10/15/04 *

# Precipitation and Latent Heating Distributions from Satellite Passive Microwave Radiometry Part I: 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, 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

*current affiliation: U.S. Army Research Laboratory, AMSRL-CI-EB,
Adelphi, Maryland

#current affiliation: Science Applications International Corporation,
General Sciences Operation
Beltsville, Maryland

To be submitted October, 2004

*Corresponding author address:* William S. Olson, NASA/GSFC, Code 912.1, Greenbelt, MD 20771, olson@agnes.gsfc.nasa.gov

## Abstract

A revised Bayesian algorithm for estimating surface rain rate, convective rain proportion, and latent heating/drying 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 non-convective 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 instantaneous rain rate estimates at 0.5° resolution range from approximately 50% at 1 mm $h^{-1}$ to 20% at 14 mm $h^{-1}$. These errors represent about 70-90% of the mean random deviation between collocated passive microwave and spaceborne radar rain rate estimates.

The cumulative algorithm error in TMI estimates at monthly, 2.5° resolution is relatively small (less than 6% at 5 mm $day^{-1}$) compared to the random error due to infrequent satellite temporal sampling (8-35% at the same rain rate). Percentage errors due to sampling decrease with increasing rain rate, and sampling errors in latent heating rates follow the same trend. Averaging over three months reduces sampling errors in rain rates to 6-15% at 5 mm $day^{-1}$, with proportionate reductions in latent heating sampling error.

## 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), Adler et al. (2000), Curtis and Adler (2000). 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, as this partitioning is linked to the organization of convective systems and the distributions of vertical motion and latent heat release in the atmosphere. 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, a knowledge of the 4-D 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. 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. They presented the 3-D global distributions of monthly mean latent heating for 1992 and discussed the evolution of the Walker circulation for the 1992 El Niño event and the Asian monsoon circulation in the context of latent heating. 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 the Tropical Rainfall Measuring Mission (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 a passive and active microwave sensors that provide measurements of the horizontal and vertical structure of precipitation in the atmosphere at a relatively high sampling rate (minimum spacing ~5 km). The aim of the present study is to examine the potential for estimating consistent precipitation, latent heating and drying 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 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 horizontal inhomogeneity of the rain field within the satellite footprint, which is critical for establishing the link between footprint-average 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 non-convective (stratiform) rain generally indicates weak mesoscale ascent and heating in the upper troposphere with descent and evaporative cooling at lower altitudes; see Houze (1989).

In the present study, the Bayesian estimation method described in Kummerow et al. (1996), Olson et al. (1996), Olson et al. (1999), and Kummerow et al. (2001) is extended to provide estimates of precipitation, latent heating and drying, 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 high surface emissivity of land surfaces; i.e., microwave emission/absorption by liquid precipitation cannot be easily distinguished from land

emission. Aside from the adaptation to TMI observations, the method is improved by (a) expanding the algorithm's cloud-resolving model database to include a greater diversity of precipitation systems, (b) making basic adjustments to the ice precipitation microphysics in the cloud-resolving model simulations 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/drying distributions that have more fidelity with climatological distributions; see Part II of this series. The Bayesian estimation method is at the heart of the current TRMM facility algorithm (2A-12, 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 due to incomplete information in the microwave data can also be derived from the Bayesian (algorithmic) method. Both synthetic retrievals and the algorithmic method are described and applied in section 4. In section 5, the capability of the complete 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 due to infrequent sampling, are examined. The paper concludes with a brief summary and reflections on the direction of future work (section 6).

## 2. Data and Basic Method

### *a. TRMM observations*

The analyses presented in this study are based upon observations, both synthesized and actual, from the TRMM Microwave Imager (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. At 85.5 GHz, measured radiances are essentially insensitive to variations in rain path for path-integrals greater than about 1 kg $m^{-2}$, but radiances can decrease by 10's of °K for modest increases in the path-integrals of snow or graupel. Due to 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 the swath imagery. The sampling resolution of the TMI is about 14 km along-track and 5 km cross-track.

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; Table 1. 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. There are basic physical limitations on the swath width of spaceborne radars; the PR swath is 215 km wide, compared to the 760 km-wide swath of the TMI. The greater swath width and smaller mass and power consumption of microwave radiometers define their role as sampling instruments, and therefore a constellation of radiometers is planned for the future Global Precipitation Measurement (GPM) mission. Spaceborne radars, which can provide high-resolution precipitation data coincident with the passive radiometers over limited areas, are expected to be the "calibrators" of the passive microwave measurements.

*b. Estimation of cloud properties from TMI*

The TMI retrieval algorithm (2A-12, 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, cloud-resolving model 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 composited from those profiles in the database which correspond to simulated radiances consistent with the observed radiances. Formally, a TMI estimate of profile parameters, $\hat{E}[\boldsymbol{x}]$, is given by

$$\hat{\mathbf{E}}[\mathbf{x}] = \sum_k \mathbf{x}_k \frac{\exp\left\{-0.5\left(\mathbf{I_S}(\mathbf{x}_k) - \mathbf{I_0}\right)^T \left(\mathbf{S_I} + \mathbf{O_I}\right)^{-1}\left(\mathbf{I_S}(\mathbf{x}_k) - \mathbf{I_0}\right) + C\right\}}{\hat{M}}, \qquad (1)$$

where the model profile vector $\boldsymbol{x_k}$ contains all parameters, including the surface rain rate, convective rain rate, liquid/ice-phase precipitation and latent heating/drying profiles, corresponding to the simulated radiance indices, $\boldsymbol{I_S(x_k)}$. The radiance indices, constructed from radiances at the different radiometer channel frequencies/polarizations, are the normalized polarization and scattering indices defined by Petty (1994). $\boldsymbol{I_O}$ is a vector of sensor observed radiance indices, similarly defined. $\boldsymbol{S_I}$ and $\boldsymbol{O_I}$ are error covariance matrices of the simulated and observed microwave radiance indices, respectively, 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 x 14 km for TMI) and the freezing level, is included in the constraint term, $C$.

$$C = -0.5\left(\boldsymbol{f_S}(\mathbf{x}_k) - \boldsymbol{f_0}\right)^T \left(\mathbf{S}_f + \mathbf{O}_f\right)^{-1}\left(\boldsymbol{f_S}(\mathbf{x}_k) - \boldsymbol{f_0}\right) \qquad (2)$$

Here, $\boldsymbol{f_S}$ is a vector of simulated constraint parameters, and $\boldsymbol{f_O}$ is a vector of corresponding observed parameters. $\boldsymbol{S_f}$ and $\boldsymbol{O_f}$ are error covariance matrices of the simulated and observed constraint parameters, respectively. A complete description of the formulation of the constraint parameters and their errors is deferred to section 3.

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, surface rainfall rate, convective rain proportion, and profiles of precipitation, latent heating and drying can be made with different degrees of accuracy.

Since, 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 a single estimated parameter using

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

which yields a measure of the uncertainty in the estimate of $x$ due to the limited information content of the observations.

The uncertainty represented by (3) would exist even if the cloud-radiative model simulations in the GPROF 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 GPROF estimates due to errors in cloud-radiative modeling may occur. Since true validation of precipitation-related quantities using independent observations is difficult, (3) at least provides a lower bound on the random error of GPROF estimates- a basic "building block" for estimates of the random error in the derived products, as described in section 4.

## 3. Modifications of the TMI algorithm

In the development of the Version 6 TMI TRMM facility algorithm, several modifications are included to allow for latent heating/drying 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/drying, the improved sensitivity of the algorithm leads to estimates of surface rainfall rate with generally less bias with respect to independent estimates; see Part II.

*a. Greater diversity in the supporting cloud-resolving model database*

In Version 5 of the TRMM facility 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, a three-dimensional, nonhydrostatic cloud-resolving model described in Tao and Simpson (1993) and Tao (2003a). Among the physical 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), in which cloud liquid, rain, cloud ice, snow, and graupel equivalent water contents are calculated prognostically. The precipitation particle size distributions are represented by inverse exponential functions with prescribed intercept values; the slopes of the exponential distributions are adjusted to account for variations in simulated water content. 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 Dec. 19-26 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 can be extracted for algorithm applications simply 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 Penn State/NCAR Mesoscale Model 5 (MM5; see Dudhia, 1993) model is utilized, but the microphysical scheme used in the MM5 simulations is again the 3ICE scheme, which has been incorporated into MM5; see Tao et al. (2003b). MM5 simulations nested to resolutions ~ 2 km are required 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/drying 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 which extended widely over stratiform precipitation regions, far from any convective updrafts. These graupel distributions were considered erroneous because graupel growth is primarily through the accretion of liquid water, which is generally confined to regions close to convective updrafts. The relatively dense graupel (0.4 g $cm^{-3}$ in 3ICE) should generally fall out within ~ 20 km of convective updrafts, while less dense snow (0.1 g $cm^{-3}$) could be horizontally advected over much greater distances. Instead, simulations based upon 3ICE showed comparable concentrations of snow and graupel, or even graupel-dominated distributions, at distances greater than 50 km from the convective "source" regions.

The erroneous distributions of graupel were traced to the process of accretion of snow by graupel, which in 3ICE did not discriminate between the relative sizes of the snow and

graupel particles. Effectively, a graupel particle could collect a snow particle of any size, and the resulting particle would be assigned to the graupel category. This would lead to a composite particle with an unrealistically large density if the colliding graupel particle was relatively small and snow particle relatively large. Therefore, very small graupel particles advected large distances from convective regions could grow significantly due to the collection of snow.

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 gridpoints 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 kernal, 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 effect of the correction is to limit graupel distributions to regions near convective updrafts, while snow dominates ice precipitation distributions at greater distances. These relative proportions graupel and snow are consistent with *in situ* microphysical observations in stratiform regions, which indicate low liquid water contents and ice-phase precipitation distributions dominated by snow/aggregates; see Houze and Churchill (1987), McFarquhar and Heymsfield (1996), and Stith et al. (2002). The correction also improves the distributions of microwave scattering by ice-phase precipitation, since excessive scattering by graupel, previously noted in simulations of stratiform regions, is eliminated. On the other hand, there is new evidence that even with the correction, simulated microwave scattering by ice-phase precipitation is still generally greater than the observed scattering, and this discrepancy will be the subject of future investigations.

*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; i.e., there is no explicit representation of mixed-phase precipitation in these simulations. In order to account for the effects of mixed-phase precipitation, the 1-D 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.

Although the impact of melting particles on upwelling microwave radiances is generally small in tropical systems, the impact of melting can be significant in midlatitude systems. The shallower rain layer and widespread distributions of snow aloft in extratropical cyclones leads to radiance increases up to 15 °K at 19.35 GHz due to melting snow. Changes in radiances of this order are associated with changes in total precipitation optical depth ~0.06 (or ~0.43 dB), which can lead to a significant impact on retrieved rain rates at midlatitudes. Therefore the parameterized effects of melting are included in the Version 6 TMI algorithm database.

