assimilation and climate analysis.

2. Data and basic method

a. TRMM observations

The analyses presented in this study are based upon observations, both synthesized and actual, from the TMI. The TMI is one of five sensors aboard the TRMM satellite observatory, which was launched into low-Earth orbit in November 1997 to provide data on the characteristics of convection in the Tropics and subtropics (35°S–35°N). The TMI is a scanning passive microwave radiometer with dual-polarization channels at 10.65, 19.35, 37, and 85.5 GHz, and a vertical polarization channel at 21.3 GHz; see Table 1. The lower-frequency channels are primarily sensitive to the vertical path integral of liquid precipitation in the atmosphere. The channels become increasingly sensitive to the vertical path integral of ice-phase precipitation as the channel frequency increases, while the range of sensitivity to rain decreases. Because of these sensitivities, the TMI has a crude precipitation-profiling capability, which is somewhat compromised by limited spatial resolution at the lower frequencies. In addition to the information provided by TMI on vertical precipitation structure, horizontal structure information can be gleaned from variations of precipitation signatures in

TABLE 1. Characteristics of the TMI and PR. The horizontal resolution specifications of the PR are at nadir view. H: horizontal; V: vertical.

TMI		
Channel frequency (GHz)/polarizations	Horizontal resolution (km)	Noise (K)
10.65/H, V	37 × 63	0.6
19.35/H, V	18 × 30	0.5
21.3/V	18 × 23	0.7
37.0/H, V	9 × 16	0.3
85.5/H, V	5 × 7	0.7
PR		
Frequency (GHz)	Horizontal resolution at surface/range resolution (km)	Uncertainty resulting from sampling (dB)
13.8	4.3/0.25	0.7

the swath imagery or by deconvolution of the multiresolution observations, as in Petty (1994b). The along-track sampling resolution of all TMI channels is about 14 km; the cross-track sampling is 4.5 (85.5 GHz) and 9 (remaining channels) km.

Another TRMM instrument, the precipitation radar (PR), measures precipitation backscatter at 13.8 GHz near nadir and is thus used to infer profiles of precipitation water content; see Table 1. The PR's 215-km-wide swath is centered within the TMI's 760-km swath, providing coincident coverage within the PR swath. The PR's greater range resolution (0.25 km) and horizontal sampling resolution (4.3 km) with respect to the TMI lead to more structural detail in retrieved precipitation fields, and therefore precipitation estimates from the PR are used as a comparative reference in the current study.

b. Estimation of cloud properties from TMI

The TMI retrieval algorithm (2A12, version 6) is based upon a Bayesian technique described in Kummerow et al. (1996, 2001) with an extension to latent heating estimation by Olson et al. (1999). In the algorithm, CRM simulations, coupled to a radiative transfer code, are used to generate a large supporting database of simulated precipitation/latent heating vertical profiles and corresponding upwelling microwave radiances. Given a set of observed multichannel microwave radiances from a particular sensor, the entire database of simulated radiances is scanned; the "retrieved" profile is a composite of those profiles in the database that correspond to simulated radiances consistent with the observed radiances. Formally, a TMI estimate of profile parameters $\hat{E}[\mathbf{x}]$ is given by

$$\hat{E}[\mathbf{x}] = \sum_k \mathbf{x}_k \frac{\exp\{-0.5[\mathbf{I}_s(\mathbf{x}_k) - \mathbf{I}_o]^{\mathrm{T}}(\mathbf{S}_I + \mathbf{O}_I)^{-1}[\mathbf{I}_s(\mathbf{x}_k) - \mathbf{I}_o] + C\}}{\hat{M}}, \tag{1}$$

where the model profile vector $\mathbf{x}_k$ contains all of the parameters, including the surface rain rate, convective rain rate, liquid-/ice-phase precipitation, and latent heating profiles, corresponding to the simulated radiance indices $\mathbf{I}_S(\mathbf{x}_k)$. The radiance indices, constructed from radiances at the different radiometer channel frequencies/polarizations, are the normalized polarization and scattering indices defined by Petty (1994a). Similarly defined, $\mathbf{I}_O$ is a vector of sensor-observed radiance indices. Respectively, $\mathbf{S}_I$ and $\mathbf{O}_I$ are error covariance matrices of the simulated and observed microwave radiance indices, and $\hat{M}$ is a normalization factor.

Additional information regarding the observed profile, such as estimates of the area fractions of rain and convective rain within the nominal satellite footprint (14 km $\times$ 14 km for TMI), is included in the constraint term C,

$$C = -0.5[\mathbf{f}_s(\mathbf{x}_k) - \mathbf{f}_o]^{\mathrm{T}}(\mathbf{S}_f + \mathbf{O}_f)^{-1}[\mathbf{f}_s(\mathbf{x}_k) - \mathbf{f}_o]. \tag{2}$$

Here, $\mathbf{f}_S$ is a vector of simulated constraint parameters, and $\mathbf{f}_O$ is a vector of corresponding observed parameters; $\mathbf{S}_f$ and $\mathbf{O}_f$ are error covariance matrices of the simulated and observed constraint parameters, respectively. A description of the rain area and convective rain area constraint parameters is deferred to section 3. Only the diagonal terms of the error covariance matrices, $\mathbf{S}_I$, $\mathbf{O}_I$, $\mathbf{S}_f$, and $\mathbf{O}_f$, are specified as in the version-5 TMI algorithm; see Kummerow et al. (2001).

The summation in (1) is over all simulated profiles/radiance indices in the supporting cloud-radiative model database. In principle, any cloud property represented in the supporting cloud-radiative model simulations can be estimated using (1) to the extent that there is sufficient sensitivity of the passive microwave observations to variations in that property. In this way, estimates of surface rainfall rate, convective rain proportion, and profiles of precipitation and latent heating can be made with different degrees of accuracy.

Because, in general, multichannel passive microwave observations contain limited information regarding precipitation and related cloud parameters, there are, in fact, a distribution of these parameters that are consistent with any set of observations at a given footprint location. The expression (1) gives the mean of this distribution, but it is also possible to calculate the variance of the distribution for any single estimated parameter x using

$$\hat{\sigma}^2[x] = \hat{E}\{(x - \hat{E}[x])^2\}, \tag{3}$$

which yields a measure of the uncertainty in the estimate of x because of the limited information content of the observations.

The uncertainty represented by (3) would exist even if the cloud-radiative model simulations in the algorithm's supporting database and the radiometer observations were error free. Therefore, although (3) accounts for random errors resulting from the limited information content of the observations, additional uncertainties in estimates resulting from errors in cloud-radiative modeling may occur. Because true validation of precipitation-related quantities using independent observations is difficult, (3) at least provides a lower bound on the random error of algorithm estimates—a basic "building block" for estimates of the random error in the derived products (described in section 4).

3. Modifications of the TMI algorithm

In the development of the version-6 TMI facility algorithm, several modifications are included to allow for latent heating estimation. These modifications lead to not only improved physical models that better represent the relationships between cloud properties and upwelling microwave radiances at the TMI frequencies, but also a better extraction of information from the TMI observations to isolate the dependencies of the observations on the cloud properties that are estimated. Although the main driver for these modifications is the estimation of latent heating, the improved sensitivity of the algorithm leads to estimates of the surface rainfall rate with generally less bias with respect to independent estimates; see Part II.

a. Greater diversity in the supporting cloud-radiative model database

In version 5 of the TMI algorithm, the Bayesian estimation method was supported by cloud simulations corresponding to only three different observed environments in the Tropics and subtropics for ocean applications. These simulations are replaced by the six simulations listed in Table 2. Four of the new simulations are performed using the Goddard Cumulus Ensemble (GCE) model, which is a three-dimensional, nonhydrostatic cloud-resolving model described in Tao and Simpson (1993) and Tao (2003a). Among the physical

TABLE 2. Cloud-resolving model simulations that currently support the version-6 TMI algorithm. Locations of simulations are often indicated by the field campaign from which environmental forcing data were obtained.

Classification	Description	Model
Tropical cyclone	Hurricane Bob, 1991	MM5
Tropical squall line	GATE, 12 Sep 1974	GCE
Tropical squall line	TOGA COARE, 22 Feb 1993	GCE
Tropical convection	TOGA COARE, 19–26 Dec 1992	GCE
Extratropical cyclone	North Atlantic, 6 Dec 1992	MM5
Extratropical cyclone	North Atlantic, 23 Jul 1999	MM5

parameterizations in this model is a description of cloud microphysical processes based upon Lin et al. (1983), with additional processes derived from Rutledge and Hobbs (1984). Bulk water contents of cloud liquid, rain, cloud ice, snow, and graupel are calculated prognostically. This single-moment microphysical scheme used in the GCE model is called the "3ICE" scheme.

GCE short-term simulations of mesoscale convective systems are usually initiated with a spreading cool pool within a fixed environment obtained from rawinsonde data. However, the longer-term evolution of cloud systems can be simulated by nudging toward environmental conditions using observed large-scale advective tendencies of temperature and humidity and observed horizontal momentum (Johnson et al. 2002). The 19–26 December Tropical Ocean and Global Atmosphere Coupled Ocean–Atmosphere Response Experiment (TOGA COARE) simulation listed in Table 2 was performed in this manner. The utility of the longer-term simulations is that a greater spectrum of convective system types can be represented because of changing environmental conditions in the model, and these varied systems can be extracted for algorithm applications by sampling the simulations over time. In the present study, less organized convection, not available from the cool-pool–forced simulations, is drawn from three periods on 23 December of the TOGA COARE simulation.

Although convective systems up to the scale of squall lines can be successfully simulated using the GCE model, the evolution of convection embedded in larger-scale systems requires a model that can represent scale interaction. For this purpose, the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5; see Dudhia 1993) is utilized, but the microphysical scheme used in the MM5 simulations is again the 3ICE scheme, which has been incorporated into MM5; see Tao (2003b). MM5 simulations nested to ~2 km resolutions are re-

quired to capture the cloud and precipitation spatial structures necessary for the proper calculation of microwave radiative transfer through the simulated structures. Using nested grids, MM5 simulations of the inner core of Hurricane Bob (1991) and the warm and cold frontal regions of wintertime and summertime extratropical cyclones are performed.

