TRMM Sampling of Radar–AMeDAS Rainfall Using the Threshold Method

RIKO OKI*

NASA/Goddard Space Flight Center, Laboratory for Atmospheres, Greenbelt, Maryland

AKIMASA SUMI

Center for Climate System Research, University of Tokyo, Meguro, Tokyo, Japan

DAVID A. SHORT

NASA/Goddard Space Flight Center, Laboratory for Atmospheres, Greenbelt, Maryland

(Manuscript received 7 November 1996, in final form 28 March 1997)

ABSTRACT

It is known that spatially averaged rainfall rate $\langle R\rangle$ is highly correlated with the fractional area (F) of rain rate exceeding a preset threshold (τ), when the area is large enough to include numerous convective systems in various stages of their life cycles. Using this fact, a method to estimate area-averaged rain rate from $F(\tau)$, which is obtained from satellite observations, is proposed for Tropical Rainfall Measuring Mission (TRMM). There have been numerous studies investigating F–$\langle R\rangle$ relationships and optimal thresholds at several radar observation sites around the world but no studies to confirm the performance of the method within Japan. In this study an analysis of radar–AMeDAS (Automatic Meteorological Data Acquisition System) precipitation data is presented. The F–$\langle R\rangle$ relationships of radar–AMeDAS rain data have been examined systematically, with the result that the optimum threshold that maximizes the correlation between $\langle R\rangle$ and $F(\tau)$ is near 3.5 mm h^{-1} in every year and season of available data.

Using the threshold method with the average coefficients obtained when the threshold is set to 3.5 mm h^{-1}, TRMM sampling of radar–AMeDAS rainfall is simulated. Fixing $5^\circ \times 5^\circ$ areas, monthly mean area-averaged rain rate is estimated from the observational coverage that would be obtained by TRMM during a month. The errors from the threshold method are only 3%–4% larger than the sampling errors (14%–19% on average) obtained by using the full dynamic range of observed rain rates. Considering the dynamic range of TRMM sensors, the threshold method would be an effective method to estimate area-average rain rate.

1. Introduction

It is known that spatially averaged rain rate $\langle R\rangle$ is highly correlated with the fractional area (F) of rain rates exceeding a preset threshold (τ), when the area is large enough to include numerous convective systems in various stages of their life cycles (Chiu 1988; Atlas et al. 1990).

Using this fact, estimation of area-averaged rain rate $\langle R\rangle$ from $F(\tau)$, which is obtained from satellites and ground-based radars is possible. This threshold method is thought to be one of the more effective means for rain-rate estimation over large areas using satellite observations.

Because of the limitation of the dynamic range of satellite sensors, it is especially difficult to design an instrument that can accurately measure both light and heavy rain. The threshold method does not require exact rain-rate measurements but only needs information on whether rain rate is larger than the preset threshold. Thus, it may work well to overcome instrument limitations and is proposed for the Tropical Rainfall Measuring Mission (TRMM; Simpson et al. 1988).

Empirical studies of thresholding similar to Chiu's (1988) were reported by Doneaud et al. (1981, 1984). They showed that the area–time integral (ATI) method can estimate the total rainfall of a storm throughout its lifetime by integrating rain rates greater than a threshold in area and time. Chiu (1988) showed the empirical fact that spatially averaged rain rate $\langle R\rangle$ is highly correlated with the fractional area (F) of rain rates exceeding a preset threshold using the Global Atlantic Tropical Experiment (GATE) radar data and explored the threshold method to estimate instantaneous area mean rain rate on the basis of the ATI. Atlas et al. (1990) proposed a

* Additional affiliation: Universities Space Research Association.

Corresponding author address: David A. Short, Code 910.1, NASA/GSFC, Greenbelt, MD 20771.
E-mail: short@trmm.gsfc.nasa.gov

unified theory for the estimation of both the total rainfall from an individual convective storm over its lifetime and the areawide instantaneous rain rate by use of measurements of the area coverage of the storms exceeding a threshold rain intensity (Chiu 1988). Those methods are based upon the existence of a well-defined probability density function (pdf) of rain rate appropriate to the climatic regime. Rosenfeld et al. (1990) have shown high correlations between area-averaged rain rates and the fractional area covered by rain in excess of a specified isopleth of rain rate τ using radar data from several radar sites. They also showed that to realize high accuracy, the domain has to be large enough (about 10^4 km^2) to include a representative sample of rain cells in different stages of their life cycles. Braud et al. (1993) and Braud et al. (1994) investigated the spatial distribution and the variability of rain rates within τ-thresholded areas.

Kedem et al. (1990) gave a theoretical explanation of the linear relationship between F and $\langle R\rangle$, assuming that rain rate has a mixed distribution. Also they have outlined a method for choosing the threshold level optimally by minimizing the change in the slope $[\beta(\tau)]$ across distributions. Kedem and Pavlopoulos (1991) and Short et al. (1993a) suggest two asymptotic procedures for deriving optimal thresholds under the same assumption as Kedem et al. (1990). The level that minimizes the variance of $\beta(\tau)$ also maximizes the correlation between F and $\langle R\rangle$. Short et al. (1993a) demonstrated that theoretically derived thresholds determined by utilizing the developments of Kedem and Pavlopoulos (1991) agree with empirically observed optimal thresholds. Short et al. (1993b) examined the relations between $\langle R\rangle$ and $F(\tau)$ using a network of rain gauges in Darwin, Australia, and showed excellent agreement between those F–$\langle R\rangle$ points and the slope predicted from the RRD of a different period. Shimizu et al. (1993) proposed single and double thresholds for the quadratic estimation of the area-average rain-rate variance. Shimizu and Kayano (1994) examined the maximum-likelihood estimator of the slopes for the first and second moments. Linear relationships between time-averaged rain rate and the fraction of time raining have also been observed from rain gauge data (Morrissey et al. 1994).

Some papers focused on actual applications of the threshold method. Krajewski et al. (1992) investigated the accuracy of the method accounting for ''radar effects'' on reflectivity measurement errors and the conversion of radar reflectivity to rain rate with the use of a stochastic space-time model of rainfall (Bell 1987) as the true rainfall-field generator. They suggested that for even high errors in the Z–R relationship, the area-threshold method with a low threshold (near zero) provides good performance, since radar is better at distinguishing rain/no rain than rain above or below a threshold rain rate. Recently, Morrissey (1994) studied the effect of data resolution on the area-threshold method using a stochastic model (Bell 1987) of radar snapshots of rainfall and showed that it is essential that in the application of the method, the climatologically determined $\beta(\tau)$ should have the same resolution as the satellite instruments. Meneghini and Jones (1993) have described and analyzed a multiple-threshold approach. In this method, the rain-rate distribution function is estimated with the use of multiple thresholds and a model function such as the lognormal or gamma. Any statistic of the areawide rain rate can be obtained from the fitted distribution. In the TRMM standard dataset, monthly mean rain rates over $5^\circ \times 5^\circ$ areas by this method will be included.