*d. Definition of cloud and precipitation properties in simulated footprints*

As described in Section 2, any cloud/precipitation properties $\boldsymbol{x}$ 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 x 14 km nominal "footprint" is selected for the resolution of Version 6 algorithm products.

Simulated properties such as the surface rainfall rate, cloud and precipitation water contents are horizontally averaged over 14 km x 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, and cloud drying is $Q_2$, the apparent moisture sink; see Yanai et al. (1973). Eddy flux convergence contributions to these quantities are defined relative to a 55 km-resolution mean state. Simulated cloud latent heating and drying rates are similarly averaged over 14 km x 14 km domains to represent heating and drying rates at product resolution.

The classification of convective and non-convective gridpoints 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 convective/non-convective classification of CRM gridpoints 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). First, the vertical atmospheric column above each model gridpoint is examined to find the level corresponding to the maximum magnitude of mass flux. The maximum magnitude of mass flux is compared to a prescribed threshold,

$$M_{thresh} = S \cdot 2\text{kg m}^{-2}\text{s}^{-1}, \tag{4}$$

where

$$S = 1 - \left(\frac{\Delta}{10}\right)^{1.8} \tag{5}$$

is a scaling factor that accounts for the decrease in grid-resolution updraft/downdraft strength depending on the model grid spacing. Here, $\Delta$ is the model horizontal grid spacing in kilometers. If the maximum mass flux exceeds the threshold, and the level of the maximum flux is in the lower troposphere (here defined as below the freezing level, or below 2 km altitude, whichever is higher), then the gridpoint is classified as convective. If the mass flux threshold is exceeded in the upper troposphere, then a test is performed to see if the upper-tropospheric maximum is part of a contiguous "updraft" with roots in the lower troposphere. This is accomplished by sequentially searching, level by level, from the level of maximum mass flux to ever-lower levels. If, at the next lowest level, the maximum mass flux in a 3 x 3 gridpoint neighborhood is greater than $M_{thresh}$, then the process is repeated at the next-lowest level, testing the 3 x 3 neighborhood centered on the new maximum. A 3 x 3 neighborhood of the current maximum at the next-lowest level is tested to account for significant tilting of convective updrafts. The process is continued until (a) the maximum magnitude of mass flux is less than $M_{thresh}$, at which point the process is terminated, or (b) a maximum magnitude of mass flux greater than $M_{thresh}$ is found at a level in the lower troposphere, and the original gridpoint is classified as convective.

If the tests previously described do not result in a convective classification, then further tests are applied to identify weaker convection. At the given gridpoint, the column-average mass flux and column-average cloud liquid water content below the

freezing level are calculated. The gridpoint is classified as convective if the column-average mass flux is positive, and the average cloud water content is greater than

$$\overline{CLWC}_{thresh} = S \cdot 0.08 \text{ g m}^{-3} . \quad (6)$$

Note that cloud water is generally produced if mean vertical motions are positive below the freezing level. If none of the above tests are satisfied, then the gridpoint is classified as non-convective. The definition of convection derived from these tests is generally correlated with classifications based upon rainfall rate, but it is somewhat more liberal since some gridpoints associated with stronger upper level updrafts are also classified as convective.

Using the convective/non-convective classification, the area fraction of convection within the simulated footprints (14 km x 14 km) is calculated. The area fractions of convection are compared to empirical estimates derived from the TMI data in the algorithm's constraint term, (2). The convective contributions to surface rainfall rate, $Q_1$-$Q_R$, and $Q_2$ within each simulated footprint are also computed and included in the vector $\boldsymbol{x}$ of parameters to be estimated.

In addition to the convective area fraction, the area fraction of total rainfall rate within each simulated footprint is also calculated. The area fraction of total rainfall rate is defined as the area within each simulated 14 km x 14 km footprint where the model gridpoint values of surface rainfall rate exceed 0.3 mm $h^{-1}$. The area fraction of total rainfall is compared to empirical estimates from the TMI data in the algorithm's constraint term, (2).

*e. New geographic database*

Crucial to the operation of the algorithm is the correct classification of the underlying surface as open water, land, or coast. Previously, the definition of "coast" was based on a threshold applied to the distance to the nearest land (if over water) or nearest 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" due to 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° x 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 it 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 "coastal land" if $R$ is less than 50 km; points over water are classified as "coastal water" 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 retrievals 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 "ocean" and "land" would not experience operationally significant contamination by the opposite surface type.

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

In this study, the TMI algorithm as applied to ocean locations, and therefore the radiance indices [$I_S$ and $I_O$ in (1)] are the normalized polarization and scattering indices described by Petty (1994). These indices were developed to isolate the microwave signal due to precipitation particles from variations in radiances due to varying ocean surface emissivity, sea surface temperature, and atmospheric water vapor and also to decouple the brightness temperature effects of microwave attenuation from those due to 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. Unfortunately, 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 (1994); details of the implementation for TMI, including the regression-based algorithms for $V$ and $U$ and the dependence of background radiances on $V$ and $U$, are given by Chiu and Petty (2004). 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-resolving model simulations are determined by evaluating the true background radiance field. That is, upwelling radiances at each model gridpoint are calculating with all cloud and precipitation water contents set to zero.

*g. Reformulation of the convective/non-convective rain area constraint*

The changes in the Version 6 TMI algorithm that have the greatest impact on estimates of cloud latent heating and drying 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 is computed as a minimum variance estimate, combining convective fraction estimates from both TMI polarization signatures and image texture; see Olson et al (2001c).

$$f_{conv} = \left\{ \frac{f_{pol}}{\sigma_{pol}^2} + \frac{f_{tex}}{\sigma_{tex}^2} \right\} \Bigg/ \left\{ \frac{1}{\sigma_{pol}^2} + \frac{1}{\sigma_{tex}^2} \right\} \qquad (7)$$

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

The polarization-based estimate of convective area fraction within an 85.5 GHz TMI footprint is given by Eq. (8) in Olson et al. (2001c). This estimate will be identified as $f_{polO}$ in the present study. Since the nominal TMI footprint in the Version 6 algorithm is larger (14 km x 14 km) than the 85.5 GHz footprint, a three-footprint filter is first applied to the estimates to approximate the spatial response at 37 GHz (given the spacing of the 85.5 GHz footprints along-track and cross-track, it is not possible to closely approximate a uniform response over the nominal footprint using such a filter). The filtered convective fraction estimate is given by

$$\hat{f}_{polO} = 0.2329\, f_{polO_{j-1}} + 0.5342\, f_{polO_j} + 0.2329\, f_{polO_{j+1}}. \qquad (8)$$

Here, the subscript $j$ indicates the scan position of the footprint, and $\hat{f}_{polO}$ is the filtered estimate at scan position $j$. Therefore, the filter weights the convective fraction estimates corresponding to the current, preceding, and following footprints along the TMI scan.

Finally, synthetic radiance tests are used to calibrate the filtered estimate to make it conform to the definition of convective fraction in the supporting CRM database. The synthetic radiances are based upon a subset of the CRM simulations in Table 2, for which upwelling radiance calculations at the TMI frequencies have been performed to simulate TMI observations. The polarization of radiances at 85 GHz is approximated using the empirical formula,

$$POL_{strat} = a\left(\frac{TB85_v + TB85_h}{2}\right) + b, \qquad (9)$$

at each CRM gridpoint classified as stratiform (Olson et al. 2001c). The polarization of convective gridpoints is assumed to be zero. Here, *TB85v* and *TB85h* are the microwave radiances at 85.5 GHz in the vertical and horizontal polarizations, respectively, and $a$ is equal to –0.192 and $b$ is 52.4 °K. The simulated radiances at CRM resolution are convolved by the TMI antenna patterns to synthesize TMI observations. A regression fit between the filtered convective fractions, $\hat{f}_{polO}$, derived from the synthetic radiance data and the convective fractions derived from the CRM simulations over nominal footprints yields

$$f_{pol} = 0.011665 + 1.5315\,\hat{f}_{polO} - 1.7703\,\hat{f}_{polO}^2 + 1.1993\,\hat{f}_{polO}^3, \qquad (10)$$

where $f_{pol}$ is the final convective fraction estimate. The error variance, $\sigma^2_{pol}$, of polarization-based convective fraction estimates is given by Eq. 16 in Olson et al. (2001c), propagated through (8) and (10). 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, 1994), 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.

The first index is the maximum difference between the normalized polarization at 37 GHz and the normalized polarization of its neighbors.

$$P37_{grad} = \max_{\substack{i'=i\pm1 \\ j'=j\pm3}} \left\{ P37_{i'j'} - P37_{ij} \right\} \tag{11}$$

Here, $P37$ is the normalized polarization at 37 GHz, as defined by Petty (1994), and the subscripts refer to the scan line ($i$) and position along the scan line ($j$) of a given footprint. Because the spatial separation between every third footprint along a scan line (13.7 km) is approximately equal to the spacing between scan lines (13.9 km), (8) represents the maximum difference between $P37_{ij}$ and the $P37$ values of nearly equally spaced neighbors. A large value of $P37_{grad}$ generally indicates a strong gradient of absorption by liquid precipitation; however, the edges of stratiform precipitation regions can sometimes produce significant absorption gradients. In order to eliminate these stratiform precipitation gradients from consideration, the polarization-based convective fraction estimate is used to determine which gradients are in close proximity to convection.

$$w_{pol} = \sum_{\substack{i-2\le i'\le i+2 \\ j-6\le j'\le j+6}} f_{pol_{i'j'}} \exp\left(\frac{-d_{i'j'}}{d_o}\right) \Bigg/ \sum_{\substack{i-2\le i'\le i+2 \\ j-6\le j'\le j+6}} \exp\left(\frac{-d_{i'j'}}{d_o}\right) \tag{12}$$

The distance $d_{i'j'}$ is the separation distance between the footprint indexed by $ij$ and the footprint indexed by $i'j'$, and $d_O$ is set to 4 km. Therefore the index $w_{pol}$ is a measure of the probability of convection based upon the presence of convection in the neighborhood of the footprint being analyzed.

The previous indices are useful for indicating the presence of convection if the convection produces relatively strong microwave signatures. A final index is formulated to identify footprints containing weak or isolated convection.

$$P37_{max} = \max_{\substack{i-1 \le i' \le i+1 \\ j' = j-3, j}} \left\{ 0.5 P37_{i'j'} - P37_{ij} + 0.5 P37_{2i-i'\ 2j-j'} \right\} \tag{13}$$

The index $P37_{max}$ is essentially the maximum second derivative of $P37$ in any direction, centered on the footprint being analyzed. In the search for the maximum second derivative, only values of $P37$ at footprints indexed by $i'j'$ and $2i\text{-}i'\ 2j\text{-}j'$ both greater than the value of $P37$ at footprint $ij$ are admitted, thereby ensuring that $P37_{ij}$ is a local minimum. Weak or isolated convection will produce locally greater liquid water absorption than neighboring footprints, and therefore a smaller value of $P37$.