Although the simulations listed in Table 2 represent only a small fraction of systems that might be observed by the TMI, inclusion of these simulations in the Bayesian method's supporting database leads to much greater diversity in candidate precipitation and latent heating profiles than in earlier algorithm implementations.

b. Adjustment of cloud-resolving model ice microphysics

Prior to this study, a significant shortcoming of simulations based upon the 3ICE microphysics were distributions of graupel that extended widely over stratiform precipitation regions, far from any convective updrafts. These graupel distributions were considered erroneous because graupel grows primarily through the accretion of liquid water, which is generally confined to regions close to convective updrafts.

To correct this problem within the constraints of the 3ICE bulk microphysical parameterization, the efficiency for collection of snow by graupel is set to zero at model grid points where concentrations of cloud liquid water and rain are negligible. Although this correction is an oversimplification, it avoids the difficulty of reformulating the snow–graupel collection kernel and is supported by evidence that the collection efficiency of "dry" ice particles is relatively small; see Pruppacher and Klett (1997) for a general discussion. The correction also improves the distributions of microwave scattering by ice-phase precipitation, because excessive scattering by graupel, previously noted in simulations of stratiform regions, is eliminated.

c. Inclusion of the effects of mixed-phase precipitation in CRM simulations

In the CRM simulations listed in Table 2, the melting of ice-phase hydrometeors is simplified, such that any meltwater is immediately categorized as rain; that is, there is no explicit representation of mixed-phase precipitation in these simulations. To account for the effects of mixed-phase precipitation, the 1D parameterization for melting precipitation described in Olson et al. (2001a,b) is applied to all stratiform grid points in the CRM simulations of the algorithm's supporting database. The parameterization is a steady-state model that simulates the evolution of the spectra of melting

snow and graupel particles, given their initial spectra just above the freezing level as simulated by the CRM. The concentrations and electromagnetic properties of the melting particles are computed, replacing those of the original CRM.

d. Definition of cloud and precipitation properties in simulated footprints

As described in section 2, any cloud/precipitation properties **x** simulated by the CRM's can potentially be retrieved from the TMI data using (1). The specific choice of cloud/precipitation properties and their resolution is somewhat arbitrary, however. Given the sample spacing of the TMI sensor along track (13.9 km), a 14 km $\times$ 14 km nominal "footprint" is selected for the resolution of the version-6 algorithm products.

Simulated properties such as the surface rainfall rate, cloud and precipitation water contents, and latent heating are horizontally averaged over 14 km $\times$ 14 km areas in the CRM domains to represent properties at product resolution. Cloud latent heating is defined as $Q_1 - Q_R$, the apparent heat source less the contribution from radiative heating/cooling; see Yanai et al. (1973). Eddy heat flux convergence contributions to $Q_1 - Q_R$ are defined relative to a 55-km-resolution mean state.

The classification of convective and nonconvective grid points in the CRM simulations is required not only for determining convective and stratiform rain proportions, which are estimated using the TMI algorithm, but also for determining the area coverage of convection, which, incorporated into (2), is used to constrain algorithm estimates. The new convective/nonconvective classification of CRM grid points in version 6 is based upon an evaluation of the vertical structure of model-simulated mass fluxes, inspired by the partitioning method of Xu (1995). Using this convective/nonconvective classification, the area fraction of convection within the simulated footprints (14 km $\times$ 14 km) is calculated. In addition to the convective area fraction, the area fraction of total rainfall rate, defined as the area within each simulated 14 km $\times$ 14 km footprint where the model gridpoint values of surface rainfall rate exceed 0.3 mm h^{-1}, is also calculated. The area fractions of convection and total rainfall are compared with empirical estimates derived from the TMI data in the algorithm's constraint term, (2). The convective contributions to surface rainfall rate and $Q_1 - Q_R$ within each simulated footprint are also computed and included in the vector **x** of parameters to be estimated.

e. New geographic database

Crucial to the operation of the algorithm is the correct classification of the earth's surface within the satellite footprint as predominantly ocean, coast, or land. Previously, the definition of "coast" was based on a threshold applied to the distance to the nearest land (if over water) or water (if over land). This definition has the drawback that a rather large area of open water centered on even a very small island will be classified as coast, preventing the utilization of the more appropriate ocean algorithm for this case. Additionally, large areas of interior land are classified as coast because of the presence of nearby rivers and lakes.

A revised geographic database is therefore developed based on the minimum radius of a circle encompassing a specified minimum fraction of the opposite surface type. The starting point of the geographic database is the U.S. Navy's 1/6° $\times$ 1/6° global "elevation" dataset, which includes a terrain classification of each grid point. For each 1/6° grid point over water, the radius R of a circle centered on that location is increased until the circle encompasses 5% land area. The final radius is then recorded at that grid point as the effective distance from significant land contamination. For points over land, a similar procedure is used, except that the threshold for water coverage is set to 20%. A grid point over land is then classified as coast if R is less than 50 km; points over water are classified as ocean if R is less than 30 km.

The above thresholds for land/water fraction and for radius R are determined subjectively, based on examination of TMI images and rain-rate estimates in coastal areas. The intent is to classify as few grid points as possible as coast, while still ensuring that TMI footprints centered over locations classified as ocean and land would not experience operationally significant contamination by the opposite surface type.

f. Calculation of the "background" rain-free radiance field

In this study, the TMI algorithm is applied to ocean locations, and therefore the radiance indices [$\mathbf{I}_S$ and $\mathbf{I}_O$ in (1)] are the normalized polarization and scattering indices described by Petty (1994a). These indices were developed to isolate the microwave signal resulting from precipitation particles from variations in radiances because of varying ocean surface emissivity, sea surface temperature, and atmospheric water vapor, and also to decouple the brightness temperature effects of microwave attenuation from those resulting from scattering. Both indices depend upon reasonable estimates of "background" radiances in the absence of clouds or precipitation.

In version 5, the background radiances were estimated from actual radiances at nearby locations deemed to be relatively free of rain and cloud. Unfor-

tunately, the criteria used for determining such points were unreliable and, in any case, even the existence of cloud-free pixels could not be guaranteed within any reasonable radius. Consequently, estimates of background radiances were commonly too warm and too depolarized, by as much as 10 K or more.

In version 6, an improved method is implemented based on direct TMI estimates of column water vapor V and ocean surface wind speed U, both of which are possible even in cloudy areas as long as precipitation contamination is not severe. This method greatly increases the fraction of the ocean area for which reliable local estimates of background radiances can be made; these are then interpolated spatially into the relatively small areas for which precipitation contamination is significant. The basic procedure is outlined by Petty (1994a). The net result of the revision is that fields of background radiances are both smoother and significantly more realistic than those obtained using the version-5 method.

The radiance indices in the cloud-radiative model simulations of the algorithm's database are determined by evaluating the true background radiance field. That is, upwelling radiances at each model grid point are calculated with all cloud and precipitation water contents set to zero.

g. Reformulation of the convective/nonconvective rain area constraint

The changes in the version-6 TMI algorithm that have the greatest impact on estimates of cloud latent heating profiles are those related to the estimation of the parameters describing the convective coverage from TMI observations, and the evaluation of the same parameters in the CRM simulations. In previous versions of the TMI algorithm, only estimates of the area fraction of convection within the nominal TMI footprint were used as a constraint in (2). In version 6, the area fractions of both convective rain and total rain are estimated, and these are included in (2).

As in version 5, the area fraction of convective rain within the TMI footprint in version 6 is computed as a

minimum variance estimate, combining convective fraction estimates from both TMI polarization signatures and image texture (see Olson et al. 2001c);

$$ f_{conv} = \left(\frac{f_{pol}}{\sigma_{pol}^2} + \frac{f_{tex}}{\sigma_{tex}^2} \right) \Big/ \left(\frac{1}{\sigma_{pol}^2} + \frac{1}{\sigma_{tex}^2} \right). \tag{4} $$

Here, f_{pol} is the convective-fraction estimate based upon TMI polarization data, f_{tex} is the estimate based upon image texture information, and σ_{tex} and σ_{pol} are the uncertainties corresponding to these estimates.

In version 6, the polarization-based estimate of convective area fraction f_{pol} is given by (8) in Olson et al. (2001c), but with a slight along-scan filtering to accommodate the nominal footprint dimensions (14 km × 14 km) of the version-6 algorithm. A statistical adjustment of f_{pol} is applied based upon synthetic radiance data to make it consistent with cloud model–based convective fractions. The error variance σ_{pol}^2 of polarization-based convective fraction estimates is given by (16) in Olson et al. (2001c), propagated through the same along-scan filter and statistical adjustment.

The polarization-based convective fraction estimate is reliable only if significant scattering by precipitation-sized ice is observed. If not, then microwave image texture and liquid precipitation emission signatures must be used to identify convection. In the present study, if the 85-GHz scattering index S85 (from Petty 1994a) is less than 40 K, then an alternative texture-based estimate of convective fraction is calculated. In estimating the convective fraction based upon texture information, three indices are used to first decide whether or not convection might be present in the TMI footprint—(a) P37$_{grad}$, the maximum difference between the normalized polarization at 37 GHz (Petty 1994a) and the normalized polarizations of neighboring footprints, (b) w_{pol}, a "probability of convection" based upon the spatial distribution of f_{pol}, and (c) P37$_{max}$, the maximum second spatial derivative of 37-GHz normalized polarization in any direction.

The texture-based area fraction of convection within the footprint is estimated using the normalized polarization difference at 37 GHz, as follows:

$$ f_{tex} = \begin{cases} 1 - \text{P37}, & \text{for S85} < 40 \text{ K} \quad \text{and} \\ & [(\text{P37}_{grad} > 0.2 \quad \text{and} \quad w_{pol} > 0.1) \quad \text{or} \quad \text{P37}_{max} > 0.3] \\ 0, & \text{otherwise.} \end{cases} \tag{5} $$

For optically thick precipitation regions partially filling a radiometer footprint, $1 - \text{P37}$ is approximately equal to the area fraction of precipitation within the footprint (Petty 1994a). However, even if convection is indi-

cated by the index tests on the right-hand side of (5), not all of the rain within the footprint is necessarily convective. Therefore, a statistical adjustment of f_{tex} is applied based upon synthetic radiance data to make it

consistent with model-generated convective fractions. If a texture-based convective fraction estimate is allowed by (5), then the error variance σ^2_{tex} of the estimate is assigned a value of 0.2 based upon previous work by Hong et al. (1999).