In this paper, radar–AMeDAS precipitation data, which are radar-based rain maps over Japan, were used. Previously there has been no study of the threshold method using rain data from Japan. One purpose of the present investigation is to determine the statistical parameters and performance of the threshold method in a climatic regime to be observed by TRMM. Parameters of the threshold method were determined empirically from the data and also determined assuming the lognormal distribution following some previous theoretical studies, and then consistency between those was checked. Also, the southern part of Japan will be observed by the TRMM satellite; therefore, this data have been used for a realistic simulation of TRMM sampling with the threshold method. There have been no previous attempts at applying the threshold method to the realistic situation of sparse observations from a low inclination, low elevation satellite. We report the results of area-averaged rain-rate estimation by the threshold method using radar–AMeDAS precipitation data and the TRMM orbit.

Due to the limitation of satellite-borne sensors, the dynamic range of rain-rate measurements from the TRMM satellite will be bounded. For example, the TRMM precipitation radar (PR) operating at 14 GHz will observe well from about 1 to 30 mm h^{-1} at the top of rain and to 20 mm h^{-1} at the bottom of rain column where a linear relationship between echo intensity and rain rate is expected (Okamoto et al. 1988). Also in the case of the TRMM Microwave Imager (TMI), brightness temperatures will be saturated around 20 mm h^{-1} (Yamasaki and Wilheit 1990).

However, in the threshold procedure, the only information needed is whether the observed rain rate is larger than the threshold or not. It can be said that the threshold method covers deficiencies of satellite sensors by the use of statistics. Even if advanced satellite sensors with increased dynamic range are developed in the near future, the method will be still an effective approach to estimate area-average rain rates due to its simplicity.

Section 2 describes the data and its rain-rate distribution (RRD) characteristics. Optimal thresholds of the radar–AMeDAS data are obtained empirically (section 3) and theoretically (section 4). Using the thresholds and regression coefficients, a realistic TRMM sampling simulation with the threshold method is performed and results described in section 5. Section 6 presents a summary.

2. Data

a. Radar–AMeDAS precipitation data

The radar–AMeDAS precipitation data represent 1-h accumulations estimated from observations of radars and rain gauges (Obayashi 1991; Takemura et al. 1984; Takase et al. 1988; Makihara et al. 1995; Makihara et al. 1996; Makihara 1996). The data cover all of Japan and its surrounding coastal waters. The spatial resolution is approximately 5 km × 5 km. The Japan Meteorological Agency (JMA) produces digital radar–AMeDAS precipitation data and charts routinely for the purpose of short-term rainfall prediction.

As of April 1991, JMA operated 18 radar sites on land and two on ships, covering the entire Japanese archipelago. The operation of shipborne radars is limited to winter and the Baiu season (early summer). There are eight observations (six observations at the site of Mount Fuji) per hour. By integrating eight (or six) observations, 1-h radar estimated rain-rate data are computed. AMeDAS is the network consisting of an automated operation center and automatic observation stations located all over Japan, linked through a telephone network. The time interval of data acquisition is once per hour. The AMeDAS network includes about 1300 rain gauge stations all over Japan, with a mean spatial interval of 17 km.

Rain rates observed by radars are adjusted to have the same value at each point in overlapping areas covered by multiple radars and calibrated by the use of AMeDAS rain gauge data. The difficulties inherent in quantitative radar observations of rainfall are compensated for by the calibration process with the use of rain gauge data. This estimated 1-h rain-rate database has wide-area coverage and long temporal coverage, preferable for satellite validation studies. The data has been generated since 1988. The period of the data used in this study covers 43 months, from March 1988 to September 1991. An example of the radar–AMeDAS data used here can be seen in Oki and Sumi (1994; see their Fig. 1).

The GATE radar data, which Chiu (1988) used and in which the highly correlated F–$\langle R \rangle$ relationship was found, were snapshots taken every 15 min. On the other hand, differing from the GATE data, radar–AMeDAS is 1-h accumulated rainfall. Nevertheless, it can be shown that any temporally averaged data can be used to find a F–$\langle R \rangle$ relationship, using the threshold method. However, we have to be careful that the same effect that Morrissey (1994) studied regarding spatial resolution can occur due to temporal resolution when the results here are compared with other results. The radar–AMeDAS spatial resolution is nearly the same as the GATE radar data or the TRMM PR.

There have been several reports regarding the accuracy of the radar–AMeDAS precipitation data (Kitabatake and Obayashi 1991; Yamamoto 1991; Makihara 1993; Makihara 1996). Compared with observations of rain gauges in the Tokyo metropolitan area, the estimated rain rates of radar–AMeDAS tend to be larger than those by gauges, however the overall accuracy is reliable (Kitabatake and Obayashi 1991). On a monthly mean basis, estimated monthly mean rain rates by radar–AMeDAS are larger than monthly mean rain rates from AMeDAS rain gauges by 0%–40%. Also, there is a data jump in April 1990 due to the change in data processing procedures. Before that the radar–AMeDAS monthly rain rates agree with AMeDAS estimated monthly rain rates, and after that radar–AMeDAS rain rates are larger than AMeDAS gauges by about 20%. Makihara (1993) showed that radar–AMeDAS composites correspond to 105% of AMeDAS gauges in the same grids at weak rain rates and 120% at strong rain rates using 3-h composite radar–AMeDAS data. In the data processing to make the digital radar–AMeDAS precipitation data, the grid size is changed from the original 2.5-km radar grid into a 5-km grid. In this procedure, the maximum value among the four 2.5-km pixels is chosen in order not to fail to detect local severe precipitation. According to Makihara's (1993) account, this error is unavoidable because of the maximum value selection in the data processing.

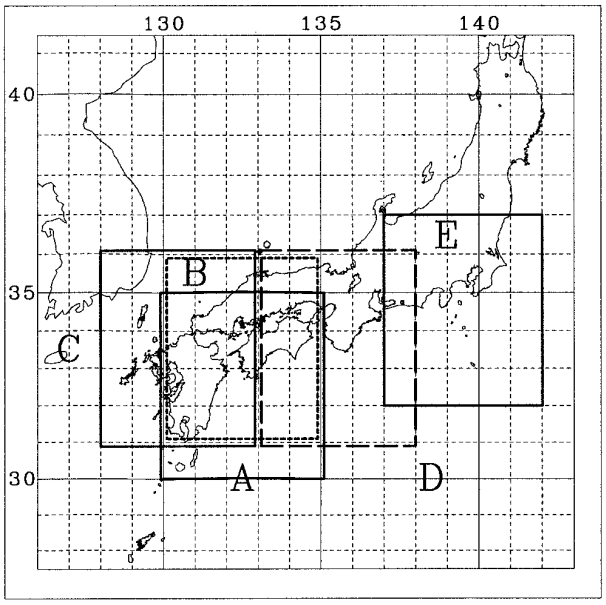

FIG. 1. The five test areas of 5° × 5° from A to E.

b. Conditional rain-rate distributions (RRD)

To determine the distribution characteristics of rain rate, when it is raining, normalized distributions of the cumulative rainfall by depth and by number of observations were obtained from the hourly, 5-km pixels within 5° × 5° longitude and latitude areas. Five 5° × 5° areas named from A to E, respectively, are shown in Fig. 1.