Based upon $P37_{grad}$, $w_{pol}$, and $P37_{max}$, the possibility of convection within the analyzed footprint is ascertained. If it is determined that convection within the footprint is possible, then the area fraction of convection within the footprint is estimated using the normalized polarization difference at 37 GHz, as follows.

$$\hat{f}_{tex} = \begin{cases} 1 - P37, & \text{for } S85 < 40\text{K and} \\ & \left\{ \left( P37_{grad} > 0.2 \text{ and } w_{pol} > 0.1 \right) \text{ or } P37_{max} > 0.3 \right\} \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

For optically thick precipitation regions partially filling a radiometer footprint, $1 - P37$ is approximately equal to the area fraction of precipitation within the footprint; Petty (1994). Even if convection is indicated by the index tests on the right-hand-side of (14), not all of the rain within the footprint is necessarily convective. In addition, the resolution of the nominal footprint (14 km x 14 km) is different from the TMI spatial response at 37 GHz. Therefore, the convective fraction estimates are calibrated against model-simulated convective fractions, using synthetic data as in the derivation of (10). The final texture-based convective fraction estimate is given by

$$f_{tex} = -0.12272 + 0.91249 \hat{f}_{tex} . \tag{15}$$

If a texture-based convective fraction estimate is allowed by (14), then the error variance, $\sigma^2_{tex}$, of the estimate is assigned a value of 0.2 $^{\circ}K^2$ 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 non-convective 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/non-convective rain flux directly, the *area* proportions of convective and non-convective 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 (1994), the total rain area fraction in the plane-parallel limit is first approximated by

$$\hat{f}_{rain} = 1 - P37, \tag{16}$$

and then calibrated using synthetic data, yielding

$$f_{rain} = -0.0031981 + 0.28436\,\hat{f}_{rain} + 1.7964\,\hat{f}_{rain}^2 - 1.1102\,\hat{f}_{rain}^3. \tag{17}$$

Estimates of convective area fraction, given by (7), and total rain fraction, given by (17), are incorporated into the constraint term, (2), of the Bayesian method.

## 4. Error Estimates

Errors in retrieved precipitation-related parameters can be estimated (a) from comparisons of TMI-retrieved parameters to independent data, (b) from algorithm applications to synthetic data, and (c) from the algorithm itself, drawing upon information provided by the Bayesian formulation. Since independent estimates of rain rate from ground-based radars or raingages are not always reliable, the alternatives (b) and (c) are emphasized here. 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 3-D model volumes at two time periods each from the Hurricane Bob, GATE Sept. 12, TOGA COARE Feb. 22, and TOGA COARE Dec. 19-26 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. Since the polarization of scattered radiances at 85.5 GHz is used in the TMI algorithm (section 3f) but not calculated using the Eddington method, the empirical function (9) 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 to 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 (Fig. 1a) are comparisons of true and estimated parameters at footprint-scale. True and estimated parameters in the middle panels (Fig. 1b) have been averaged to 28 km resolution, and parameters in the lower panels (Fig. 1c) have been averaged to 56 km resolution. Note that although the overall magnitude 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 to 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 Fig. 2. Plotted in the top panel 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 yields the relative contribution of the rain rate (along the abscissa in 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 histogram of true rain rates is plotted. Here, the plotted points can be interpreted as the contribution of the differences between estimated and true rain rates to the total bias, corresponding to the true rain rates along the abscissa. The Bayesian formulation 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 $Q_I$-$Q_R$ 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 $Q_I$-$Q_R$, as well as the contributions to the bias of $Q_I$-$Q_R$, at all altitudes where algorithm estimates are made (0 – 18 km). Note from the mean contribution plots that the range of estimated $Q_I$-$Q_R$ values is more limited than the true range. The contributions to $Q_I$-$Q_R$ 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 non-convective contributions to $Q_1$-$Q_R$, from applications of the algorithm to the synthetic Hurricane Bob data (see Table 1). 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. Since 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 non-specific 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. Algorithmic 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 due to deficient information contained in the microwave radiance data. Although addition random error in TMI estimates can result from errors in the algorithm's supporting database of cloud-resolving model simulations, relation (3) provides a useful "building block" for the derivation of errors in space-time 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 TMI observations (over ocean footprints only) spanning the month of July 2000. Note that the errors due to 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 calculated algorithmic 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. Also, Haddad (personal communication) has modified (3) to include the additional error contribution due to uncertainties in the cloud-resolving model simulations of the algorithm's supporting database. Uncertainties in the model hydrometeor profiles have not yet been quantified, however.

*c. Errors in half-degree, instantaneous 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 degree spatial resolution are consistent with the grid resolution of a prospective analysis. In this subsection, the algorithmic estimates of footprint-scale error are propagated to 0.5° resolution and compared to TMI-PR differences.

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

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i , \tag{18}$$

then the error variance of the spatial average is

$$\sigma_{\bar{x}}^2 = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_i \sigma_j r_{ij} , \tag{19}$$

where $\sigma_i$ and $\sigma_j$ are the errors of the $i$th and $j$th 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 algorithmic estimates of footprint-scale random error given by (3), the error variance of the spatial average, (19), 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, 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 x 55 km rectangular box. Note that although both curves exhibit the familiar exponential decrease in correlation with footprint separation distance (e.g. Bell et al. 1990, 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 x 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-average 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-average rain rate, decreases even more rapidly with footprint separation distance than the unconditional error correlation. The reason for this behavior may be traced to similar behavior in the spatial correlation of rain rates: as the average rain rate is restricted to some small interval, the variance of average rain rate must also be small. Therefore higher rain rates in one part of a specified area must be compensated by lower rain rates in another part, resulting in positive correlations between footprints at small spatial separations and negative correlations for larger separations. Although the correlation of errors does not become negative at larger spatial separations, the correlation of errors becomes small, in consonance with small or negative rain correlations.

If the mean correlation function based upon the rainfall-stratified data is incorporated in (19), and the footprint-scale uncertainties are derived from (3), then (19) can be evaluated from applications of the TMI algorithm to radiance data. As a test of (19), 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 into 1 mm $h^{-1}$ intervals. Based upon the TMI-PR pairs in each rain interval, the mean half-degree TMI rain rate error from (19) is plotted against the mean half-degree PR rain rate; see Fig. 7. For comparison, the standard deviation of TMI-PR differences is also plotted for each rain rate interval.

Comparing the half-degree resolution TMI rain rate random errors in Fig. 7 to 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 curves in a quantitative way, since the distribution of rain rates narrows with increasing averaging area; however, the minimum percentage error at half-degree resolution is ~20%, compared to ~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 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 ranged from roughly 50% at 1 mm $h^{-1}$, to about 20% at 14 mm $h^{-1}$. TMI errors also appear to explain 70 - 90% of the TMI-PR random deviation; see Fig. 7. The residual 10 - 30% is likely due to random errors in PR rain estimates and differences in spatial sampling of half-degree boxes by the TMI and PR.

*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. (2000). As the basis of a preliminary analysis of the uncertainties in longer time/space average precipitation/heating estimates, the random errors in monthly mean, 2.5° resolution estimates of surface rainfall rate are examined in this subsection. Also, as in previous work [Bell et al. (1990), Li et al. (1998), Bell and Kundu (2000)], it will be assumed that the errors due to the infrequent sampling of low earth orbiting sensors such as TMI are independent of retrieval errors. It follows that the total error variance of the time-average estimate is equal to the sum of the retrieval and sampling error variances,

$$\sigma^2_{\langle \overline{P} \rangle} = \sigma^2_{retr} + \sigma^2_{samp} , \tag{20}$$

where $\overline{P}$ is the instantaneous, area-average rain rate in a 2.5° x 2.5° box, and the brackets, $\langle\ \rangle$, indicate a monthly average. The accumulated retrieval error variance, $\sigma^2_{retr}$, over the period of a month, is evaluated by calculating the retrieval error variance of each 2.5° box-average instantaneous estimate, as described in section 4c, and then taking the average of these variances over the month.

Regarding the sampling error, $\sigma^2_{samp}$, it will be assumed that the estimate of the monthly-mean rain rate over a given 2.5° x 2.5° box is the area-weighted average,

$$\langle \overline{P} \rangle = \frac{1}{S} \sum_{i=1}^{M} \left( \frac{A_i}{A} \right) \overline{P}_i , \tag{21}$$

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

$$S \equiv \sum_{i=1}^{M} \left( \frac{A_i}{A} \right) . \tag{22}$$

Here, $A$ is the area enclosed by the 2.5° x 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 $i$th overpass, and the summations are over all full or partial observations of the box during the month.

The error in $\langle \overline{P} \rangle$ due to the relatively infrequent sampling of rainfall by low earth-orbiting satellite radiometers has been studied by several investigators; e.g., Laughlin

(1981), Shin and North (1988), Bell et al. (1990), Kedem et al. (1990), Oki and Sumi (1994), Huffman (1997), Li et al. (1998), Steiner et al. (2003). Investigations by Bell and Kundu (1996) and Bell and Kundu (2000), hereafter 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 the following expression for the percentage sampling error,

$$\frac{\sigma_{samp}}{\langle \overline{P} \rangle} = \frac{\left\{ \sigma_A \langle \overline{P} \rangle^{-1/2} S^{-1/2} \left( \coth\left[ \frac{T}{2S\tau_A} \right] - \frac{2S\tau_A}{T} \right)^{1/2} \right\}}{\langle \overline{P} \rangle^{1/2}}. \quad (23)$$

Here, $\sigma^2_A$ is the variance of the 2.5° instantaneous rain estimates, $\tau_A$ is the autocorrelation time of instantaneous box-average rain rates, $\overline{P}$, and $T/S$, the period of observation $T$ (1 month) divided by the effective number of visits, has been used to approximate the sampling time interval ($\Delta t$) in the original expression from BK00.

Based upon a limited number of radar and raingage studies, BK00 argued that the quantity in curly brackets in (23) is relatively constant for a given geographic location, leading to an approximate $\langle \overline{P} \rangle^{-1/2}$ dependence of percentage error on monthly-mean rain rate. In the present study, (23) is evaluated with the help of the TMI rain rate estimates themselves. The monthly-mean rain rate is estimated using (21), while $S$ is given by (22), and $\sigma^2_A$ is approximated by the variance of the instantaneous 2.5° TMI rain estimates. The autocorrelation time, $\tau_A$, 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 GARP Atlantic Tropical Experiment (GATE) radar observations,

$$\tau_A = 0.394 \left( \sqrt{A} \right)^{0.525}, \quad (24)$$

is employed, where $\tau_A$ is in hours and $A$ is in km$^2$. BK00 reported autocorrelation times as short as 4 h over 2.5° boxes during TOGA COARE, while autocorrelation times as long as 14 h over 5° boxes in southern Japan during summer were found by Oki and Sumi (1996). 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 (24) is an obvious simplification.