It has been demonstrated that algorithm constraints on convective rain area lead to more accurate passive microwave estimates of precipitation and latent heating (Olson et al. 1999); however, the relative proportions of convective and stratiform rain flux are more closely related to vertical latent heating structure (see Tao et al. 1993). Although microwave radiometric signatures cannot be interpreted in terms of convective/stratiform rain flux directly, the *area* proportions of convective and stratiform precipitation coverage can be inferred if, in addition to convective area coverage, the total rain coverage can be estimated. Therefore, precipitation-related quantities retrieved from TMI data are also constrained by estimates of the total rain area within the nominal footprint. Following Petty (1994a), the total rain area fraction in the plane-parallel limit is first approximated by

$$f_{\text{rain}} = 1 - \text{P37}, \tag{6}$$

and then adjusted using synthetic data.

Estimates of convective area fraction and total rain fraction are incorporated into the constraint term [(2)] of the Bayesian method.

4. Error estimates

Errors in retrieved precipitation-related parameters can be estimated from (a) comparisons of TMI-retrieved parameters with independent data, (b) algorithm applications to synthetic data, and (c) the algorithm itself, drawing upon information provided by the Bayesian formulation. Because independent estimates of rain rate from ground-based radars or rain gauges may contain considerable random error, it is not always possible to distinguish these errors from algorithm errors. Therefore, alternatives (b) and (c) are explored here to independently estimate errors that are intrinsic to the algorithm and to provide a basis for the development of error models. Evaluation of TMI surface rain-rate, convective rain-rate, and $Q_1 - Q_R$ estimates using independent, ground-based observations is performed in Part II of this series.

a. Synthetic data

TMI observations are synthesized using the same procedure utilized to create the algorithm's supporting database (see section 3a). A subset of the cloud-resolving model simulations listed in Table 2, but separate from the simulations used in the algorithm's database, are set aside to represent "true" fields of cloud, precipitation, and latent heating. The subset consists of 3D model volumes at two time periods each from the Hurricane Bob, Global Atmospheric Research Program (GARP) Atlantic Tropical Experiment (GATE) 12 September simulation, and the TOGA COARE 22 February and 19–26 December simulations. Upwelling radiances at the TMI frequencies/polarizations are calculated using Eddington's second approximation, and these radiances are convolved by functions approximating the TMI antenna patterns to synthesize radiances as they might be measured by TMI. Because the polarization of scattered radiances at 85.5 GHz is used in the TMI algorithm (section 3g), but not calculated using the Eddington method, the empirical function [Olson et al. 2001c; their (10)] is utilized to simulate the polarization of scattered radiances. A nominal level of Gaussian-distributed noise, with a standard deviation of 1 K, is added to the convolved radiances to simulate sensor noise.

The TMI algorithm is applied to the synthesized microwave radiances, and the estimated parameters are compared with the "true" parameters from the corresponding model simulations. Although footprint-scale (14 km) estimates of precipitation and latent heating may be of interest in studies of storm structure, applications such as global data assimilation generally require lower-resolution estimates. Figure 1 illustrates the impact of averaging on the random error of TMI surface rain-rate and $Q_1 - Q_R$ estimates. The top panels are comparisons of true and estimated parameters at footprint scale. True and estimated parameters in the middle panels have been averaged to 28-km resolution, and parameters in the bottom panels have been averaged to 56-km resolution. Note that although the magnitudes of rain-rate and $Q_1 - Q_R$ estimates are diminished by averaging, the scatter of estimates relative to the variance of values is reduced. The correlation coefficient of rain rate increases from 0.88 to 0.92 and 0.95 as the averaging area is increased.

Although random errors in rain-rate and $Q_1 - Q_R$ estimates decrease as the effective resolution of the estimates decreases, systematic errors remain. The general trend of these errors is illustrated in Figs. 2 and 3. Plotted in the upper panel of Fig. 2 are rain-rate-weighted histograms of both the true and estimated rain rates at footprint resolution (14 km), derived from applications of the TMI algorithm to the synthetic data. The weighting of the histogram by rain rate in each histogram bin (1 mm h^{-1} bin interval) yields the relative contribution of the rain rate (along the abscissa in
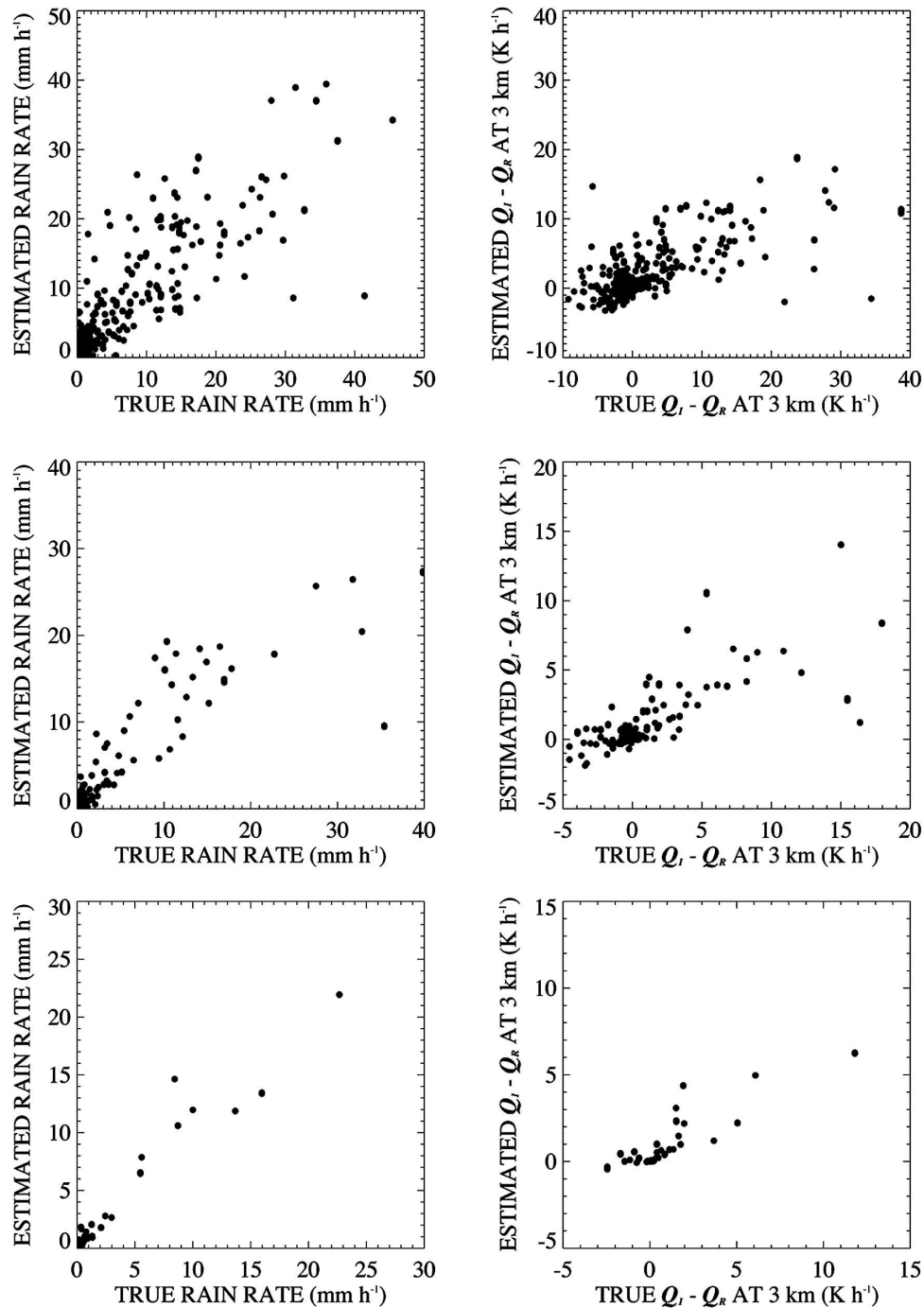
FIG. 1. Scatterplots of estimated vs (left) true rain rate and (right) $Q_1 - Q_R$ at 3-km altitude at (top) footprint scale (14 km) and at (middle) 28- and (bottom) 56-km resolution, based upon applications of the microwave radiometer algorithm to synthetic radiance data.

the figure) to the mean rain rate. Although there are obvious differences between the histograms of true and estimated rain rates, there are no apparent systematic differences.