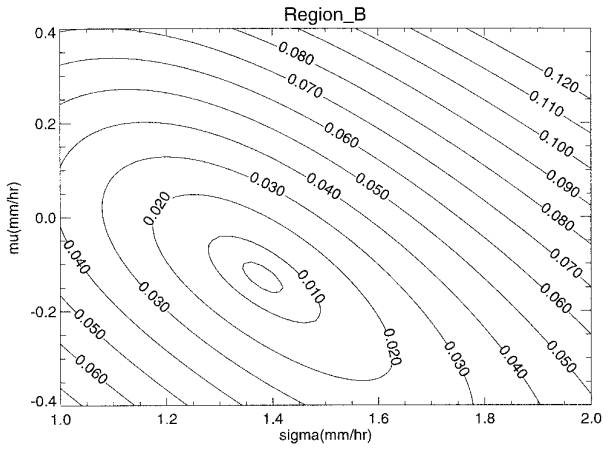

FIG. 2. The rms difference between CDF from the data and a model (lognormal) CDF in μ and σ space. For region B, the best fit is found for $\mu = -0.13$ and $\sigma = 1.38$.

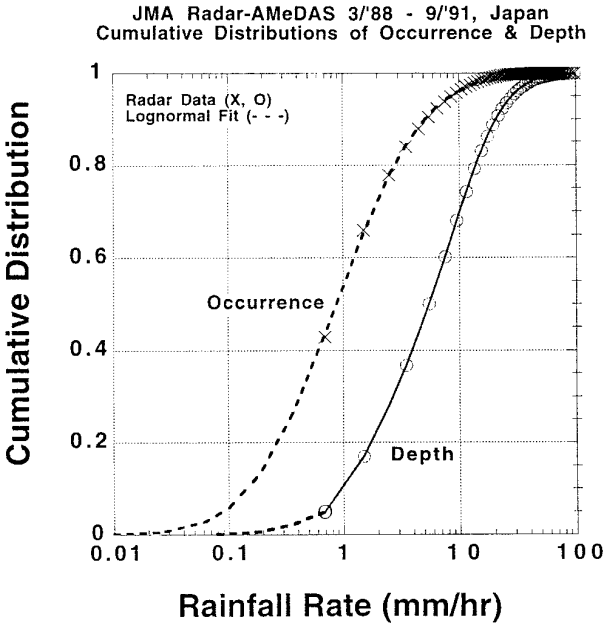

FIG. 3. Normalized distributions of the cumulative rainfall by depth and occurrence. The percentage of rain rates observed below 1 mm h^{-1} is 53% in occurrence and only about 12% in depth.

The weakest rain-rate category of radar–AMeDAS is 0.0 mm $h^{-1} < r < 0.7$ mm h^{-1}. There is no information on the distribution between 0.0 and 0.7 mm h^{-1}. By fitting a lognormal distribution to the available categories and extrapolating to the lowest category, the shape of the distribution between 0 and 0.7 mm h^{-1} was obtained. Applying the method of Meneghini and Jones (1993), for the range of values $-0.5 < \mu < 0.5$, $1.0 < \sigma < 2.0$, we searched for the pair of parameters (μ and σ) that realized the minimum rms difference between the lognormal function and the cumulative distribution from actual data ranging from 0.7 to 14.5 mm h^{-1}. Figure 2 shows the rms contours in μ and σ space. For region B, the best fit is found for $\mu = -0.13$ and $\sigma = 1.38$.

Using these obtained μ and σ, the cumulative distribution of occurrence in the lowest category was estimated. The cumulative distribution of depth was obtained from the occurrence. The graph for region B is shown in Fig. 3. The percentage of rain rates observed below 1 mm h^{-1} is 53%. The conditional fraction of rain rates less than a specified threshold (τ) will be referred to as $f(\tau)$. Rain rates below 1 mm h^{-1} accounted for 12% of rain depth.

c. Conditional mean rain rate

It is known that optimal thresholds that give the highest correlation coefficients in F–$\langle R \rangle$ relationships are near the conditional mean rain rate (Kedem and Pavlopoulos 1991; Short et al. 1993a). Conditional mean rain rates from radar–AMeDAS are shown in Table 1. A conditional mean was calculated by multiplying the middle value of each rain intensity category by the count in that category (simple average). However, simple averaging of radar–AMeDAS data can cause an overestimation of the conditional mean rain rate because radar–AMeDAS does not have information on the distribution below 0.7 mm h^{-1} as mentioned before. Another estimate of the conditional mean was calculated (from the CDF), using the extrapolated rain distribution, which can be seen in the lowest rain rates of Fig. 3. This value is 2.135 mm h^{-1}. Also, if a uniform distribution of rain rates below 0.7 mm h^{-1} was assumed, the conditional mean is 2.176 mm h^{-1}. If all rain rates below 0.7 mm h^{-1} were truncated to zero, the conditional mean is 3.547 mm h^{-1}.

3. Optimal threshold of radar–AMeDAS precipitation

In order to apply the threshold method, a high correlation between F and $\langle R \rangle$ is needed, because the F–$\langle R \rangle$ relationship is the basis of the method. However, there have been no studies of F–$\langle R \rangle$ relationships of rain in the Japan area until now. The following sections describe an analysis of F–$\langle R \rangle$ relationships using the Japanese radar–AMeDAS precipitation data.

a. Test areas

Five 5° × 5° test areas were chosen in the area of the radar–AMeDAS precipitation charts (already shown

TABLE 1. The conditional mean rain rate in region B (mm h^{-1}).