Shown in Fig. 8 are estimates of the monthly-mean surface rain rate, convective rain proportion, and $Q_I$-$Q_R$ at 7 km and 3 km altitude, based upon all TMI over-ocean observations from July 2000. The main climatological features of tropical rain distributions are seen in the surface rain map, including the intertropical convergence zone (ITCZ), more widespread rains over the Western Pacific warm pool region, and the intense rains over the Bay of Bengal signifying the summer phase of the Indian Monsoon. To the west of the Americas and Africa, the regions of the subtropical highs are nearly rain-free. In the subtropics of the southern hemisphere, weaker rains associated with the northern fringe of the midlatitude storm tracks are also seen. Rain rates are predominantly convective in the northern hemisphere and the portion of the southern hemisphere down to a latitude of 20° S. Note the relative minima of convective proportion along the 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, since 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 retrieval and sampling errors plotted as functions of the monthly, 2.5° resolution rain rate estimates. Although retrieval errors are not entirely negligible at 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 retrieval error contribution to the monthly rain error, the remainder of this section will focus on sampling errors.

Distributions of sampling error in monthly, 2.5°-resolution surface rain rates, convective rain proportions, and $Q_I$-$Q_R$ at 7 km and 3 km altitudes are provided in Fig. 10. The sampling error of surface rain rates is computed using (23); the sampling errors of convective rain rate and $Q_I$-$Q_R$ are calculated by substituting the variances of these quantities for $\sigma^2_A$ in (23) and assuming that the autocorrelation times are the same as the rain rate autocorrelation time. Note that the sampling errors for rain rate and $Q_I$-$Q_R$ closely follow the patterns of rain rate and heating rate shown in Fig. 8; however, the percentage errors decrease with increasing rain rate; see Fig. 9. This behavior of the sampling error was noted by 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 due to 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 (23). Convective rain proportion, being 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. 10. Aside from variations in sampling error due to 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 retrieval 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 retrieval and sampling errors of estimated June-August 2000 average rain rates and $Q_I$-$Q_R$ at 7 km altitude. Although

there is some reduction of the variance of 3-month average rain rates relative to the 1-month averages, a significant reduction in sampling error in the 3-month averages is also 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/drying ($Q_1$-$Q_R$ and $Q_2$) 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-resolving 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 non-convective regions of 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 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-resolving model simulations, and through construction, starting with algorithmic 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 algorithmic 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 climate

studies or large-scale analyses. 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 $h^{-1}$. These errors accounted for 70-90% of the deviation between TMI and collocated PR estimates at this resolution. Random errors in monthly, 2.5° rain rate and heating estimates are due to the combination of retrieval 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 retrieval 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% to 15% at 5 mm $day^{-1}$. Latent heating errors are similarly reduced.

It should be stressed that the algorithmic 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-resolving 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 algorithmic error estimates and the TMI-PR deviations suggest that these algorithm 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 due to the lack of information in the radiance data or insufficient sampling will decrease, exposing systematic errors associated with the cloud-resolving 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 to independent ground-based observations, as well as those from the PR.

*Acknowledgments*

The authors would like to thank Peter Bauer for many helpful discussions during the course of this study regarding the electromagnetic modeling of mixed-phase particles and the analysis of rain rate errors. Arthur Hou kindly provided his expertise and perspectives regarding the assimilation of precipitation and latent heating into global models. This research was supported by the NASA TRMM, Global Water and Energy Cycle, and Global Precipitation Measurement projects.

## References


Adler, R. F., G. J. Huffman, D. T. Bolvin, S. Curtis, E. J. Nelkin, 2000: Tropical rainfall distributions determined using TRMM combined with other satellite and rain gauge information. *J. Appl. Meteor.*, **39**, 2007–2023.

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. *Q. J. R. Meteorol. 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.

Bell, T. L., and P. K. Kundu, 2000: Dependence of satellite sampling error on monthly averaged rain rates: Comparison of simple models and recent studies. *J. Climate*, **13**, 449-462.

Bell, T. L., 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.

Chiu, J.-Y. and G.W. Petty, 2004: Bayesian retrieval of complete posterior PDFs of oceanic rain rate from microwave observations: Theoretical framework and sensitivity test. *J. Appl. Meteor.* (to be submitted)

Curtis, S., and R. Adler, 2000: ENSO indices based on patterns of satellite-derived precipitation. *J. of 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. of Climate*, **13**, 4180-4195.

Hou, A. Y., S. Q. Zhang, A. M. da Silva, W. S. Olson, 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. Meteorological Soc.*, **82**, 659-679.

Hou, A. Y., S. Q. Zhang, O. Reale, 2004: Variational continuous assimilation of TMI and SSM/I rain rates: Impact on GEOS-3 hurricane analyses and forecasts. *Mon. Wea. Rev.* (accepted).

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.

Houze, R. A., Jr., and D. D. Churchill, 1987: Mesoscale organization and cloud microphysics in a Bay of Bengal depression. *J. Atmos. Sci.*, **44**, 1845-1867.

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., S. Surendran, D. W. Shin, R. J. Correa-Torres, T. S. V. V. Kumar, E. Williford, C. Kummerow, R. F. Adler, J. Simpson, R. Kakar, W. S. Olson, and F. J. Turk, 2001: Real-time multianalysis-multimodel superensemble forecasts of precipitation using TRMM and SSM/I products. *Mon. Wea. Rev.*, **129**, 2861-2883.

Kummerow, C., Y. Hong, W. S. Olson, S. Yang, R. F. Adler, J. McCollum, R. Ferraro, G. Petty, D.-B. Shin, and T. T. Wilheit, 2001: The evolution of the Goddard Profiling Algorithm (GPROF) for rainfall estimation from passive microwave sensors. *J. Appl. Meteor.*, **40**, 1801-1820.

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 Sensing*, **34**, 1213-1232.

Laughlin, C. R., 1981: On the effect of temporal sampling on the observation of rainfall. *Precipitation Measurements from Space, Workshop Report*, D. Atlas and O. W. Thiele, Eds., 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, 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.

McFarquhar, G. M., and A. J. Heymsfield, 1996: Microphysical characteristics of three anvils sampled during the Central Equatorial Pacific Experiment. *J. Atmos. Sci.*, **53**, 2401-2423.

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., 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.

Olson, W. S., P. Bauer, 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.

Olson, W. S., 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.

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.

Olson, W. S., C. D. Kummerow, G. M. Heymsfield, and L. Giglio, 1996: A method for combined passive-active microwave retrievals of cloud and precipitation profiles. *J. Appl. Meteor.*, **35**, 1763-1789.

Petty, G. W., 1994: 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.

Pruppacher, H. R., and J. D. Klett, 1997: *Microphysics of Clouds and Precipitation, 2nd Rev. and Enl. Ed.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 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.

Rutledge, S. A., R. A. Houze, Jr., A. J. Heymsfield, and M. I. Biggerstaff, 1988: Dual-Doppler and airborne microphysical observations in the stratiform region of the 10-11 June MCS over Kansas during PRE-STORM. Preprints, *Tenth Int. Cloud Physics Conf.*, Offenbach am Main, Deutscher Wetterdienst, 705-707.

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.

Tao, W.-K., 2003b: Regional-scale modeling at NASA Goddard Space Flight Center, *Recent Devel. Atmos. Sci.*, **2**, 1-52.

Tao, W.-K., and J. Simpson, 1993: The Goddard Cumulus Ensemble Model. Part I: Model description. *Terr. Atmos. Oceanic Sci.*, **4**, 19-54.

Tao, W.-K., S. Lang, J. Simpson, and R. Adler, 1993: Retrieval algorithms for estimating the vertical profiles of latent heat release. *J. Met. Soc. Japan*, **71**, 685-700.

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. Met. Soc.*, **69**, 278-295.

Stith, J. L., J. E. Dye, A. Bansemer, A. J. Heymsfield, C. A. Grainger, W. A. Petersen, and R. Cifelli, 2002: Microphysical observations of tropical clouds. *J. Appl. Meteor.*, **41**, 97-117.

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

Yang, S., and E.A. Smith, 1999b: Four dimensional structure of monthly latent heating derived from SSM/I satellite measurements. *J. Climate*, **12**, 1016-1037.

Yang, S., and E. A. Smith, 2000: Vertical Structure and transient behavior of convective-stratiform heating in TOGA COARE from combined satellite-sounding analysis. *J. Appl. Meteor.*, **39**, 1491-1513.

**Tables**

Table 1. Characteristics of the TRMM Microwave Imager and Precipitation Radar. The horizontal resolution specifications of the PR are at nadir view.

*TRMM Microwave Imager (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 |

*Precipitation Radar (PR)*

| Frequency [GHz] | Horizontal Resolution at Surface/ Range Resolution [km] | Uncertainty Due to Sampling [dB] |
|---|---|---|
| 13.8 | 4.3/0.25 | 0.7 |

Table 2. Cloud-resolving model simulations that currently support the Version 6 TMI algorithm.

| Classification | Description | Model |
|---|---|---|
| Tropical Cyclone | Hurricane Bob, 1991 | MM5 |
| Tropical Squall Line | GATE, 12 Sept. 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 July 1999 | MM5 |

*Note:* locations of simulations are often indicated by the field campaign from which environmental forcing data were obtained. GATE is the Global Atmosphere Research Program's Atlantic Tropical Experiment., and TOGA COARE is the Tropical Ocean Global Atmosphere Coupled Ocean Atmosphere Response Experiment.

**Figure Captions**

Fig. 1. Scatterplots of estimated versus true rain rate and $Q_I$-$Q_R$ at 3 km altitude at (a) footprint resolution (14 km), (b) 28-km resolution, and (c) 56 km resolution, based upon applications of the microwave radiometer algorithm to synthetic radiance data.

Fig. 2. Contributions of estimated and true rain rates to their mean values (upper panel), and contribution of estimated minus true rain rates to the total bias, plotted versus the true rain rate (lower panel), based upon applications of the microwave radiometer algorithm to synthetic radiance data.

Fig. 3. Contributions of estimated (upper panel) and true (middle) $Q_I$-$Q_R$ to their mean values at different altitudes, plotted as function of the estimated and true $Q_I$-$Q_R$, respectively. In the lower panel, the contribution of estimated minus true $Q_I$-$Q_R$ to the total bias is plotted as a function of the true $Q_I$-$Q_R$ at different altitudes. Estimates are based upon applications of the microwave radiometer algorithm to synthetic radiance data.

Fig. 4. Vertical profiles of the mean estimated (right panel) and true (left panel) $Q_I$-$Q_R$, 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, non-convective, and total $Q_I$-$Q_R$ profiles over the entire simulation domain. Mean estimated and true surface rain rates are also indicated.

Fig. 5. Algorithmic estimates of random error in footprint-scale rain rates, based upon microwave radiometer algorithm applications to a subset of July 2000 TMI observations.

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° 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 versus 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.