In the lower panel of Fig. 2, the bias-weighted histo-

gram of true rain rates is plotted. Here, the plotted points can be interpreted as the contributions of the differences between estimated and true rain rates to the total bias, binned in 1 mm h^{-1} intervals of the true rain rate (plotted on the abscissa). The Bayesian formula-
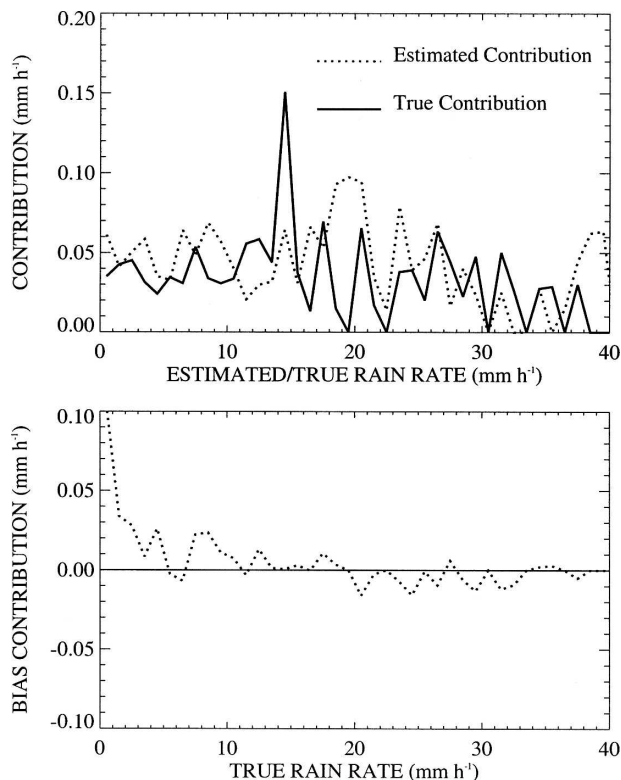
FIG. 2. (top) Contributions of estimated and true rain rates to their mean values, and (bottom) contribution of estimated minus true rain rates to the total bias, plotted vs the true rain rate, based upon applications of the microwave radiometer algorithm to synthetic radiance data.
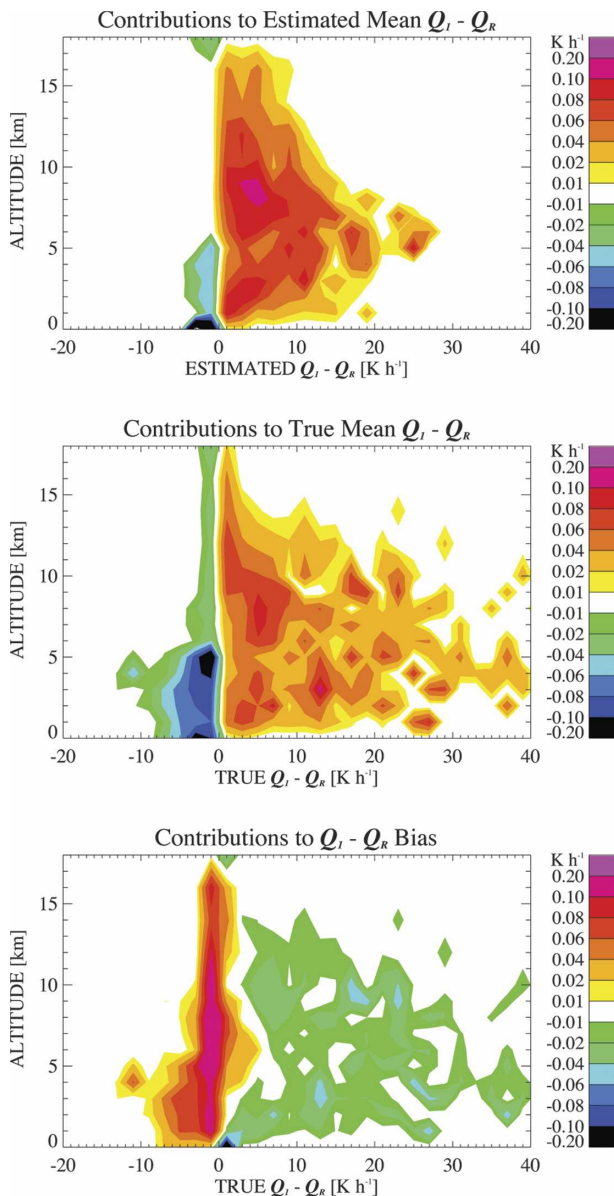

FIG. 3. Contributions of (top) estimated and (middle) true $Q_1 - Q_R$ to their mean values at different altitudes, plotted as functions of the estimated and true $Q_1 - Q_R$, respectively. (bottom) The contribution of estimated minus true $Q_1 - Q_R$ to the total bias is plotted as a function of the true $Q_1 - Q_R$ at different altitudes. Estimates are based upon applications of the microwave radiometer algorithm to synthetic radiance data.

tion of the TMI algorithm is designed to yield an unbiased rain-rate estimate given a set of observations, and the total bias is only 0.016 mm h^{-1}, or 0.9% of the true mean rain rate. However, significant biases may occur for certain subpopulations. It may be noted from Fig. 2 that lower rain rates tend to be overestimated while higher rain rates are underestimated. The trend of over-/underestimation is a consequence of the ambiguity of the input radiance data, which do not uniquely specify a particular rain rate. Given this ambiguity, the algorithm provides a rain estimate that is roughly the "average" of all rain rates in the supporting database that are consistent with the input observations. In the range of lower rain rates, the algorithm averages light rains in the database consistent with the input observations, but the distribution of these light rains is limited by the physical constraint that rain rate must be greater than or equal to zero. The algorithm's averaging over this asymmetric distribution leads to a positive bias of the rain-rate estimate. In the range of higher rain rates, the algorithm's average of rains consistent with the input observations tends to favor less intense rain rates,

which occur more frequently than rains of greater intensity in the supporting database. The result is a negative bias of rain estimates in this range.

An analysis of $Q_1 - Q_R$ estimates based upon synthetic data leads to similar bias trends. Presented in the panels of Fig. 3 are the contributions to the mean estimated and true $Q_1 - Q_R$, as well as the contributions to the bias of $Q_1 - Q_R$, in bin intervals of 2 K h^{-1} along
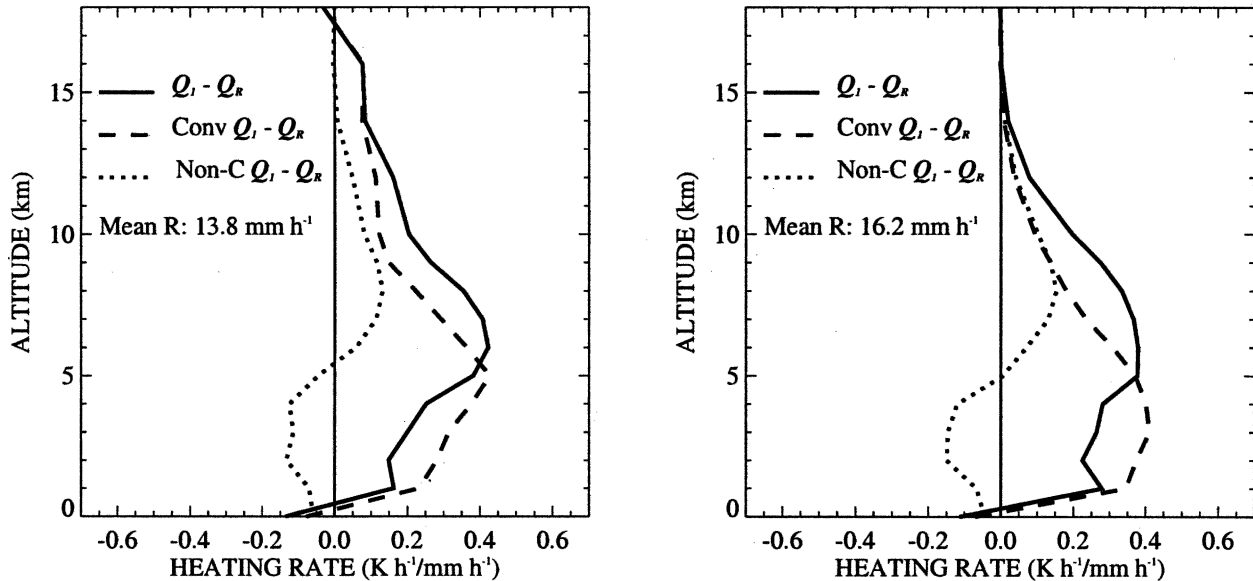
FIG. 4. Vertical profiles of the mean (left) estimated and (right) true $Q_1 - Q_R$, based upon applications of the microwave radiometer algorithm to synthetic radiances, derived from a cloud-radiative model simulation of Hurricane Bob (1991). Plotted are the mean convective, nonconvective, and total $Q_1 - Q_R$ profiles over the entire simulation domain. Mean estimated and true surface rain rates are also indicated.

the abscissa and at all altitudes where algorithm estimates are made (0–18 km). Note from the mean contribution plots that the range of estimated $Q_1 - Q_R$ values is more limited than the true range. The contributions to $Q_1 - Q_R$ bias indicate overestimation of weak heating rates, while stronger heating and cooling rates are systematically underestimated. A consequence of these bias trends is illustrated in Fig. 4. Plotted in the figure are the mean vertical profiles of $Q_1 - Q_R$ and the convective and nonconvective contributions to $Q_1 - Q_R$ from applications of the algorithm to the synthetic Hurricane Bob data (see Table 2). For comparison, the true mean $Q_1 - Q_R$ profiles from Hurricane Bob are also plotted, and in each case the $Q_1 - Q_R$ profiles have been normalized by the mean rain rate to help isolate differences in profile shape. Because the domain of the Bob simulation covers only the inner core of the storm, strong convective heating is expected to dominate the vertical profile of $Q_1 - Q_R$, and indeed this is indicated by the true mean profiles. The underestimation of extreme heating by the algorithm, however, leads to a low bias of estimated convective and total $Q_1 - Q_R$ in the lower troposphere. Mean rain rates are similarly underestimated.

In the foregoing examples, deficient information contained in upwelling microwave radiances was shown to lead not only to significant random errors at the scale of microwave footprints, but also to biases in estimates of rain rate and latent heating. It should be noted that

these biases occur in spite of the fact that the physical models incorporated in the algorithm are consistent with the true precipitation/heating distributions and radiative transfer calculations utilized in the creation of the synthetic data. In applications of the TMI algorithm to actual TMI radiance data, errors in the physical models will lead to additional error. However, the focus of the present work is on the random error associated with nonspecific information in the microwave data, and how this error propagates to time/space averages of TMI estimates. Described in the next subsection are methods for extracting information from the TMI algorithm itself to help estimate random errors in specific applications of the algorithm.

b. Algorithm-based methods

In section 2b, the Bayesian formulation of the TMI algorithm was used to derive an estimate of the random error (3) in footprint-scale-retrieved parameters resulting from deficient information contained in the microwave radiance data. Although additional random error in TMI estimates can result from errors in the algorithm's supporting database of cloud-radiative model simulations, relation (3) provides a useful "building block" for the derivation of errors in time–space average precipitation or latent heating estimates.