From CDF			Simply averaged
Lognormal	Uniform	Truncated	
2.135	2.176	3.547	2.157

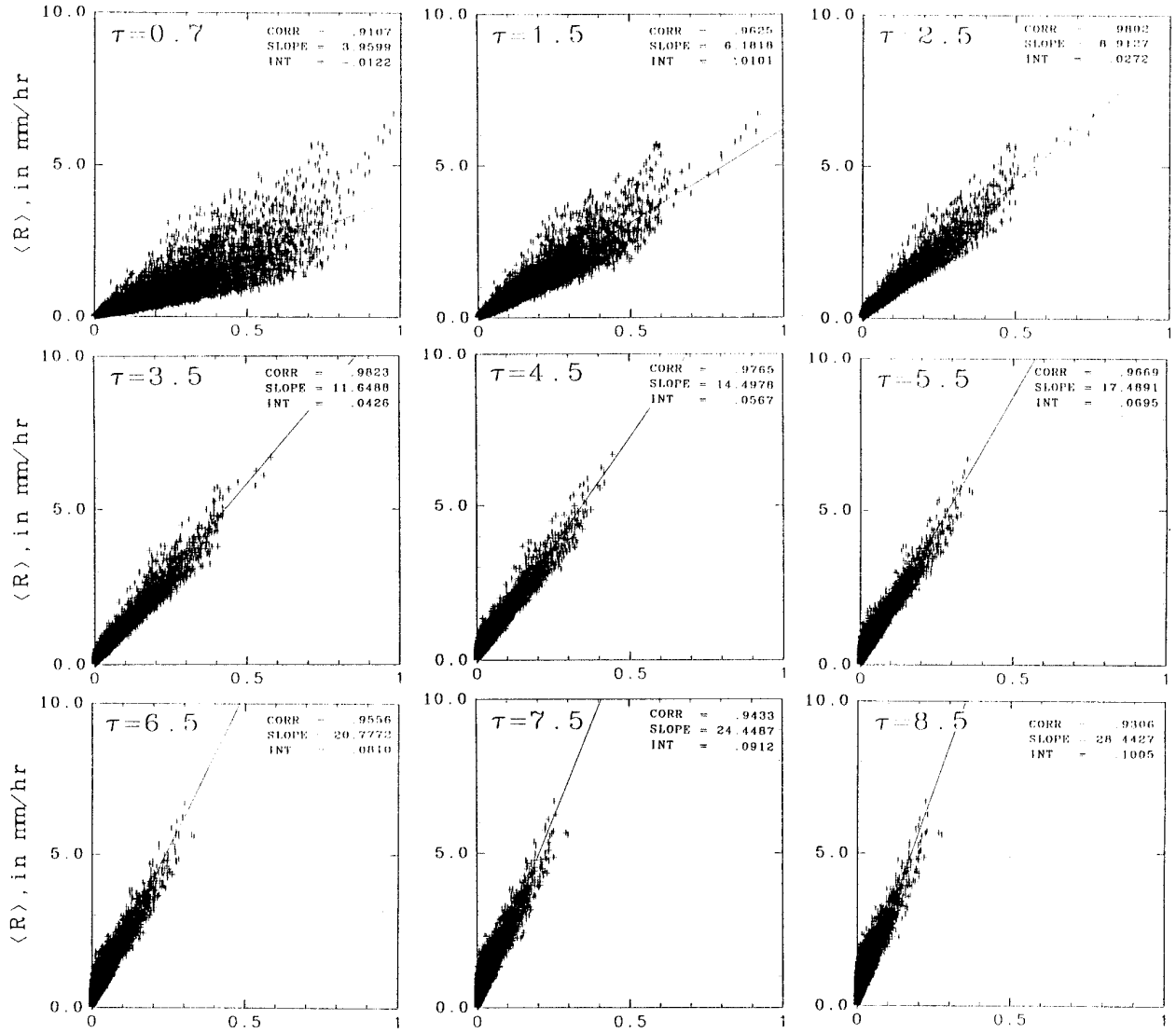

FIG. 4. Area-average rain rate versus fractional coverage of rain rates exceeding τ mm h^{-1} (0.7–14.5 mm h^{-1}) over a 5° × 5° subset of the radar–AMeDAS hourly rain rates, from March 1988 to September 1991, region B.

in Fig. 1). TRMM sensors will cover the area from 38°N to 38°S in latitude. Thus these 5° × 5° areas will be observed by the TRMM satellite. These areas are the same ones used in Oki and Sumi (1994), in which TRMM sampling errors were estimated. Also, five areas of 2.5°, 1.0°, and 0.5° latitude–longitude were selected (not shown).

b. F–⟨R⟩ relationship of radar–AMeDAS

A systematic study of the relationship between $F(\tau)$ and $\langle R\rangle$, similar in nature to those previously done for GATE and other areas, was performed on the radar–AMeDAS data. Setting a threshold value (mm h^{-1}), the number of pixels where the rain rates exceeded the threshold were counted in each hourly rain-rate map of each test region. The ratio of the counted number to the total number of pixels in a test region is $F(\tau)$, the fractional area exceeding the threshold. At the same time, $\langle R\rangle$ was calculated. Applying linear regression to the F–$\langle R\rangle$ pairs, that is,

$$\langle R\rangle = \alpha(\tau) + \beta(\tau)F(\tau), \qquad (1)$$

the coefficients (slope β and intercept α) were obtained. Also, correlation coefficients were calculated. From previous studies, the threshold that gives the highest correlation is considered as the optimal threshold. In this study, the same criterion regarding the optimal threshold was used. Changing the threshold level from 0.7 to 14.5 mm h^{-1}, the threshold giving the highest correlation coefficient was sought in each region.

All F–$\langle R\rangle$ pairs during the whole period in region B are plotted in Fig. 4. In each graph, the correlation coefficient (CORR), the slope (SLOPE) and the intercept

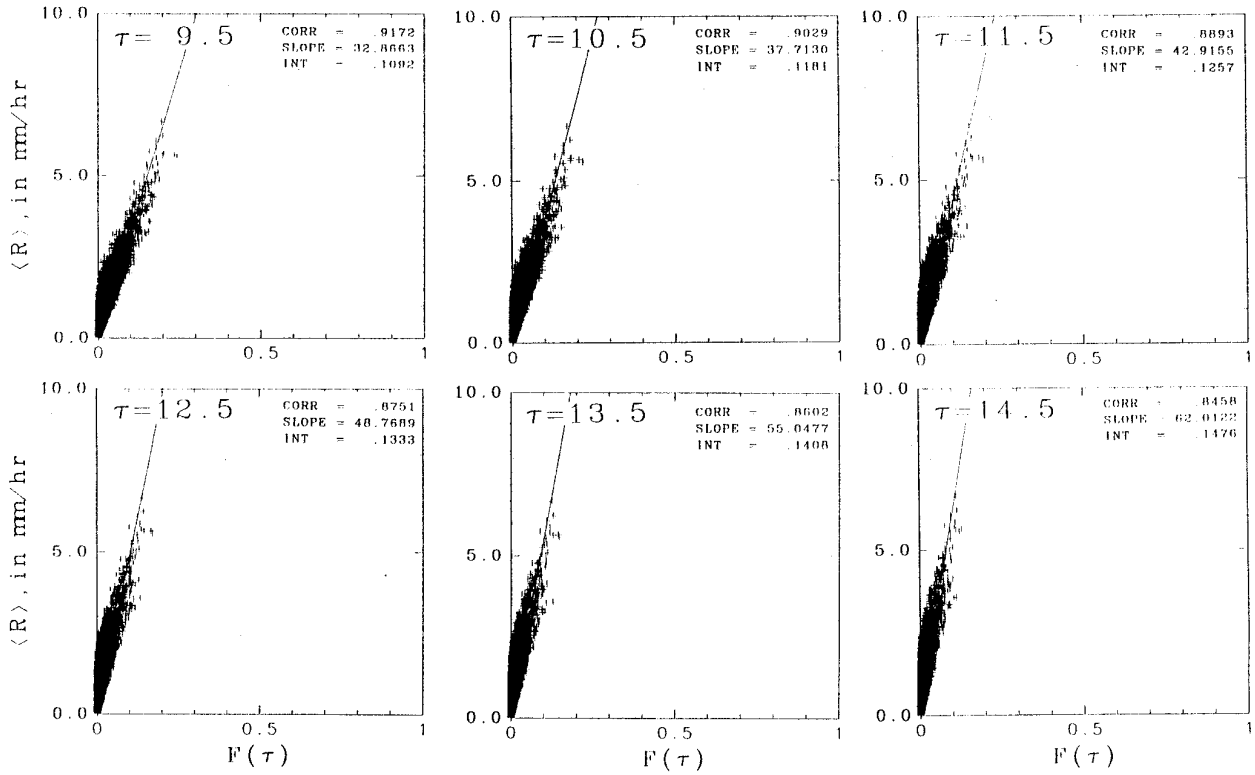

FIG. 4. *(Continued)*

(INT) are written. These coefficients are summarized in Table 2. The correlation coefficient is maximized when τ is 3.5 mm h^{-1}. For this optimal threshold, the correlation coefficient is very high and larger than 0.98. The intercepts are quite close to 0 mm h^{-1}. These results are similar to previous studies [for example Chiu (1988) and Rosenfeld et al. (1990)], showing that the threshold method will work well also in the Japan area. As τ becomes larger, the slope coefficient increases. Intercepts are almost zero for all τ.