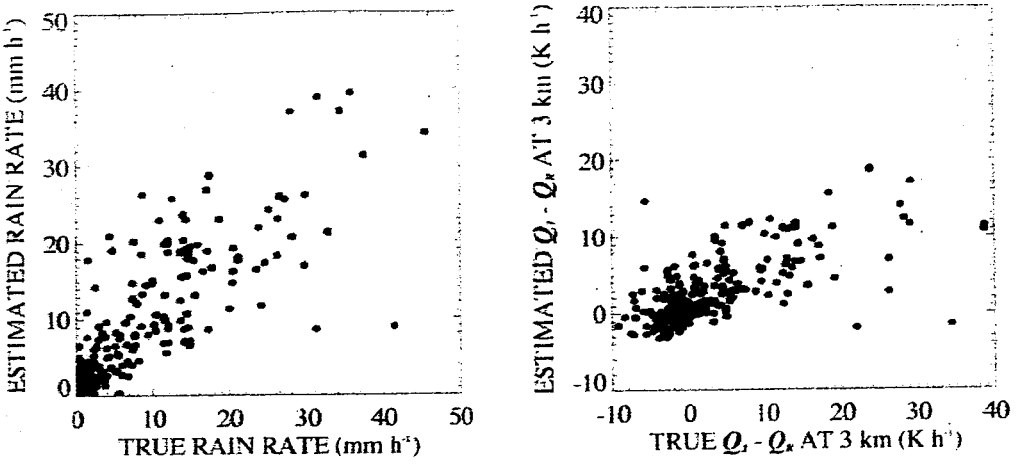
Fig. 8. Monthly-mean, 2.5° resolution surface rain rates, convective rain proportions, and $Q_I$-$Q_R$ at 7 km and 3 km, derived from algorithm applications to TMI observations over ocean from July 2000.

Fig. 9. Percentage errors in monthly-mean, 2.5° resolution surface rain rate estimates due to retrieval errors and incomplete temporal sampling, plotted versus the estimated rain rate (left panel). Monthly statistics are based upon applications of the microwave radiometer algorithm to TMI observations over ocean from July 2000. Also plotted are

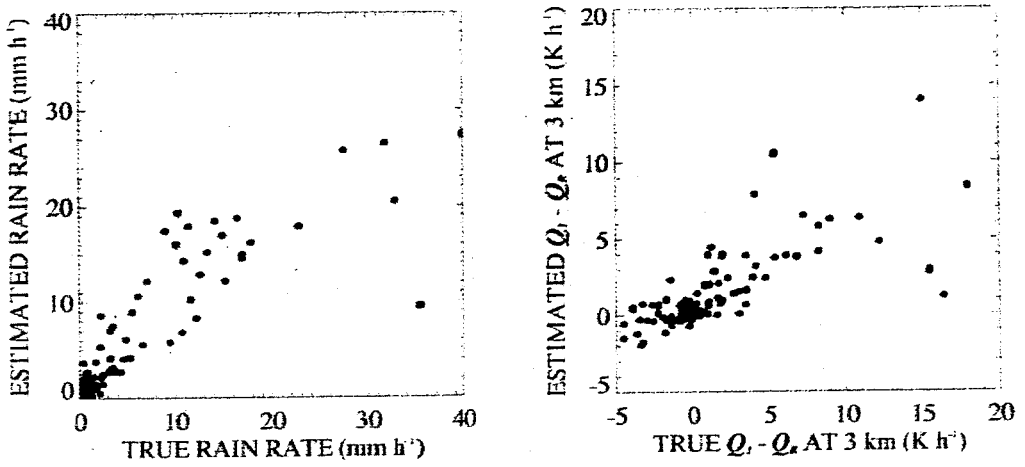
retrieval and sampling errors in June-August 2000 mean surface rain rates at 2.5° resolution (right panel).

Fig. 10. Estimated sampling errors in monthly-mean, 2.5° resolution surface rain rates, convective rain proportions, and $Q_1$-$Q_R$ at 7 km and 3 km, derived from algorithm applications to TMI observations over ocean from July 2000.

(a)


ESTIMATED RAIN RATE (mm h⁻¹)
TRUE RAIN RATE (mm h⁻¹)
ESTIMATED $Q_t$ - $Q_e$ AT 3 km (K h⁻¹)
TRUE $Q_t$ - $Q_e$ AT 3 km (K h⁻¹)

(b)


ESTIMATED RAIN RATE (mm h⁻¹)
TRUE RAIN RATE (mm h⁻¹)
ESTIMATED $Q_t$ - $Q_e$ AT 3 km (K h⁻¹)
TRUE $Q_t$ - $Q_e$ AT 3 km (K h⁻¹)

(c)

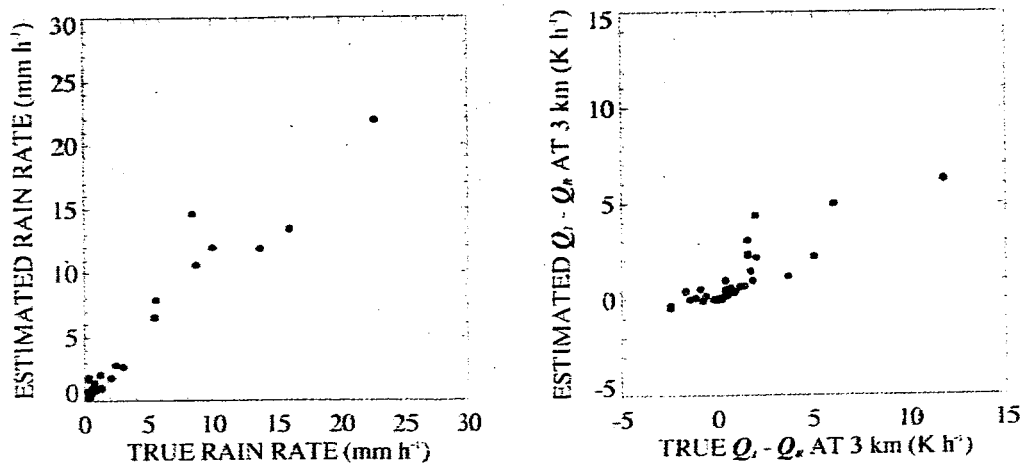

ESTIMATED RAIN RATE (mm h⁻¹)
TRUE RAIN RATE (mm h⁻¹)
ESTIMATED $Q_t$ - $Q_e$ AT 3 km (K h⁻¹)
TRUE $Q_t$ - $Q_e$ AT 3 km (K h⁻¹)

**Fig. 1**

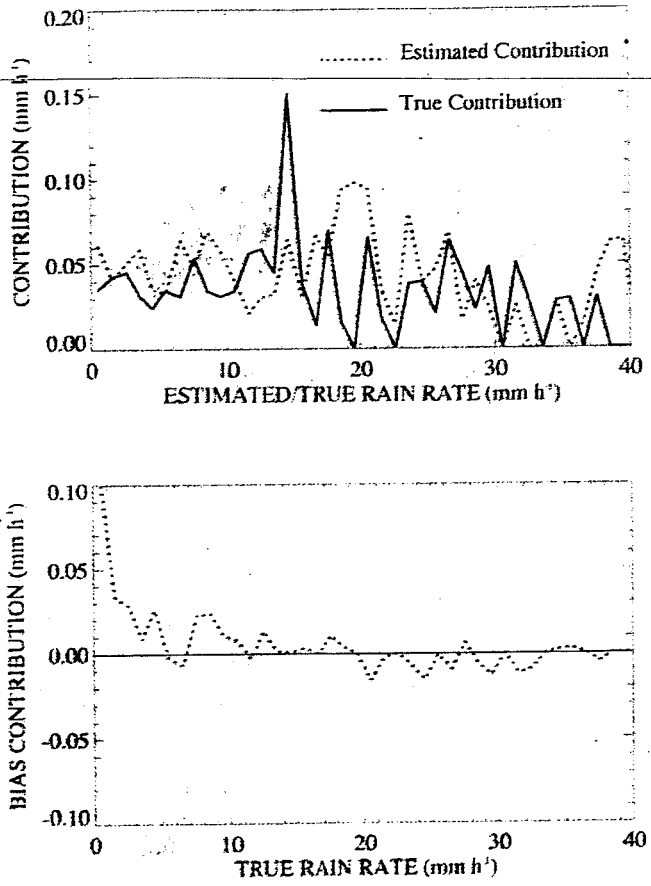

0.20
0.15
0.10
0.05
0.00
CONTRIBUTION (mm h⁻¹)
Estimated Contribution
True Contribution
0
10
20
30
40
ESTIMATED/TRUE RAIN RATE (mm h⁻¹)
0.10
0.05
0.00
-0.05
-0.10
BIAS CONTRIBUTION (mm h⁻¹)
0
10
20
30
40
TRUE RAIN RATE (mm h⁻¹)

**Fig. 2**

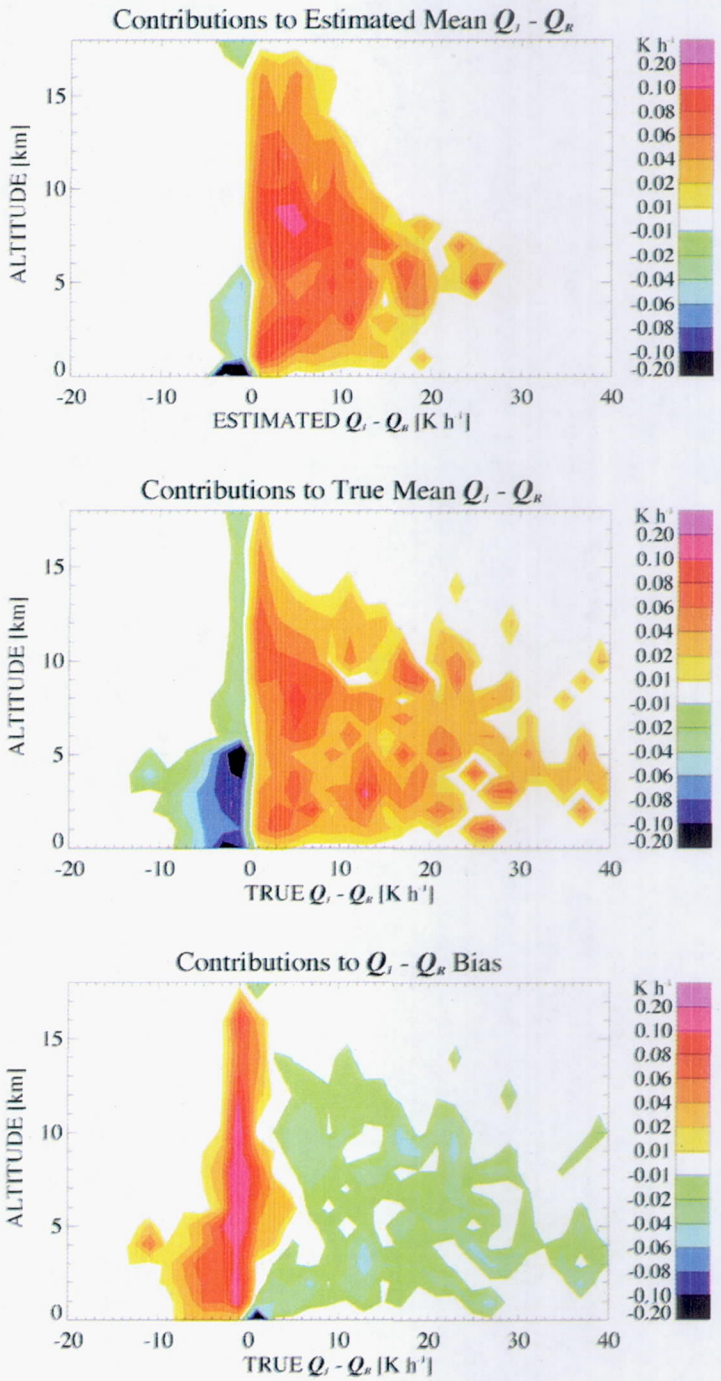
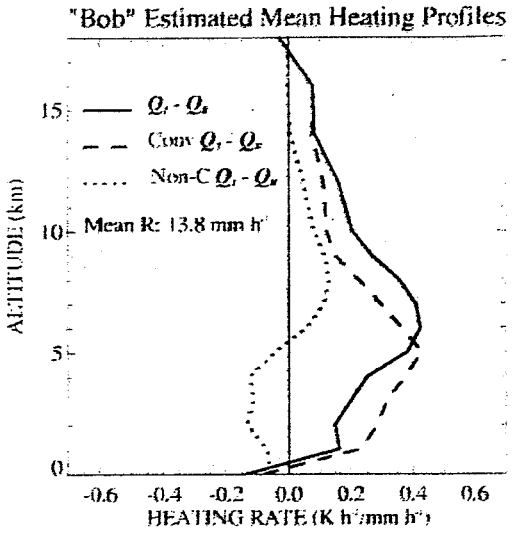