The general magnitude of errors in footprint-scale rain-rate estimates is illustrated by Fig. 5, which is based upon applications of (1) and (3) to seven orbits of
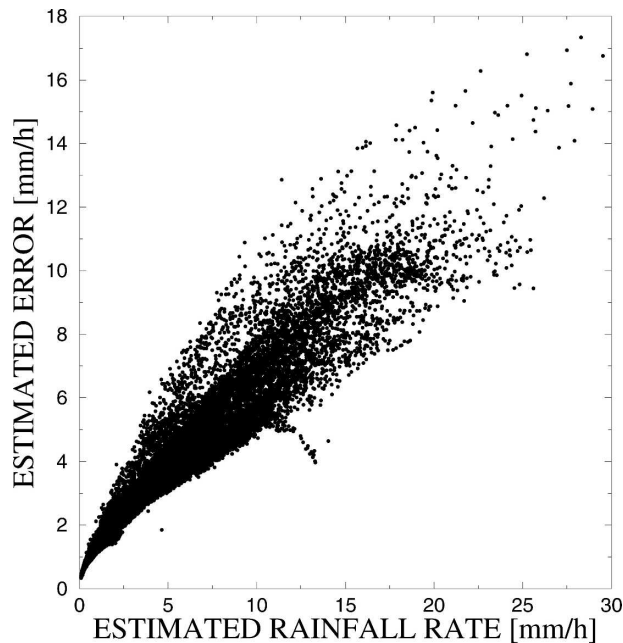
FIG. 5. Algorithm-based estimates of random error in footprint-scale rain rates based upon microwave radiometer algorithm applications to a subset of July 2000 TMI observations.

TMI observations (ocean footprints only) spanning the month of July 2000. Note that the errors resulting from deficiencies in radiance information increase with estimated rainfall rate, while percentage (relative) errors decrease. For example, the random error in a 1 mm h^{-1} rain-rate estimate is approximately 100%, but the error decreases to about 60% at 20 mm h^{-1}. These results are similar to the findings previously reported by Bauer et al. (2002), who described algorithm-based errors from the version-5 TMI algorithm applied to observations of Supertyphoon Paka. The value of these error estimates is that they reflect not only the dependence of errors on rain intensity but also the dependence on other environmental conditions specific to a given radiometer observation. Note from Fig. 5 that the range of errors corresponding to a given rain-rate estimate increases as the rain intensity increases.

c. Errors in instantaneous half-degree rain rates

It was noted previously that in applications such as data assimilation, TMI estimates at footprint resolution (14 km) are not generally required. In such applications, estimates at 0.5° or 1.0° spatial resolution are consistent with the grid resolution of a prospective analysis. In this subsection, the algorithm-based estimates of footprint-scale error are propagated to 0.5° resolution and are compared with TMI–PR differences.

If the simple spatial average of N TMI footprint-scale estimates over a given area is

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i, \qquad (7)$$

then the error variance of the spatial average is

$$\sigma_{\bar{x}}^2 = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sigma_i \, \sigma_j \, r_{ij}, \qquad (8)$$

where σ_i and σ_j are the errors of the ith and jth footprint-scale estimates, respectively, and r_{ij} is the correlation of the errors of the estimates. Note that although the TMI footprints are nearly contiguous, there is greater spatial sampling cross track than down track in the instrument swath, which could result in additional random error in the spatial average.

Given the algorithm-based estimates of footprint-scale random error given by (3), the error variance of the spatial average [(8)] can be evaluated if the spatial correlation of errors r can be estimated. In Bauer et al. (2002), it was assumed that the correlation of errors could be approximated by the correlation of retrieved parameters themselves. In the current study, the synthetic precipitation and radiance data are used to evaluate not only the correlation of surface rain rates, but also the correlation of rain-rate errors, which are derived from applications of the TMI algorithm to the synthetic radiances (see section 4a). Presented in Fig. 6 are correlations of surface rain rate and rain-rate error, plotted as functions of footprint separation distance. The correlations, given by the short- and long-dashed curves in the figure, are based upon all possible footprint pairs in the synthetic data, such that each pair is contained in the same 55 km × 55 km rectangular box. Note that although both curves exhibit the familiar quasi-exponential decrease in correlation with footprint separation distance (e.g., Bell et al. 1990, their Fig. 3), the correlation of errors decreases much more rapidly with distance than the correlation of rain rates.

Furthermore, if the synthetic data are also stratified by the average estimated rain rate in each 55 km × 55 km box, an additional set of error correlation curves is produced. Although there are insufficient synthetic data to determine the dependence of error correlation on box-averaged rain rate, the mean correlation and standard deviation values, also plotted in Fig. 6, indicate a relatively distinct family of curves. The mean error correlation, conditioned on box-averaged rain rate, decreases even more rapidly with footprint sepa-
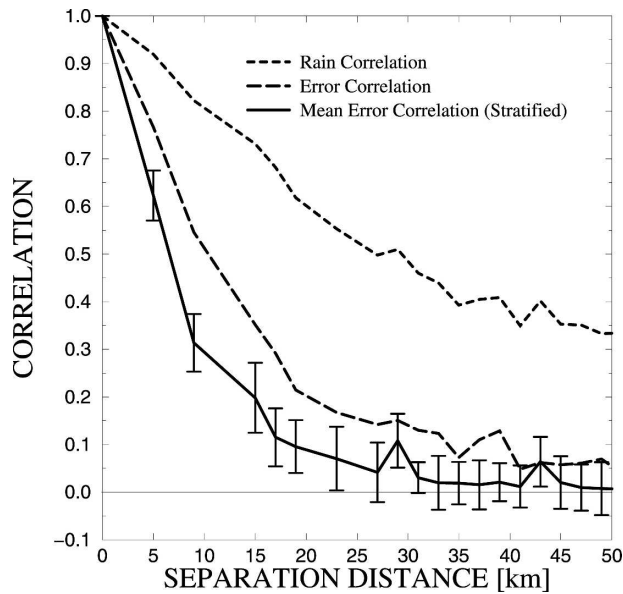
FIG. 6. Correlations of estimated rain rates and rain-rate errors for TMI footprints separated by distances plotted on the abscissa. The solid curve indicates the mean error correlation (at the specified distance) for data stratified by instantaneous mean rain rate in $0.5° \times 0.5°$ grid boxes; bars indicate the standard deviation of correlations at the specified distance. Please see text for a complete description.
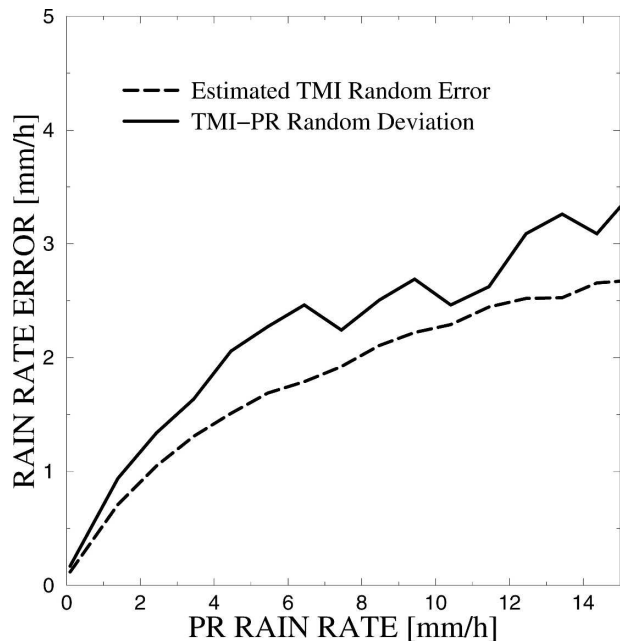

FIG. 7. Estimated mean random errors in TMI instantaneous rain rates at 0.5° resolution, plotted vs collocated PR rain-rate estimates (dashed line). The TMI–PR random deviation is plotted for comparison (solid line). The statistics are based upon TMI and PR observations over the ocean from July 2000.

ration distance than does the unconditional error correlation.

If the mean correlation function based upon the rainfall-stratified data is incorporated in (8), and the footprint-scale uncertainties are derived from (3), then (8) can be evaluated from applications of the TMI algorithm to the radiance data. As a test of (8), the algorithm is applied to all over-ocean TMI observations from the month of July 2000. The TMI half-degree instantaneous rain rates and their estimated errors are then collocated with half-degree instantaneous PR rain rates; only TMI–PR pairs for which both instruments observed at least 90% of the same half-degree box are included in the analysis. The TMI–PR pairs are then binned by the PR half-degree rain rate in 1 mm h^{-1} intervals. Based upon the TMI–PR pairs in each rain interval, the mean half-degree TMI rain-rate error from (8) is plotted against the mean half-degree PR rain rate (see Fig. 7). For comparison, the standard deviation of the TMI–PR differences is also plotted for each rain-rate interval.

In comparing the half-degree-resolution TMI rain-rate random errors in Fig. 7 with the errors at footprint scale (Fig. 5), it is apparent that in both cases, errors increase with increasing rain rate, but percentage errors decrease with rain rate. It is difficult to compare the plots in a quantitative way, because the distribution of rain rates narrows with increasing averaging area; however, the minimum percentage error at half-degree resolution is ~20%, as compared with ~60% at footprint scale. Therefore, averaging appears to significantly reduce the random error of rain estimates, even considering the footprint-scale error spatial correlations that would tend to limit such a reduction. The larger minimum percentage error of ~40%, reported by Bauer et al. (2002) for version-5 TMI estimates averaged to 60-km resolution, is due to the greater error correlations assumed in that study. Bauer et al. (2002) assumed that rain-rate and error correlations were equal, leading to a likely overestimate of error correlations.