There can be some influences of data accuracy on parameters of the threshold method obtained here. Makihara (1996) showed that radar–AMeDAS precipitation overestimates AMeDAS (rain gauge) measurement by 11% at 10 mm h^{-1} and by 12% at 40 mm h^{-1}, although overestimation decreases to 8% at 5 mm h^{-1}. Thus, the amount of overestimation is small and the overestimation occurs almost uniformly in every category. In such a case, true parameters could be recalculated from ''adjusted'' data.

TABLE 2. Thresholds, correlation coefficients, slopes, and intercepts for region B.

		$R = \alpha + \beta F$		$R = \beta F$
τ (mm h^{-1})	Correlation	Slope (mm h^{-1})	Intercept (mm h^{-1})	Slope (mm h^{-1})
0.7	0.911	3.96	−0.012	3.92
1.5	0.962	6.18	0.011	6.23
2.5	0.980	8.91	0.027	9.07
3.5	0.982	11.65	0.043	11.97
4.5	0.976	14.50	0.057	15.01
5.5	0.967	17.49	0.070	18.23
6.5	0.956	20.78	0.081	21.79
7.5	0.943	24.45	0.091	25.77
8.5	0.931	28.44	0.100	30.12
9.5	0.917	32.87	0.109	34.96
10.5	0.903	37.71	0.118	40.29
11.5	0.889	42.92	0.126	46.01
12.5	0.875	48.77	0.133	52.48
13.5	0.860	55.05	0.141	59.46
14.5	0.846	62.01	0.148	67.19

c. Seasonal dependence of the optimal threshold

The variability of correlation coefficients between F and $\langle R \rangle$ with season is shown in Table 3. This is the case of region A. The thick font expresses the highest correlation coefficient in each season. The optimal thresholds are a little bit larger in summer seasons, however seasonal dependence is not so clear.

Even as the threshold is changed from 1.5 to 4.5 mm h^{-1}, the correlation coefficient is very high and still larger than 0.96. Thus in the Japan area, we can fix the threshold to 2.5 or 3.5 mm h^{-1} through all seasons. This robust characteristic is one advantage of the method.

TABLE 3. The variability of correlation coefficient with season.

τ	Mar–May			Jun–Aug			Sep–Nov			Dec–Feb		
(mm h^{-1})	1988	1989	1990	1988	1989	1990	1988	1989	1990	1988	1989	1990
0.7	0.92	0.97	0.95	0.96	0.90	0.94	0.94	0.96	0.95	0.97	0.97	0.90
1.5	0.97	**0.99**	0.98	0.98	0.96	0.97	0.97	0.98	0.97	**0.99**	0.99	0.94
2.5	**0.98**	0.98	**0.98**	0.98	0.98	0.99	**0.98**	0.99	0.99	0.98	**0.99**	0.97
3.5	0.97	0.97	0.98	0.98	**0.98**	**0.99**	0.98	**0.99**	0.99	0.97	0.98	0.98
4.5	0.95	0.95	0.97	**0.99**	0.98	0.99	0.97	0.99	**0.99**	0.96	0.96	**0.98**
5.5	0.93	0.92	0.95	0.98	0.98	0.99	0.96	0.99	0.99	0.95	0.95	0.98
6.5	0.91	0.89	0.93	0.98	0.97	0.98	0.94	0.98	0.99	0.93	0.93	0.97
7.5	0.89	0.86	0.92	0.97	0.96	0.98	0.93	0.97	0.98	0.92	0.92	0.97
8.5	0.87	0.84	0.90	0.96	0.95	0.97	0.91	0.97	0.98	0.91	0.91	0.96
9.5	0.85	0.81	0.88	0.95	0.93	0.97	0.90	0.96	0.97	0.89	0.90	0.95

d. Variability of β

In practical applications of the method, it would be preferable to have minimum changes in threshold levels. In the previous section, it was found that the threshold level near Japan can be fixed without sacrificing accuracy. If the threshold level is fixed, how much does the slope of the linear regression change from place to place or season to season? In Table 4, slopes for each season and year are shown for thresholds from 1.5 to 4.5 mm h^{-1}. Values written in bold fonts are the optimal thresholds in each season. One expression for the variability of beta is the following:

$$\frac{\beta - \overline{\beta}}{\overline{\beta}}.$$

Figure 5 shows this variability fixing the threshold levels from 2.5 to 4.5 mm h^{-1}. When τ is 3.5 mm h^{-1}, the variability of β is smallest.

Using a β whose variability is small will lead to minimal estimation errors if a fixed β is used in the method. Kedem and Pavlopoulos (1991) showed that an optimal threshold can be defined where the variance of $\beta(\tau)$ is minimized.

When τ is 2.5 mm h^{-1}, the variability has seasonal changes and a small trend. At other thresholds such seasonal change doesn't appear. Considering the trend, the quality of the data also should be checked.

e. Area size

Optimal thresholds for various size areas were calculated and the correlation coefficients of F–$\langle R\rangle$ relationships are shown in Table 5. The area sizes are 2.5° × 2.5°, 1.0° × 1.0°, 0.5° × 0.5°, and 5.0° × 5.0°. It is obvious that higher correlation coefficients are obtained in larger areas. The optimal thresholds that realize the highest correlation are 3.5 mm h^{-1} in 5.0° × 5.0° regions and 4.5 mm h^{-1} in 1.0° × 1.0° regions. The optimal threshold does not change much with area size.

f. Slopes predicted from the RRDs

Using the RRD graph obtained in section 2, slopes of F–$\langle R\rangle$ relationship were obtained following the method of Short et al. (1993a). The general thresholding model is

$$\langle R\rangle = \beta(\tau)F(\tau).$$

For region B, the conditional mean rain rate $[R]$ is 2.135 mm h^{-1}, calculated using an interpolation at weak rain rates. The slope predicted from the RRD is given by the following:

$$\beta(0.7) = \frac{[R]}{f(0.7)} = \frac{2.135 \text{ mm h}^{-1}}{0.47} = 4.54 \text{ mm h}^{-1}. \quad (2)$$

These values are summarized in Table 6. Slopes obtained here are slightly larger than those observed in the F–$\langle R\rangle$ relationships determined by regression (Table 2); however, it can be said that estimates by the two different techniques still have a good agreement.

4. Theoretically derived optimal thresholds

Optimal thresholds were determined theoretically assuming the lognormal distribution. The derivation of the following formulas is shown in Kedem and Pavlopoulos (1991) and Short et al. (1993a). The lognormal proba-

TABLE 4. The variability of β in each season, region A.