Contributions to Estimated Mean $Q_J$ - $Q_R$
ALTITUDE [km]
ESTIMATED $Q_J$ - $Q_R$ [K h$^{-1}$]
Contributions to True Mean $Q_J$ - $Q_R$
ALTITUDE [km]
TRUE $Q_J$ - $Q_R$ [K h$^{-1}$]
Contributions to $Q_J$ - $Q_R$ Bias
ALTITUDE [km]
TRUE $Q_J$ - $Q_R$ [K h$^{-1}$]
K h$^{-1}$
0.20
0.10
0.08
0.06
0.04
0.02
0.01
-0.01
-0.02
-0.04
-0.06
-0.08
-0.10
-0.20
0
5
10
15
-20
-10
0
10
20
30
40

**Fig. 3**


"Bob" Estimated Mean Heating Profiles
$Q_1 - Q_2$
Conv $Q_1 - Q_2$
Non-C $Q_1 - Q_2$
Mean R: 13.8 mm h$^{-1}$
ALTITUDE (km)
15
10
5
0
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
HEATING RATE (K h$^{-1}$/mm h$^{-1}$)

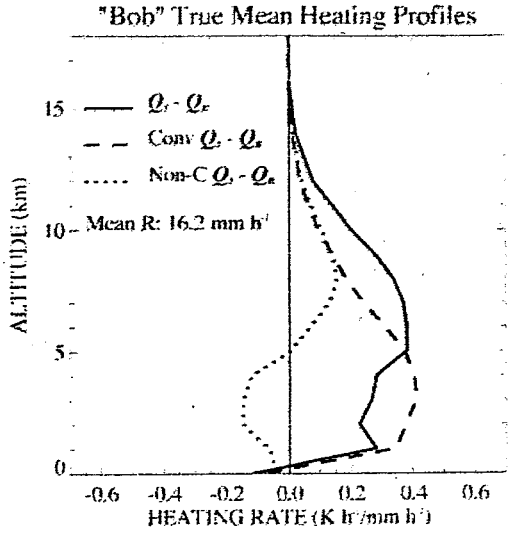

"Bob" True Mean Heating Profiles
$Q_1 - Q_2$
Conv $Q_1 - Q_2$
Non-C $Q_1 - Q_2$
Mean R: 16.2 mm h$^{-1}$
ALTITUDE (km)
15
10
5
0
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
HEATING RATE (K h$^{-1}$/mm h$^{-1}$)

**Fig. 4**

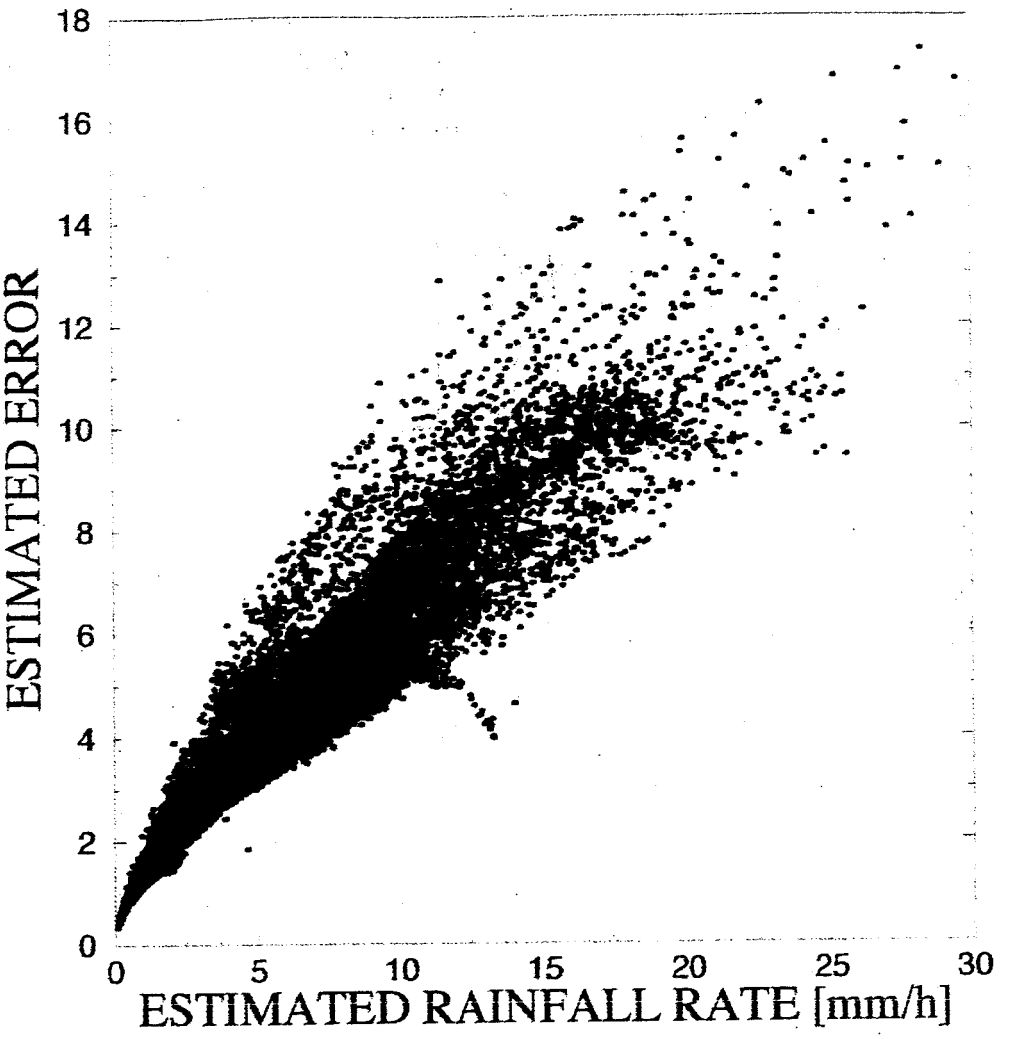

ESTIMATED ERROR
0
2
4
6
8
10
12
14
16
18
0
5
10
15
20
25
30
ESTIMATED RAINFALL RATE [mm/h]

**Fig. 5**

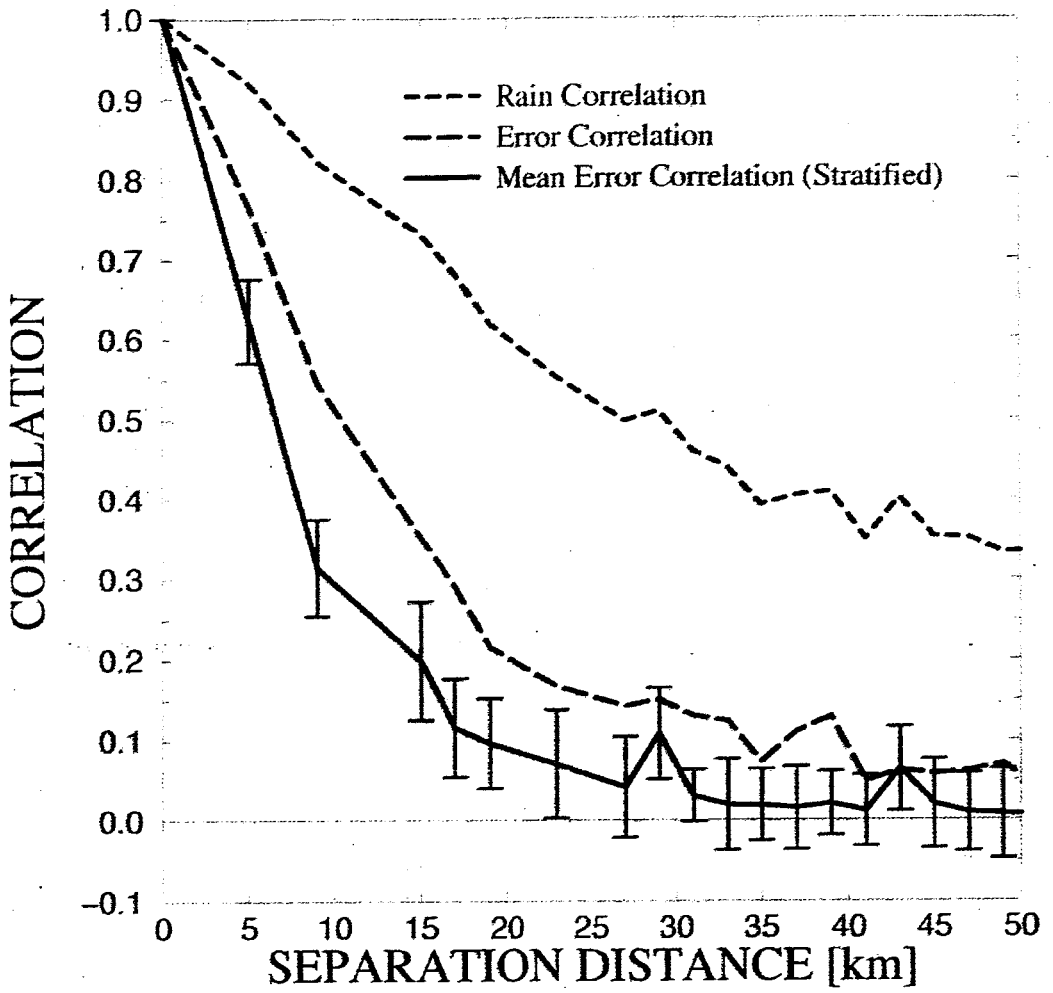

Rain Correlation
Error Correlation
Mean Error Correlation (Stratified)
CORRELATION
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
0
5
10
15
20
25
30
35
40
45
50
SEPARATION DISTANCE [km]