Based upon applications of the algorithm to TMI observations over ocean from July 2000, random errors in instantaneous 0.5° rain-rate estimates range from roughly 50% at 1 mm h^{-1}, to about 20% at 14 mm h^{-1}. TMI errors also appear to explain about 60%–80% of the variance of the TMI–PR differences; see Fig. 7. The residual 20%–40% is likely because of random errors in PR rain estimates and differences in spatial sampling of half-degree boxes by the TMI and PR. This implies that the standard deviation of PR random errors (plus spatial sampling differences) is roughly 50%–80% of the TMI error standard deviation.

d. Random errors in monthly mean 2.5°-resolution estimates

In applications of satellite rain-rate and latent heating estimates to large-scale analysis or climate studies, data are often averaged to coarse spatial resolution and then over monthly or longer time periods (e.g., Xie and Arkin 1997; Adler et al. 2003). As the basis of a preliminary analysis of the uncertainties in longer time-/space-averaged precipitation/heating estimates, the random errors in monthly mean 2.5°-resolution TMI-derived estimates are examined in this subsection. Also, as in previous work (Bell et al. 1990; Li et al. 1998; Bell and Kundu 2000, hereinafter BK00), it will be assumed that the errors resulting from the infrequent sampling of low-Earth-orbiting sensors, such as TMI, are independent of algorithm errors. It follows that the total error variance of the time-averaged estimate is equal to the sum of the algorithm and sampling error variances,

$$\sigma^2_{\langle \overline{P} \rangle} = \sigma^2_{\text{alg}} + \sigma^2_{\text{samp}}, \tag{9}$$

where $\overline{P}$ is the instantaneous area-averaged rain rate in a 2.5° × 2.5° box, and the brackets $\langle \rangle$ indicate a monthly average. The accumulated algorithm error variance σ^2_{alg} over the period of a month is evaluated using (8), but in this case the summations are extended over all footprint-scale rain estimates in a given 2.5° × 2.5° box over the month.

Regarding the sampling error σ^2_{samp}, it is assumed that the estimate of the monthly mean rain rate over a given 2.5° × 2.5° box is the area-weighted average,

$$\langle \overline{P} \rangle = \frac{1}{S} \sum_{i=1}^{M} \left(\frac{A_i}{A} \right) \overline{P}_i, \tag{10}$$

where the "effective number of visits" S is given by

$$S \equiv \sum_{i=1}^{M} \left(\frac{A_i}{A} \right). \tag{11}$$

Here, A is the area enclosed by the 2.5° × 2.5° box, A_i and $\overline{P}_i$ are the area of observation within the box and the mean estimated rain rate within that area, respectively, on the ith overpass, and the summations are over all full or partial observations of the box during the month.

The error in $\langle \overline{P} \rangle$ resulting from the relatively infrequent sampling of rainfall by low-Earth-orbiting satellite radiometers has been studied by several investigators, for example, Laughlin (1981), Shin and North (1988), Bell et al. (1990), Kedem et al. (1990), Oki and Sumi (1994), Huffman (1997), Li et al. (1998), and Steiner et al. (2003). Investigations by Bell and Kundu

(1996) and BK00 have demonstrated the applicability of a relatively simple analytical model to the sampling problem. Incorporating the uniform sampling approximation of Laughlin (1981) in this model, BK00 derived an expression for the percentage sampling error in monthly mean estimates $\sigma_{\text{samp}}/\langle \overline{P} \rangle$ in terms of the variance of the 2.5° instantaneous rain estimates σ^2_A, the autocorrelation time τ_A of instantaneous box-averaged rain rates $\overline{P}$, separated by an interval T/S, where the period of observation T (1 month) divided by the effective number of satellite visits S has been used to approximate the sampling time interval (Δt) in the original expression from BK00 [see (2.22) and (2.23) in BK00].

Based upon a limited number of radar and rain gauge studies, BK00 argued that the percentage sampling error varied approximately as $\langle \overline{P} \rangle^{-1/2}$, but with a coefficient of proportionality that varied with geographic location. In the present study, the BK00 expression for the percentage error in monthly estimates is evaluated with the help of the TMI rain-rate estimates themselves. The monthly mean rain rate is estimated using (10), while S is given by (11), and σ^2_A is approximated by the variance of the instantaneous 2.5° TMI rain estimates. The autocorrelation time τ_A is typically much shorter (several hours) than the effective sampling time interval T/S of a single low-Earth-orbiting sensor (~1 day), and so the autocorrelation time cannot be estimated from the sensor observation time series. As an alternative, the expression from Bell et al. (1990) that was derived from GATE radar observations,

$$\tau_A = 0.394(\sqrt{A})^{0.525}, \tag{12}$$

is employed, where τ_A is in hours and A is in squared kilometers. BK00 reported autocorrelation times as short as 4 h over 2.5° × 2.5° boxes during TOGA COARE, while autocorrelation times as long as 8 h (14 h) over 5° × 5° boxes in southern Japan during winter (summer) were found by Oki and Sumi (1994). Therefore, the autocorrelation time associated with a specific region is dependent not only on box area but also the climatology of rain systems in that region, and the use of (12) is an obvious simplification.

Shown in the left panels of Fig. 8 are estimates of the monthly mean surface rain rate, convective rain proportion, and $Q_1 - Q_R$ at 7- and 3-km altitudes, based upon all TMI over-ocean observations from July 2000. The main features seen in the surface rain map correspond to those seen in global climatologies (e.g., Adler et al. 2003). Note the relative minima of convective proportion along the intertropical convergence zone
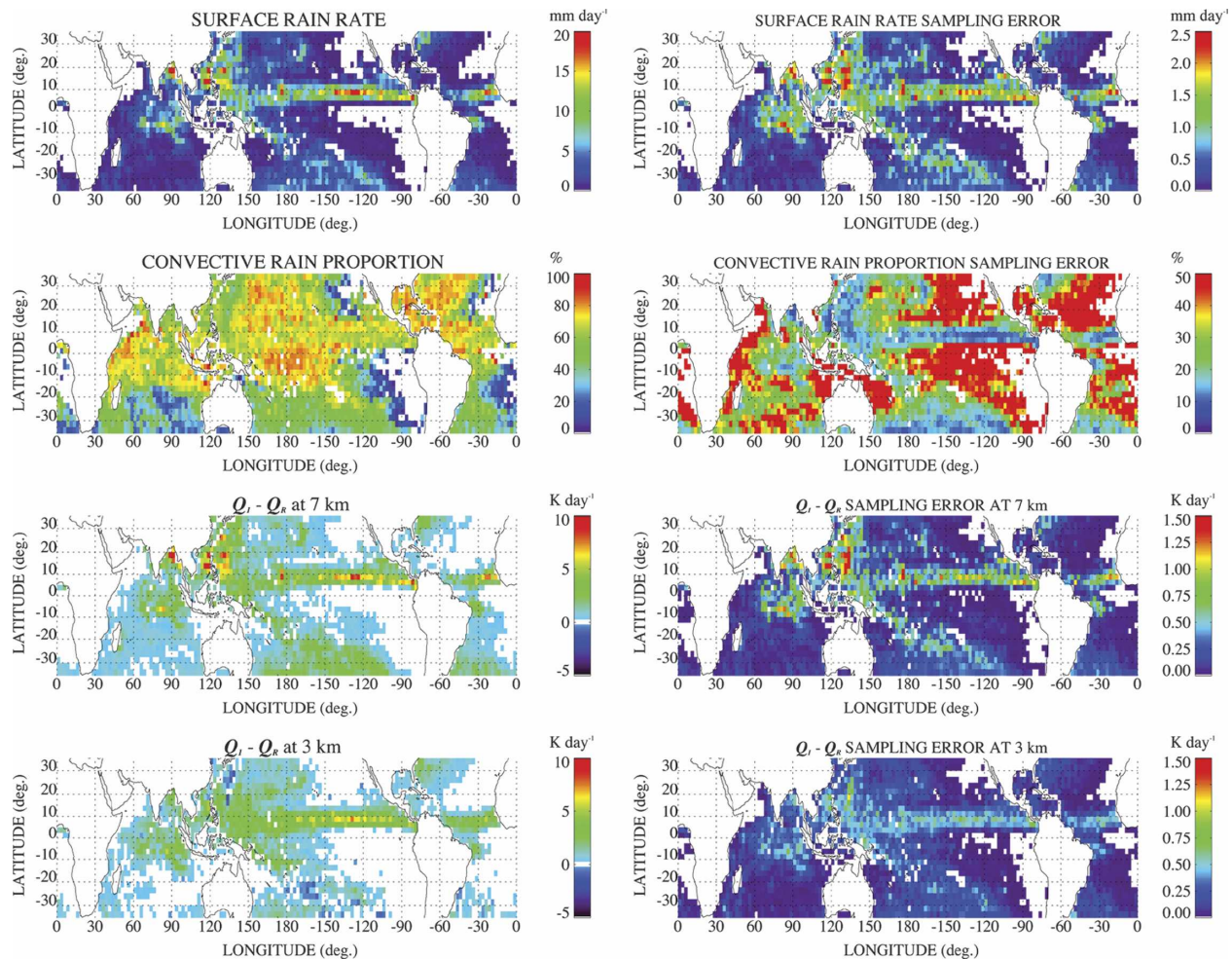
FIG. 8. (left) Monthly mean 2.5°-resolution (top) surface rain rates, (middle) convective rain proportions, and (bottom) $Q_1 - Q_R$ at altitudes of 7 and 3 km, derived from algorithm applications to TMI observations over ocean from July 2000. (right) Corresponding sampling errors in monthly mean 2.5° estimates.

(ITCZ) and in the more intense rain regions of the western Pacific and Indian Ocean. These relative minima indicate a significant contribution to the total rainfall by organized mesoscale convective systems, as described by Rickenbach and Rutledge (1998) in their analysis of radar observations from the TOGA COARE field campaign (located in the western Pacific warm-pool region).

Latent heating distributions, to first order, follow the patterns of surface rain rate, because the vertically integrated heating is approximately equal to $L_v P$, where L_v is the latent heat of vaporization and P is the precipitation rate. Note that in the more intense rain regions, the heating at 7-km altitude exceeds that at 3 km. Where organized mesoscale convective systems produce a large proportion of the rainfall, the contribution of stratiform rains to total rainfall is significant, and the altitude of maximum latent heating rate is generally

elevated; see Lin and Johnson (1996). However, in regions of weaker rains in the Northern Hemisphere, the rain spectrum is predominantly convective, and the heating at 3 km is often greater than that at 7 km. In the northern fringe of the Southern Hemisphere storm tracks, large baroclinic systems dominate and largely stratiform rains produce a characteristic weak heating at upper levels and evaporative cooling below.