τ	Mar–May			Jun–Aug			Sep–Nov			Dec–Feb		
(mm h^{-1})	1988	1989	1990	1988	1989	1990	1988	1989	1990	1988	1989	1990
1.5	5.5	**4.9**	5.6	6.5	6.8	7.4	5.9	6.8	7.3	**5.1**	5.5	6.2
2.5	**8.0**	7.4	**8.3**	8.7	8.9	10.0	**8.8**	8.9	9.8	7.8	**8.2**	9.9
3.5	10.9	10.9	11.1	11.4	**11.1**	**12.2**	12.2	**11.0**	12.1	11.5	11.1	13.5
4.5	14.6	16.6	14.3	**14.8**	13.4	14.6	16.5	13.3	**14.5**	16.9	14.3	**16.6**

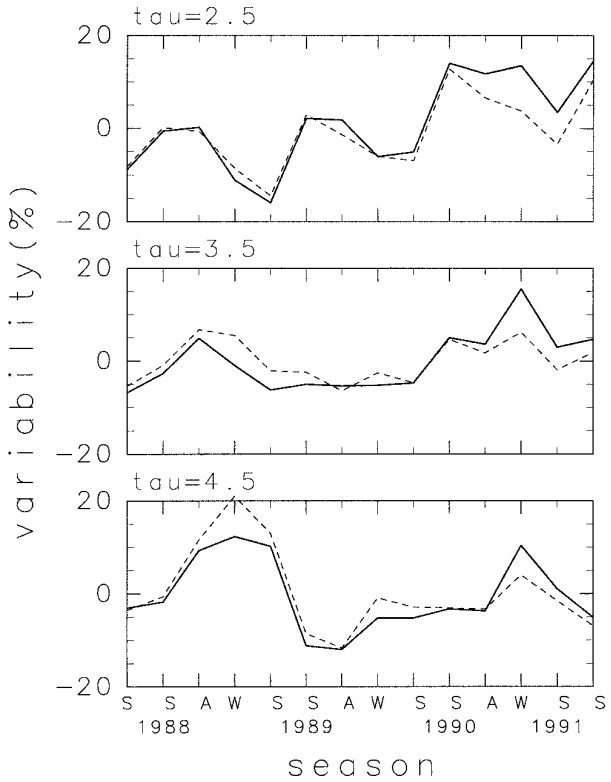

FIG. 5. The variability of beta for each year and season. Solid line: region A; dashed line: region B.

bility density function is given below with $-\infty < \mu < \infty$, $\sigma > 0$ (Crow and Shimizu 1988; Aitchison and Brown 1963).

$$g(x) = \frac{1}{\sqrt{2\pi}\,\sigma x} \exp\left\{\frac{-[\log(x) - \mu]^2}{2\sigma^2}\right\}. \quad (3)$$

It is useful to define

$$u \equiv \frac{\log(\tau) - \mu}{\sigma}. \quad (4)$$

The use of (3) and (4) in evaluating the slope coefficient for the probability distribution of rain rate results in

$$\beta_{\boldsymbol{\theta}}(\tau) = \frac{\exp[\mu + (\sigma^2/2)]}{1 - \Phi(u)}, \quad (5)$$

where $\Phi(u)$ is the cumulative distribution function for the standard normal distribution, and $\boldsymbol{\theta}$ is the parameter vector (μ, σ). If μ and σ are known, $\beta_{\boldsymbol{\theta}}$ can be calculated using (5).

By a polynomial fit to values in Table 2 of Kedem and Pavlopoulos (1991), Kedem and Short (1993) obtained the following approximation:

$$\tau_{\text{optimal}} = \exp(-0.322 - 0.014\sigma + 0.973\sigma^2 + \mu). \quad (6)$$

For region B, using the μ and σ obtained from the rain-rate distribution in this approximation, τ_{optimal} was calculated for region B and the value was 3.98 mm h^{-1}. There is a good agreement with the τ_{optimal} from regression.

TABLE 5. Dependence of correlation coefficient on the domain size.

τ (mm h^{-1})	Area size 5.0°	2.5°	1.0°	0.5°
1.5	0.96	0.94	0.89	0.82
2.5	0.98	0.96	0.93	0.87
3.5	0.98	0.97	0.95	0.90
4.5	0.98	0.97	0.95	0.91
5.5	0.97	0.96	0.95	0.91
6.5	0.96	0.95	0.94	0.91
7.5	0.95	0.94	0.93	0.90

Also using the obtained values of σ and μ, we can calculate the normalized asymptotic variance (Kedem and Pavlopoulos 1991; Short et al. 1993a),

$$\frac{\nu_{\boldsymbol{\theta}}(\tau)}{\beta^2_{\boldsymbol{\theta}}(\tau)} = \frac{\sigma^2}{[1 - \Phi(u)]^2}\left\langle\left\{[1 - \Phi(u)] - \frac{1}{\sigma}\phi(u)\right\}^2 + \frac{1}{2}\left\{\sigma[1 - \Phi(u)] - \frac{1}{\sigma}u\phi(u)\right\}^2\right\rangle, \quad (7)$$

and find its minimum, where $\phi(u)$ is the probability density function of the standard normal distribution, $\Phi(u)$. Figure 6 shows this variance for region B. Correlations are also shown in the same figure. The empirically determined optimal threshold level and theoretically derived level are both near 4 mm h^{-1}. This result confirms the sensibility of (7). The strength of the theoretical result is that it provides a method for deriving optimal thresholds from the probability distribution of rain rate, which can be found from rain gauge observations in any climatic regime.

5. Estimation of area-averaged monthly mean rain rates by the threshold method

Using the threshold method with optimal thresholds and parameters determined by linear regression from the previous section, a TRMM observational simulation was made.

In most previous papers on the threshold method,

TABLE 6. Slopes estimated from conditional mean rain rates for region B. The $F(\tau)$'s are obtained from the RRD graph in section 2.

τ (mm h^{-1})	$F(\tau)$	$f(\tau)$	RRD slope
1.0	0.53	0.47	4.54
2.0	0.73	0.27	7.91
3.0	0.81	0.19	11.24
3.5	0.84	0.16	13.34
4.0	0.86	0.14	15.25
5.0	0.89	0.11	19.41

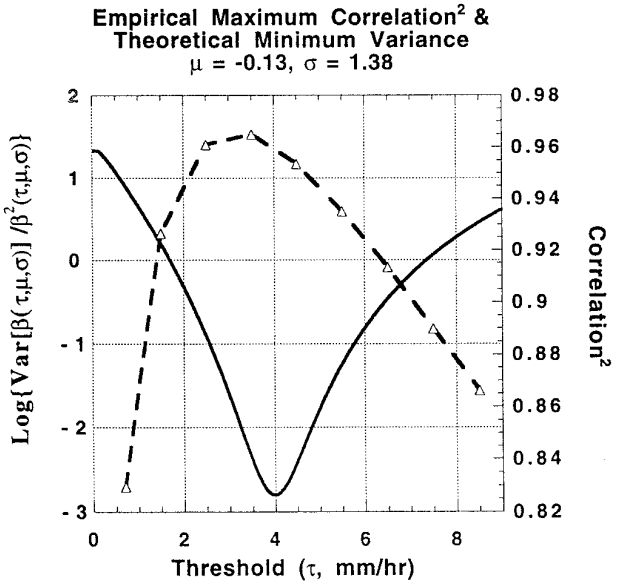

FIG. 6. Empirical correlation and normalized asymptotic variance versus threshold for region B.

realistic applications of the method using simulated satellite sampling and real rainfall data have not been explored. Considering realistic satellite observations, an entire region, for example, a 5° × 5° area, will not always be 100% covered by every overpass. Thus, we estimated the monthly mean area-averaged rain rate from observations of partial satellite visits to fixed areas. The threshold method was used only in the observed portion of each area.

a. Procedure of the estimation

1) TRMM ORBIT

We generated four types of orbits of the TRMM satellite, using an altitude of 350 km and an inclination of 35°. The four orbit types differ from each other by 90° of longitude in their starting point.

The data period used in this simulation was the same 43 months from March 1988 to September 1991 that was used for obtaining the optimal threshold values. The same 5° × 5° areas from A to E used in section 2 were also used. The swath width was set to 700 km, simulating the TMI.