**Fig. 6**

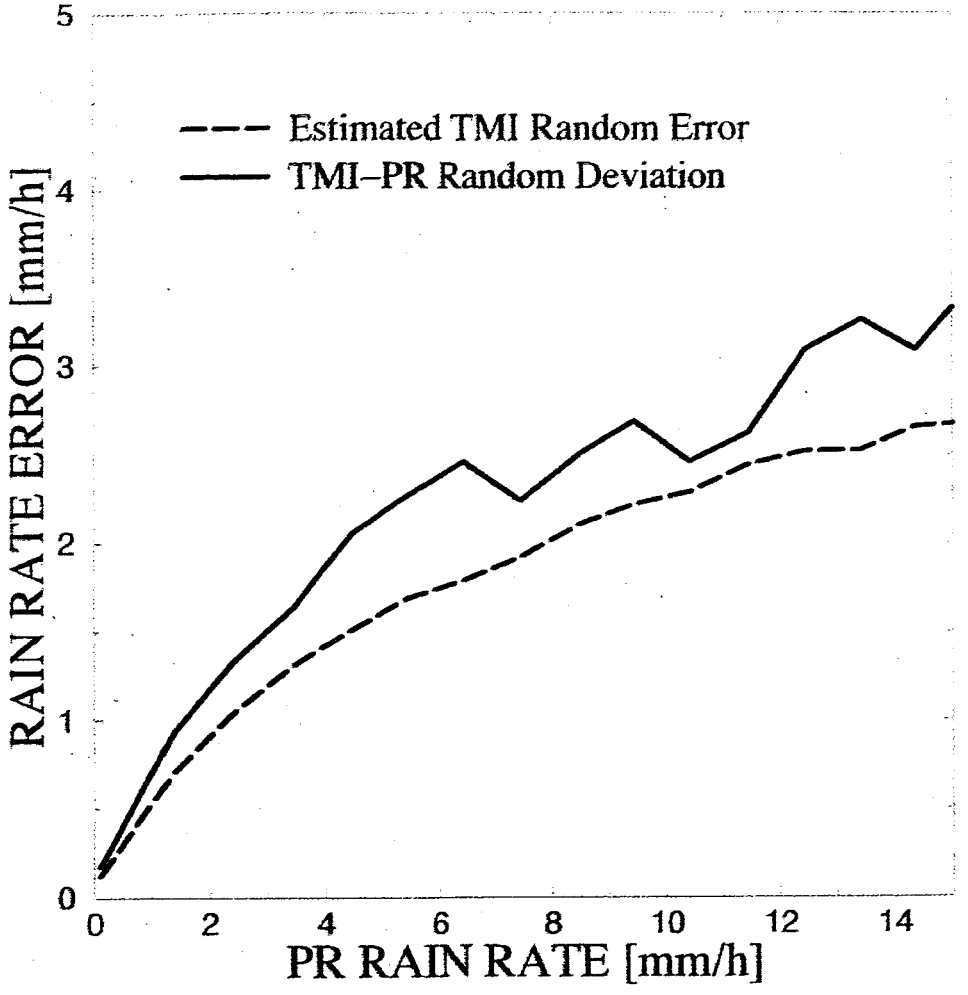

Estimated TMI Random Error
TMI–PR Random Deviation
RAIN RATE ERROR [mm/h]
PR RAIN RATE [mm/h]
0
1
2
3
4
5
0
2
4
6
8
10
12
14

**Fig. 7**

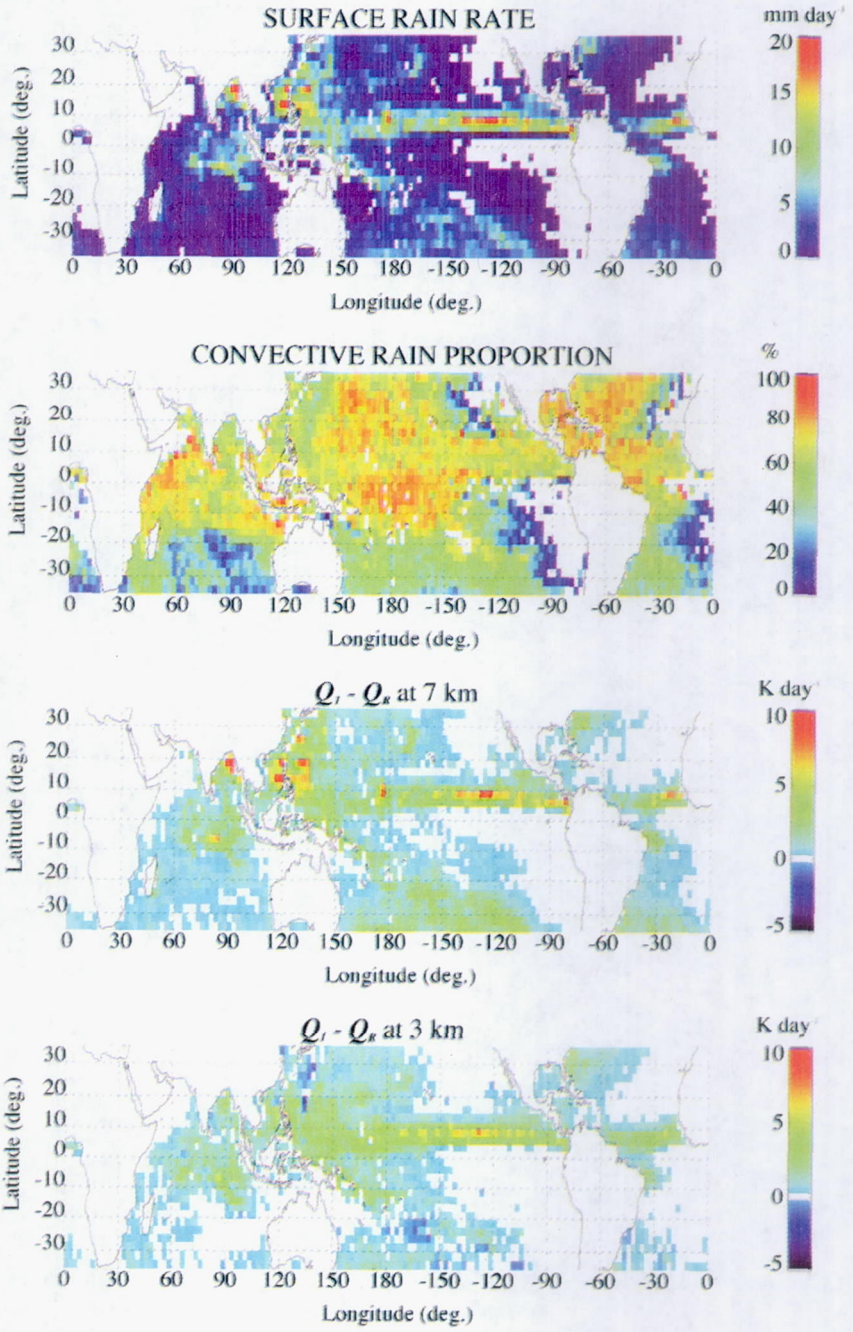

SURFACE RAIN RATE
mm day⁻¹
20
15
10
5
0
Latitude (deg.)
30
20
10
0
-10
-20
-30
0 30 60 90 120 150 180 -150 -120 -90 -60 -30 0
Longitude (deg.)
CONVECTIVE RAIN PROPORTION
%
100
80
60
40
20
0
Latitude (deg.)
Longitude (deg.)
$Q_1$ - $Q_R$ at 7 km
K day⁻¹
10
5
0
-5
Latitude (deg.)
Longitude (deg.)
$Q_1$ - $Q_R$ at 3 km
K day⁻¹
10
5
0
-5
Latitude (deg.)
Longitude (deg.)