Presented in Fig. 9 is a comparison of algorithm and sampling errors plotted as functions of the monthly, 2.5°-resolution rain-rate estimates. Although algorithm random errors are not entirely negligible at the monthly scale at the prescribed resolution, they are relatively small—less than 6% of the monthly total for rain rates greater than 5 mm day^{-1}. Because of the small algorithm random error contribution to the monthly rain error, the remainder of this section will focus on sampling errors. Note, however, that it has so far been as-
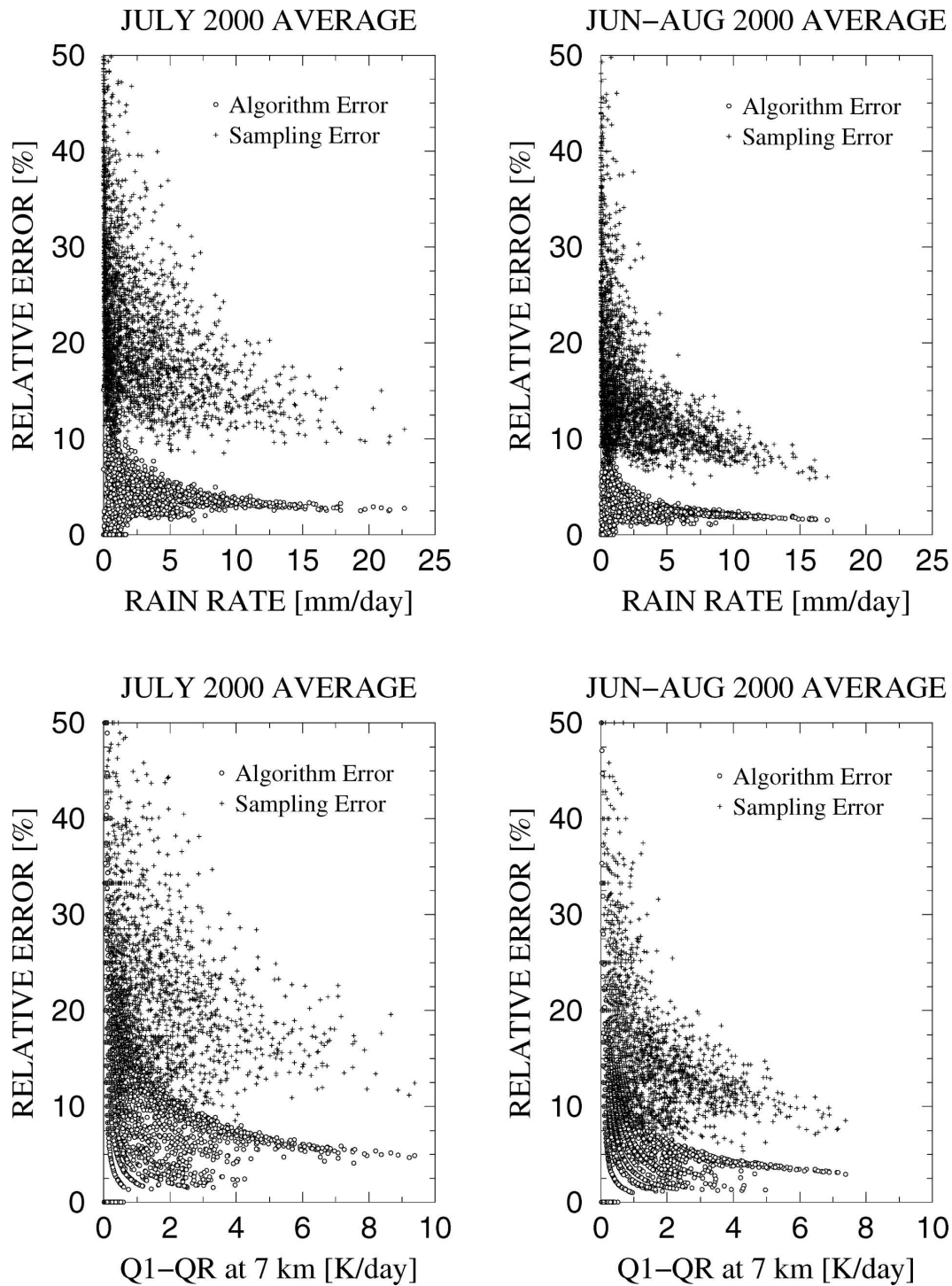
FIG. 9. Estimated percentage errors in (top left) monthly mean 2.5°-resolution surface rain rates and (bottom left) $Q_1 - Q_R$ at 7-km altitude resulting from algorithm random errors and incomplete temporal sampling, plotted vs their estimated values. Monthly statistics are based upon applications of the microwave radiometer algorithm to TMI observations over ocean from July 2000. Also plotted are (top right) algorithm and sampling errors in June–August 2000 mean surface rain rates and (bottom right) $Q_1 - Q_R$ at 7-km altitude at 2.5° resolution.

sumed in this theoretical development that algorithm errors are random with zero mean; in Part II, algorithm errors are shown to have an additional systematic component that is significant.

Distributions of sampling error in monthly, 2.5°-resolution surface rain rates, convective rain proportions, and $Q_1 - Q_R$ at 7- and 3-km altitudes are provided in the right panels of Fig. 8. The sampling error of surface rain rates is computed using BK00 [their (2.22), (2.23)]; the sampling errors of convective rain rate and $Q_1 - Q_R$ are calculated by substituting the variances of these quantities for σ_A^2 in BK00's expression and assuming that their autocorrelation times are the same as the rain-rate autocorrelation time. Note that the sampling errors for rain rate and $Q_1 - Q_R$ closely follow the patterns of rain rate and heating rate; however, the percentage errors decrease with increasing rain rate (see Fig. 9). This behavior of the sampling error was noted by Chang et al. (1993), BK00, and others. The range of sampling error is roughly 8%–35% at a rain rate of 5 mm day^{-1}, and it tapers to about 12% at 20 mm day^{-1}. The large spread of sampling error at low rain rates is mainly because of geographic variations of the variance of rain rates and the frequency of sampling by the TMI; even though rain rates are relatively low at higher latitudes, the generally low rain-rate variance and relatively high sampling rate tend to reduce percentage errors, according to (2.22) and (2.23) of BK00. Convective rain proportion—the ratio of the convective and total rain rate—has a different sampling error distribution, with lowest errors in regions of high rain rate (see Fig. 8). Aside from variations in sampling error resulting from error contributions from rain rate and convective rain-rate estimates, taking the ratio of the two quantities introduces an approximate $\langle \overline{P} \rangle^{-1}$ dependence of convective rain proportion sampling error on rain rate. Errors in monthly latent heating estimates follow the same trends as rain-rate errors, although they are proportionately greater, in general (see lower panels of Fig. 9). Also, latent heating algorithm errors cannot be neglected in relation to sampling errors.

Sampling error can be reduced further by taking longer-term averages of the satellite estimates. Shown in the right-hand panels of Fig. 9 are the algorithm and sampling random errors of estimated June–August 2000 average rain rates and $Q_1 - Q_R$ at 7-km altitude. Even considering the reduction of the variance of 3-month-averaged rain rates relative to the 1-month averages, a significant reduction of sampling error in the 3-month averages is evident, with a range of 6%–15% at an average rain rate of 5 mm day^{-1}. Latent heating errors are similarly reduced.

5. Summary and concluding remarks

A revision of the TRMM facility precipitation algorithm for applications to TMI observations (version 6) is described. The primary objective is to adapt this algorithm for the estimation of consistent convective rain proportion and cloud latent heating ($Q_1 - Q_R$) profiles, in addition to surface rainfall rate and precipitation profiles. The extension and generalization of the algorithm are accomplished by (a) increasing the diversity of cloud-radiative model simulations supporting the algorithm, (b) adjusting the cloud-resolving model microphysics to produce more realistic graupel and snow distributions, (c) including the effects of mixed-phase precipitation in nonconvective regions of the cloud-resolving model simulations, (d) utilizing consistent definitions of convective rain area and total rain area in the context of cloud-resolving model simulations and satellite observations, (e) including a new geographic database to better separate ocean, coast, and land areas, and (f) establishing a more consistent microwave radiance "background" in rain-free regions from passive microwave observations. Uncertainties in retrieved parameters are estimated by applying the algorithm to synthetic radiance data based upon a subset of cloud-radiative model simulations and through construction, starting with algorithm-based error estimates and propagating these to coarser time and space resolutions. Synthetic data applications indicate suppression of random errors with averaging, although systematic overestimation (underestimation) of the lowest (highest) rain intensities and latent heating rates is also indicated. Biases are attributed to the relatively small precipitation/heating "signal" in passive microwave observations at both extremes.

The propagation of algorithm-based random errors is computed to provide baseline uncertainties in two relevant products—an instantaneous 0.5°-resolution product suitable for data assimilation applications, and a monthly 2.5° product required for large-scale analyses or climate studies. Based upon applications of the algorithm to TMI observations over ocean from July 2000, random errors in instantaneous 0.5° rain-rate estimates ranged from roughly 50% at 1 mm h^{-1}, to about 20% at 14 mm h^{-1}. Errors in collocated PR estimates are roughly 50%–80% of the TMI errors at this resolution. Random errors in monthly 2.5° rain-rate and heating estimates are due to the combination of algorithm and sampling errors. Sampling errors are estimated using the Laughlin (1981) model, in which the variance of rain rate or heating rate over the month is supplied by the TMI estimates themselves. Although rain-rate algorithm errors are not negligible (up to 6%

at a rain rate of 5 mm day^{-1}), sampling errors dominate at this time–space resolution. Sampling errors range from 8% to 35% at 5 mm day^{-1}, but decrease with increasing rain rate. Sampling errors in latent heating rates follow the same trend. Averaging over 3 months reduces the rain-rate sampling error to a range of 6%–15% at 5 mm day^{-1}. Latent heating errors are similarly reduced.