2) THRESHOLD PARAMETERS

Parameters obtained in the previous section for each region were averaged over all seasons and a single β value was determined for each region for the simulation.

In Table 7, parameters used for each area are summarized. In each domain, the optimum threshold was 3.5 mm h^{-1}, and average parameters (slope and intercept) obtained by linear regression in the case of the 3.5 mm h^{-1} threshold were used. The slopes β are pretty stable across the five regions.

TABLE 7. Coefficients of linear regression for the estimation.

A	3.5	0.03	11.65
B	3.5	0.04	11.65
C	3.5	0.04	11.24
D	3.5	0.04	11.90
E	3.5	0.05	11.06

3) ALGORITHMS FOR MONTHLY MEAN AREA-AVERAGED RAIN RATE

As mentioned in the beginning part of this section, in a realistic satellite observation, an entire 5° × 5° area will not always be completely observed.

In fixing a predetermined domain, it is assumed that the domain will be observed N times in a month. In N times, various coverages are included. Let $\hat{R}_{th}(i)$ $(i = 1, 2, \ldots, N)$ be the averaged rain rate estimated by the threshold method in the observed portion at each visit. Then the monthly mean rain rate $\overline{\langle R \rangle}$ is obtained using $\hat{R}_{th}(i)$, $(i = 1, N)$ values for 43 months.

Three algorithms were used to calculate monthly mean rain rate. The first one was simply,

$$\overline{\langle R \rangle} = \frac{1}{N} \sum_{i=1}^{N} \hat{R}_{th}(i), \tag{8}$$

where each term in the sum has the same weight. In other words, each overpass in a month was considered as representative of the whole area.

In the second method, we omitted the observations in which only a small portion of the area was covered. The level was varied from 30% coverage, to 50% coverage, to 70% coverage.

The third algorithm was

$$\overline{\langle R \rangle} = \frac{\sum_{i=1}^{N} f_i \hat{R}_{th}(i)}{\sum_{i=1}^{N} f_i}, \tag{9}$$

where f_i is a weighting factor equal to the fraction of the domain observed by the satellite. This algorithm was supposed a priori to be the best.

The ''true'' monthly averaged rain rate in each area was assumed to be the rain rate calculated from all hourly radar–AMeDAS data during the month. The difference between the TRMM estimate using the threshold method and such a true value was computed. To express the differences, root-mean-square errors were used in the form of the ratio of the estimated to true monthly precipitation.

It should be noted that the results of this simulation, which will be shown in the next section, will include both sampling errors and errors arising from the threshold method.

TABLE 8. Root-mean-square errors (%) between the true means and estimated results by the threshold method. Also, estimated sampling errors (OS 1994) are listed. Here, OS (1994) used the same data, the same TRMM orbits, and the same test areas. These estimated results by the threshold method include these sampling errors, so the differences between threshold estimates and OS (1994) sampling errors are corresponding to errors due to the threshold method.

Area	Algorithm	Orbit				Mean	Averaged sampling error (OS 1994)
		000	090	180	270		
A	AVR	25.64	17.79	22.09	17.13	21.05	
	CUT	20.40	18.40	19.97	17.80	19.23	
	WGT	20.01	16.10	20.14	16.65	18.38	15.0
B	AVR	21.73	22.21	18.87	18.13	20.45	
	CUT	18.84	15.93	19.73	17.08	18.12	
	WGT	19.14	15.34	18.64	15.97	17.43	14.0
C	AVR	25.49	25.22	21.45	22.40	23.58	
	CUT	24.46	17.81	18.06	22.72	20.96	
	WGT	23.91	18.38	17.53	20.06	20.11	16.0
D	AVR	22.30	20.98	17.07	21.13	20.42	
	CUT	19.32	18.38	16.33	18.47	18.19	
	WGT	18.28	18.03	16.53	18.46	17.88	15.0
E	AVR	18.87	18.08	27.91	26.50	23.16	
	CUT	22.63	19.90	26.56	25.47	23.85	
	WGT	20.54	18.65	26.07	25.06	22.85	19.0

b. Estimation results by the threshold method

Table 8 shows the root-mean-square errors (%) between the true monthly averaged rain rates and the estimated rain rates by the threshold method, expressed as the ratio of estimated to true monthly precipitation. The values for each orbit separately and the column labeled ''mean'' are the values obtained by averaging all orbits. These are ensemble mean values using 43 months.

In the right column, sampling errors estimated by using the same period of radar–AMeDAS and the same orbital parameters (Oki and Sumi 1994, hereafter OS) are shown. Here, OS (1994) estimated sampling errors in the same way. However, monthly mean rain-rate estimations by the TRMM satellite were obtained by accumulating rain rates at every grid point covered by TRMM sensors for 1 month, not using the threshold method. That is, a monthly rainfall estimate at a radar–AMeDAS grid point, P_{TRMM}, was estimated as follows:

$$P_{\text{TRMM}}(x, y) = \frac{P(x, y)}{N(x, y)} \times 24 \times \text{number of days},$$

where $P(x, y)$ is an accumulated rainfall obtained from the simulated TRMM measurement, and $N(x, y)$ is the number of visits at (x, y). Monthly precipitation over the area $\overline{P}_{\text{TRMM}}$ can be estimated by averaging all grid-point values of P_{TRMM} within a prescribed test area:

$$\overline{P}_{\text{TRMM}} = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} P_{\text{TRMM}}(x_i, y_j).$$

Then, sampling errors were estimated, comparing such estimated monthly area-average rain rates with the true monthly area averages calculated from all data during 1 month in each test area. The same data period, the same orbits, and the same validation areas were used. That is, these sampling errors are averaged values of 43 months and 4 orbits.

In Table 8, note that the estimated results by the threshold method include these sampling errors. That is, the differences between threshold estimates and OS (1994) sampling errors correspond to errors due to the threshold method itself.

The first method (AVR) that uses an estimated average rain rate in a partial observed area as a mean rain rate of the entire area gives the least accurate estimates. To see the differences between partial averaged rain rates and the entire $5^\circ \times 5^\circ$ averaged rain rates, root-mean-square errors (mm h^{-1}) of the differences are shown in Fig. 7. The abscissa is the ratio of observed area to whole area. This is the case of region B. In the observations in which only small areas were observed, assuming a partial averaged rain rate to be a $5^\circ \times 5^\circ$ averaged rain rate can cause a large estimation error.

In the second method, the ratio that was ignored in the monthly averaging process was changed from 30%, to 50%, to 70%. Among these coverages, the result of 30% shown in Table 8 was the best. Figure 8 shows how many times region B was observed at each TRMM overpass with a 700-km swath for the whole period of data (43 months). The $5^\circ \times 5^\circ$ areas in Japan will have more than 80% coverage for most TRMM visits. The latitude around Japan is that observed most frequently by the TRMM satellite. The relationship between latitude and number of observations can be seen in OS (1994) or Bell et al. (1990). Thus, ignoring the data from small coverage observations, such as 30%, does not have much effect on estimates of the monthly mean rain rate.