**Fig. 8**

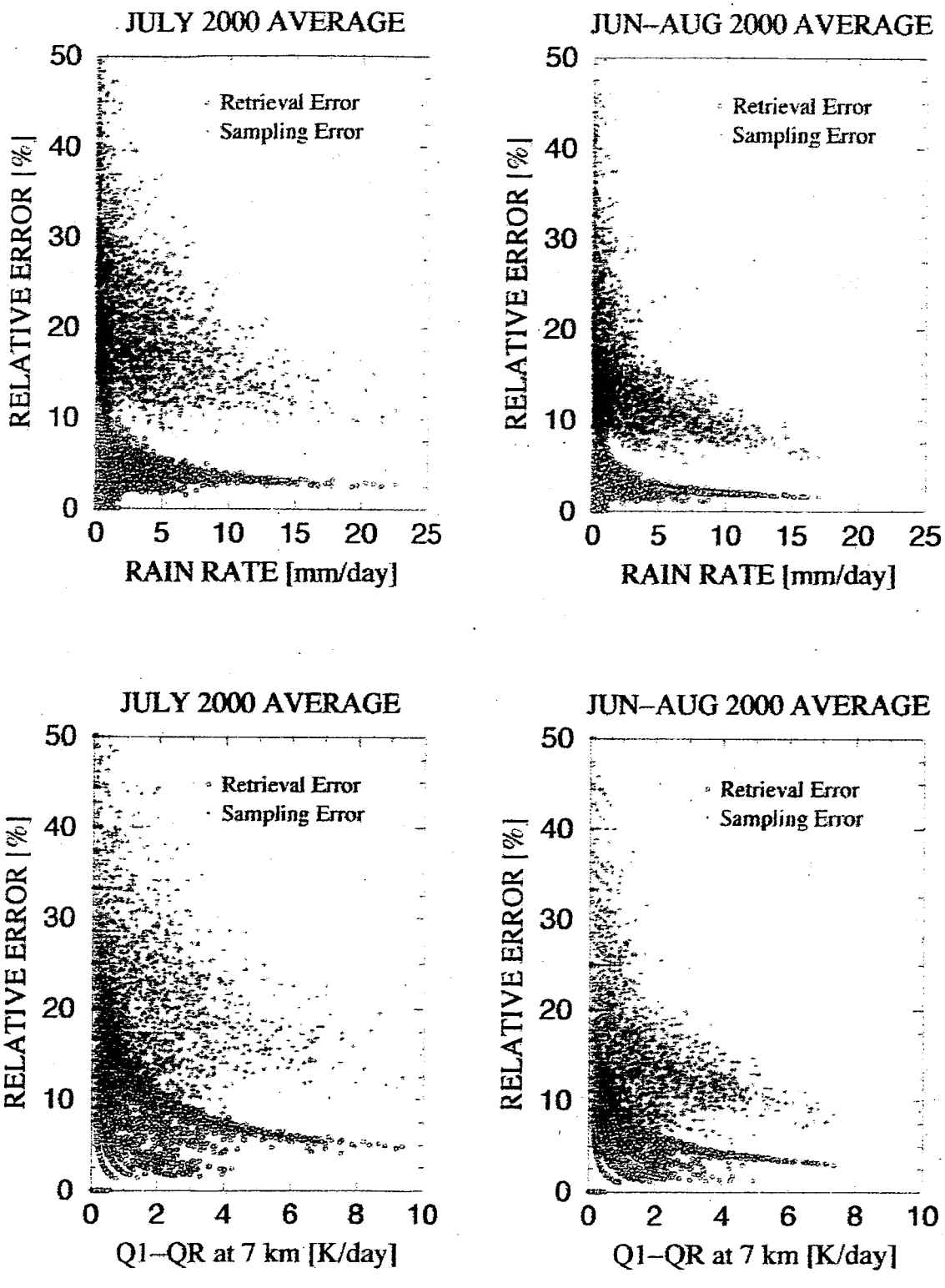

JULY 2000 AVERAGE
JUN–AUG 2000 AVERAGE
Retrieval Error
Sampling Error
RELATIVE ERROR [%]
RAIN RATE [mm/day]
Q1–QR at 7 km [K/day]
0
10
20
30
40
50
0
5
10
15
20
25
2
4
6
8

**Fig. 9**

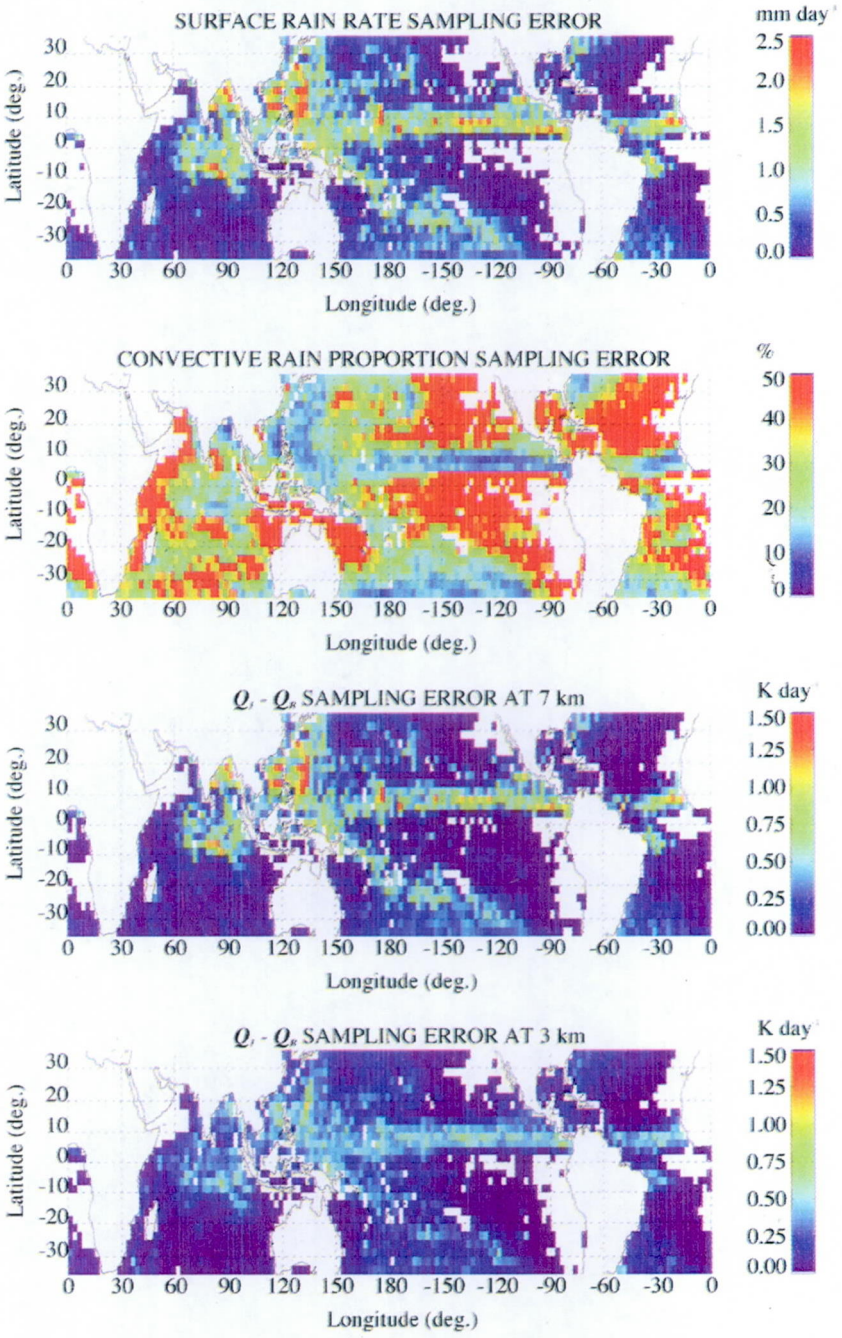

SURFACE RAIN RATE SAMPLING ERROR
mm day⁻¹
2.5
2.0
1.5
1.0
0.5
0.0
Latitude (deg.)
30
20
10
0
-10
-20
-30
0 30 60 90 120 150 180 -150 -120 -90 -60 -30 0
Longitude (deg.)
CONVECTIVE RAIN PROPORTION SAMPLING ERROR
%
50
40
30
20
10
0
Latitude (deg.)
Longitude (deg.)
$Q_l$ - $Q_s$ SAMPLING ERROR AT 7 km
K day
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Latitude (deg.)
Longitude (deg.)
$Q_l$ - $Q_s$ SAMPLING ERROR AT 3 km
K day
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Latitude (deg.)
Longitude (deg.)

**Fig. 10**

10/15/04

## Popular Summary

## Precipitation and Latent Heating Distributions from Satellite Passive Microwave Radiometry Part I: Method and Uncertainties, and Part II: Evaluation of Estimates Using Independent Data

A basic process in clouds is the transformation of water in the vapor phase to the liquid or solid phase. This process produces not only precipitation (rain or snow) but also latent heat, which is released into the atmosphere as the water molecules go from the higher to the lower energy states. A knowledge of the distributions of precipitation over the globe is essential for studying the earth's climate and the cycling of water through the earth-atmosphere system. The distributions of latent heating tell us how variations in buoyancy drive circulations in the atmosphere and how the sun's energy (which is consumed in the evaporation of surface moisture) is redistributed by the atmosphere to regions of more intense precipitation (and thus more intense latent heating).

In Part I of this study, an improved method for estimating precipitation and latent heating distributions from satellite-borne passive microwave radiometer measurements is described. A key feature of this method are the improved models of clouds and precipitation that are used to create a kind of reference "library" of possible solutions that the method can draw upon. If simulated microwave radiances associated with a particular cloud/precipitation model in the library are consistent with radiances actually observed by the satellite microwave radiometer, then it is assumed that that particular model is a likely solution. The models of cloud and precipitation are not perfect, however: errors in the models can result in errors in inferred precipitation or latent heating from the estimation method. In the current study, the physics of precipitating ice particles in the models is improved to remove a significant source of error in the modeled clouds. Another important advance is the way convective and stratiform precipitation areas are separated by the estimation method. Convective regions are associated with intense rainfall and strong upward and downward air motions in the precipitating clouds, while stratiform rain is gentle and widespread, and the associated air motions are relatively weak. By initially classifying a given region as convective or stratiform based upon the microwave radiometer observations, a more specific precipitation or latent heating estimate can be made. An improved technique for separating convective and stratiform regions using the horizontal variation of observed microwaves and their polarization is employed in the revised estimation method.

The improved method is scrutinized to determine what levels of error can be expected in estimated precipitation and latent heating. If rain rate estimates are averaged over 0.5° x 0.5° boxes, their errors are expected to be about 50% for a 1 mm per hour rain rate but decrease to 20% at 14 mm per hour. If the rain rate estimates are accumulated in 2.5° x 2.5° boxes and then averaged over 1 month, the anticipated errors in the resulting monthly rain rates can be up to 35% for moderate rain rates. The primary source of error in the monthly-average rain estimates is the limited sampling, about 1 per day, by the

satellite microwave radiometer. Latent heating estimates are expected to have errors that are slightly greater than errors in rain rate estimates.

The theoretical error estimates for precipitation and latent heating derived in Part I serve as benchmarks for estimates based upon the independent measurements described in Part II. Surface rain rates estimated using the satellite microwave radiometer observations are compared to both ground-based radar estimates and estimates from satellite-borne radar. Comparing estimates averaged over 0.5° x 0.5° boxes, the improved method yields estimates with less bias than forerunning methods. The bias of monthly-averaged rain rates in 2.5° x 2.5° boxes is similarly reduced. These are significant results, since applications such as the assimilation of rain estimates into computer weather forecasting models require estimates with a minimum of bias, although randomly varying errors can be tolerated. The estimates of error in Part I for 0.5° x 0.5° rain rates are consistent with the errors derived from independent data, with the exception of extremely light rains. The theoretical estimates of the sampling error in monthly-average rain rates over 2.5° x 2.5° represent only a fraction of the error indicated by comparisons with the independent data: apparently small biases in the algorithm estimates due to residual errors in the cloud model reference library are exposed in monthly averages.

Latent heating rates derived from the satellite passive microwave observations are compared to estimates from ground-based Doppler radars as well as those from a balloon sounding network. The Doppler radars provide measurements of the horizontal winds over a precipitation region; these winds, in turn, can be used to infer the strength of updrafts and downdrafts, since the mass of air must be conserved. Ascending air is associated with positive latent heating (warming of the air which creates buoyancy), while descending air is associated with negative latent heating (cooling from evaporation of rain which creates denser, sinking air). In general, the satellite latent heating estimates are consistent with the Doppler radar estimates, but an overestimation of heating at high altitudes (greater than 8 km) is also indicated.

The balloon sounding network measures the atmosphere's response to latent heating over a fairly large area (about 10° x 5° in this study). Again, rising and sinking air determined from the network are the primary indicators of latent heating and cooling, respectively. A balloon network in the South China Sea operated for almost two months during 1998, and the time-variation of heating derived from this network is similar to the time-variation of latent heating based upon the satellite microwave radiometer estimates. The limited sampling by the satellite radiometer contributes to differences between the network and satellite heating estimates, however.