It should be stressed that the algorithm-based error estimates described in this study represent only that portion of the random error associated with lack of precipitation or latent heating–specific information in the passive microwave observations. Two other sources of uncertainty are the systematic errors in the cloud-radiative model simulations that form the supporting database of the current algorithm and the lack of representativeness of the database for algorithm applications to a given region, climate regime, or atmospheric state, in general. Nevertheless, the comparisons of the algorithm-based error estimates and TMI–PR deviations suggest that algorithm random errors make up a significant portion of the total error at shorter time and space scales. As time and space averaging of the microwave precipitation/heating estimates increase, random errors resulting from the lack of information in the radiance data or insufficient sampling will decrease, exposing systematic errors associated with the cloud-radiative model database.

Errors in precipitation and latent heating estimates are examined further in Part II of this study, in which version-6 and version-5 TMI estimates are compared with independent ground-based observations, as well as those from the PR.

REFERENCES

Adler, R. F., and Coauthors, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–present). *J. Hydrometeor.,* **4,** 1147–1167.

Bauer, P., J.-F. Mahfouf, W. S. Olson, F. S. Marzano, S. Di Michele, A. Tassa, and A. Mugnai, 2002: Error analysis of TMI rainfall estimates over ocean for variational data assimilation. *Quart. J. Roy. Meteor. Soc.,* **128,** 2129–2144.

Bell, T. L., and P. K. Kundu, 1996: A study of the sampling error in satellite rainfall estimates using optimal averaging of data and a stochastic model. *J. Climate,* **9,** 1251–1268.

——, and ——, 2000: Dependence of satellite sampling error on monthly averaged rain rates: Comparison of simple models and recent studies. *J. Climate,* **13,** 449–462.

——, A. Abdullah, R. L. Martin, and G. R. North, 1990: Sampling errors for satellite-derived tropical rainfall: Monte Carlo study using a space-time stochastic model. *J. Geophys. Res.,* **95,** 2195–2205.

Chang, A. T. C., L. S. Chiu, and T. T. Wilheit, 1993: Random errors of oceanic monthly rainfall derived from SSM/I using probability distribution functions. *Mon. Wea. Rev.,* **121,** 2351–2354.

Curtis, S., and R. Adler, 2000: ENSO indices based on patterns of satellite-derived precipitation. *J. Climate,* **13,** 2786–2793.

Dudhia, J., 1993: A nonhydrostatic version of the Penn State–NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front. *Mon. Wea. Rev.,* **121,** 1493–1513.

Hong, Y., C. D. Kummerow, and W. S. Olson, 1999: Separation of convective and stratiform precipitation using microwave brightness temperature. *J. Appl. Meteor.,* **38,** 1195–1213.

Hou, A. Y., S. Q. Zhang, A. M. da Silva, and W. S. Olson, 2000: Improving assimilated global data sets using TMI rainfall and columnar moisture observations. *J. Climate,* **13,** 4180–4195.

——, ——, ——, ——, C. D. Kummerow, and J. Simpson, 2001: Improving global analysis and short-range forecast using rainfall and moisture observations derived from TRMM and SSM/I passive microwave instruments. *Bull. Amer. Meteor. Soc.,* **82,** 659–679.

——, ——, and O. Reale, 2004: Variational continuous assimilation of TMI and SSM/I rain rates: Impact on GEOS-3 hurricane analyses and forecasts. *Mon. Wea. Rev.,* **132,** 2094–2109.

Houze, R. A., Jr., 1989: Observed structure of mesoscale convective systems and implications for large-scale heating. *Quart. J. Roy. Meteor. Soc.,* **115,** 425–461.

Huffman, G. J., 1997: Estimates of root-mean-square random error for finite samples of estimated precipitation. *J. Appl. Meteor.,* **36,** 1191–1201.

Johnson, D. E., W.-K. Tao, J. Simpson, and C.-H. Sui, 2002: A study of the response of deep tropical clouds to large-scale thermodynamic forcings. Part I: Modeling strategies and simulations of TOGA COARE convective systems. *J. Atmos. Sci.,* **59,** 3492–3518.

Kedem, B., L. S. Chiu, and G. R. North, 1990: Estimation of mean rain rate: Application to satellite observations. *J. Geophys. Res.,* **95,** 1965–1972.

Krishnamurti, T. N., and Coauthors, 2001: Real-time multianalysis-multimodel superensemble forecasts of precipitation using TRMM and SSM/I products. *Mon. Wea. Rev.,* **129,** 2861–2883.

Kummerow, C., 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.

Laughlin, C. R., 1981: On the effect of temporal sampling on the observation of rainfall. Precipitation Measurements from Space Workshop Report, NASA Publication, D-59–D-66.

Li, Q., R. Ferraro, and N. Grody, 1998: Detailed analysis of the error associated with the rainfall retrieved by the NOAA/NESDIS SSM/I algorithm. 1. Tropical oceanic rainfall. *J. Geophys. Res.,* **103,** 11 419–11 427.

Lin, X., and R. H. Johnson, 1996: Heating, moistening, and rainfall over the Western Pacific Warm Pool during TOGA COARE. *J. Atmos. Sci.,* **53,** 3367–3383.

Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.,* **22,** 1065–1092.

Oki, R., and A. Sumi, 1994: Sampling simulation of TRMM rainfall estimation using radar–AMeDAS composites. *J. Appl. Meteor.,* **33,** 1597–1608.

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

——, P. Bauer, N. F. Viltard, D. E. Johnson, W.-K. Tao, L. Liao, and R. Meneghini, 2001a: A melting layer model for passive/active microwave remote sensing applications. Part I: Model formulation and comparison with observations. *J. Appl. Meteor.,* **40,** 1145–1163.

——, ——, C. D. Kummerow, Y. Hong, and W.-K. Tao, 2001b: A melting layer model for passive/active microwave remote sensing applications. Part II: Simulation of TRMM observations. *J. Appl. Meteor.,* **40,** 1164–1179.

——, Y. Hong, C. D. Kummerow, and J. Turk, 2001c: A texture-polarization method for estimating convective-stratiform precipitation area coverage from passive microwave radiometer data. *J. Appl. Meteor.,* **40,** 1577–1591.

Petty, G. W., 1994a: Physical retrieval of over-ocean rain rate from multichannel microwave imagery. Part I: Theoretical characteristics of normalized polarization and scattering indices. *Meteor. Atmos. Phys.,* **54,** 79–100.

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

Pruppacher, H. R., and J. D. Klett, 1997: *Microphysics of Clouds and Precipitation*. 2d ed. Kluwer Academic, 954 pp.

Rasmussen, E. M., and P. A. Arkin, 1993: A global view of large-scale precipitation variability. *J. Climate,* **6,** 1495–1522.

Rickenbach, T. M., and S. A. Rutledge, 1998: Convection in TOGA COARE: Horizontal scale, morphology, and rainfall production. *J. Atmos. Sci.,* **55,** 2715–2729.

Rutledge, S. A., and P. V. Hobbs, 1984: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude clouds. Part XII: A diagnostic modeling study of precipitation development in narrow cold frontal rainbands. *J. Atmos. Sci.,* **41,** 2949–2972.

Shin, K.-S., and G. R. North, 1988: Sampling error study for rainfall estimate by satellite using a stochastic model. *J. Appl. Meteor.,* **27,** 1218–1231.

Simpson, J., R. F. Adler, and G. R. North, 1988: A proposed Tropical Rainfall Measuring Mission (TRMM) satellite. *Bull. Amer. Meteor. Soc.,* **69,** 278–295.

Steiner, M., T. L. Bell, Y. Zhang, and E. F. Wood, 2003: Comparison of two methods for estimating sampling-related uncertainty of satellite rainfall averages based on a large radar dataset. *J. Climate,* **16,** 3759–3778.

Tao, W.-K., 2003a: Goddard Cumulus Ensemble (GCE) model: Application for understanding precipitation processes. *Cloud Systems, Hurricanes, and the Tropical Rainfall Measuring Mission (TRMM): A Tribute to Dr. Joanne Simpson, Meteor. Monogr.*, No. 51, Amer. Meteor. Soc., 103–138.

——, 2003b: Regional-scale modeling at NASA Goddard Space Flight Center. *Recent Dev. Atmos. Sci.,* **2,** 1–52.

——, and J. Simpson, 1993: The Goddard Cumulus Ensemble Model. Part I: Model description. *Terr. Atmos. Oceanic Sci.,* **4,** 19–54.

——, S. Lang, J. Simpson, and R. Adler, 1993: Retrieval algorithms for estimating the vertical profiles of latent heat release. *J. Meteor. Soc. Japan,* **71,** 685–700.

Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model output. *Bull. Amer. Meteor. Soc.,* **78,** 2539–2558.

Xu, K.-M., 1995: Partitioning mass, heat, and moisture budgets of explicitly simulated cumulus ensembles into convective and stratiform components. *J. Atmos. Sci.,* **52,** 551–573.

Yanai, M., S. Esbensen, and J.-H. Chu, 1973: Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. *J. Atmos. Sci.,* **30,** 611–627.

Yang, S., and E. A. Smith, 1999a: Moisture budget analysis of TOGA-COARE using SSM/I retrieved latent heating and large scale Q_2 estimates. *J. Atmos. Oceanic Technol.,* **16,** 633–655.

——, and ——, 1999b: Four-dimensional structure of monthly latent heating derived from SSM/I satellite measurements. *J. Climate,* **12,** 1016–1037.

——, and ——, 2000: Vertical structure and transient behavior of convective–stratiform heating in TOGA COARE from combined satellite–sounding analysis. *J. Appl. Meteor.,* **39,** 1491–1513.

——, W. S. Olson, J.-J. Wang, T. L. Bell, E. A. Smith, and C. D. Kummerow, 2006: Precipitation and latent heating distributions from satellite passive microwave radiometry. Part II: Evaluation of estimates using independent data. *J. Appl. Meteor. Climatol.,* **45,** 721–739.