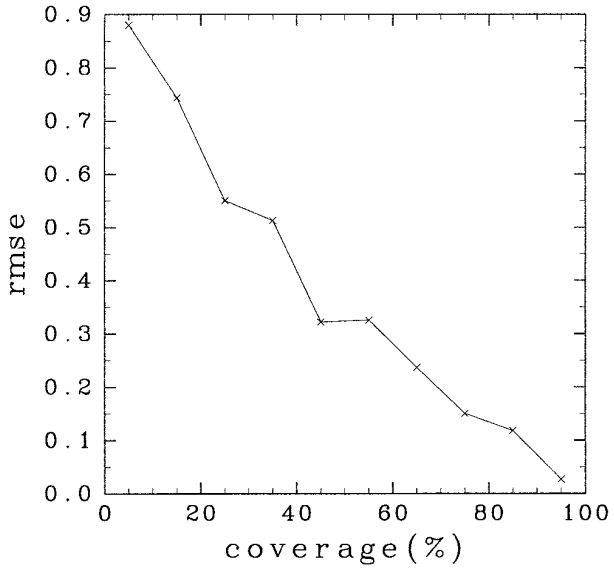

FIG. 7. Root-mean-square error (mm h^{-1}) between rain rate averaged in the entire area and estimated by the threshold method in the observed area. The *X* axis is the ratio of observed area to whole area. This is the case of region B (5° × 5°).

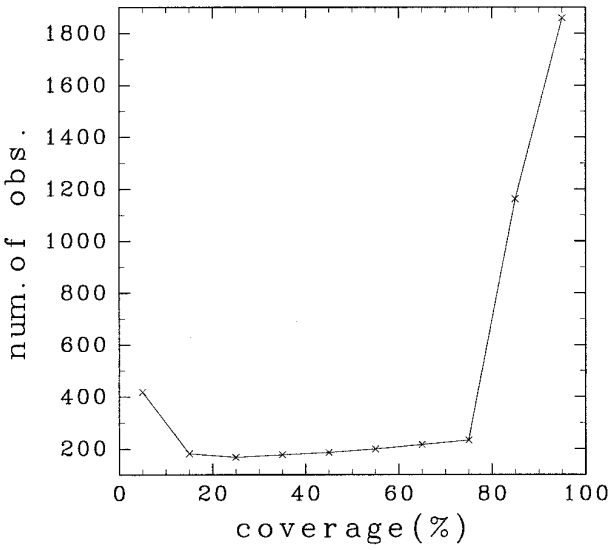

FIG. 8. The observed number of region B (5° × 5°) at each coverage during the whole period.

The third algorithm (WGT), in which a weighting factor proportional to the area coverage is multiplied by each observation, gives the most accurate estimate. If various estimates from various coverage observations are used to obtain the monthly mean area-averaged rain rate, averaging with the use of weighting factors is preferable as was anticipated. Using this algorithm, the estimated error is about 20%. Estimated root-mean-square errors are greater than sampling errors by 3%–4% in all areas. That is, these additional 3%–4% differences are the contribution of the threshold method. It may also be that 3%–4% is just the lower bound of the additional error expected if the method is applied to TRMM observations. This result suggests that even if the optimal threshold and the best value for β are used, a 3%–4% error will exist because of regression. However, realistic errors due to partial observations are considered in this result.

Here the distribution of coverage (Fig. 8) should be considered once again. The *Y* axis of this graph is the number of TRMM visits to region B during 43 months. In a month, the area will be covered more than 90% about 41 times and 80% about 28 times. Observations in which only less than 80% will be covered number about 32 times. Therefore, most of region B will be observed during two-thirds of the TRMM visits. This is the TRMM visit feature near the turning points of the orbit (±35°). At lower latitudes, a more uniform distribution is expected for each coverage percentage. Therefore, the effects of small coverages are not as large at latitudes corresponding to Japan. If areas are fixed to the same size as used in this study, errors coming from small coverage observations will be unavoidable. To understand this problem, comparisons with a simulation using low-latitude sampling characteristics will be needed.

Another realistic error in an application of the threshold method involves an error in estimates of the β coefficient. In such a case, estimated area-averaged rain rates will include bias errors. However, if careful calibrations of β were carried out using climatological information, or some preprocessing, this kind of error can be recovered, because β is quite stable climatologically at least around the southern part of Japan as shown in this paper. Another important realistic error will be, as investigated by Krajewski et al. (1992), from the high uncertainty in the *Z*–*R* relationship. Their findings about the selection of a low threshold level for good performance should be considered in an actual application of this method.

6. Summary

The threshold method to estimate area-averaged rain rates $\langle R \rangle$ from the fractional area (F) of rain rates exceeding a preset threshold (τ) was tested using digital radar–AMeDAS precipitation data in the Japan area. Those are radar-based rain data with large spatial and long temporal coverage.

The F–$\langle R \rangle$ relationship of radar–AMeDAS precipitation analysis was examined systematically. A close relationship between averaged rain rate $\langle R \rangle$ and $F(\tau)$ was found. The threshold method works well in Japan throughout the seasonal cycle in several distinct climate regimes. The optimum threshold that maximizes the correlation between area-average rain rate and the fractional coverage of rain rates exceeding that thresh-

old is 3.5 mm h^{-1}, on the average, with a variation from 1.5 to 3.5 mm h^{-1} across the seasons. However, any threshold from 1.5 to 4.5 mm h^{-1} gives a very high correlation. Considering the variation of slope coefficients of the linear relationship, the variability is smallest when 3.5 mm h^{-1} is used as the threshold. As a conclusion 3.5 mm h^{-1} is recommended as the optimal threshold for radar–AMeDAS in every year and season.

Also, optimal thresholds were determined theoretically assuming the lognormal distribution. The theoretically derived optimal threshold level, determined from the probability distribution of rain rate, agrees with the empirically determined optimal threshold level. That is, the probability distribution of rain rate, which can be found from rain gauge observations, can be used for deriving optimal thresholds.

Using the threshold method with the coefficients obtained when the threshold is set to 3.5 mm h^{-1}, TRMM sampling of radar–AMeDAS rainfall was simulated. In the simulation of TRMM observations and estimation of monthly, area-averaged rain rates over 5° by 5° boxes, the errors incurred by the threshold method were only 4% larger than those assuming perfect observations. Estimated root-mean-square errors arising from the threshold method are much smaller than the sampling error (14%–19% on average). As expected, in the use of various computational schemes for incorporating the observational coverage by each overpass during a month to estimate the area-averaged monthly mean rain rate, actual weighting factors are needed to realize the most accurate estimates. Considering the limited dynamic range of satellite sensors, the threshold method provides an effective means for estimating space–time-averaged rainfall rates.

Acknowledgments. The authors would like to thank Prof. Kunio Shimizu of Tokyo Science University, who gave helpful comments regarding the threshold method. The comments of three anonymous reviewers were very helpful to improve the original version. We also thank Dr. Y. Makihara and Dr. T. Nakazawa (Meteorological Research Institute) for providing the radar–AMeDAS data, and Dr. Utashima (National Space Development Agency of Japan) for his program to compute the TRMM orbits. R. O. wishes to thank Dr. W. K.-M. Lau and Mr. O. W. Thiele of NASA/GSFC for sponsoring her visit to GSFC.

REFERENCES

Aitchison, J., and J. A. C. Brown, 1963: *The Lognormal Distribution.* Cambridge University Press, 176 pp.

Atlas, D., D. Rosenfeld, and D. A. Short, 1990: The estimation of convective rainfall by area integrals. 1: The theoretical and empirical basis. *J. Geophys. Res.,* **95,** 2153–2160.

Bell, T. L., 1987: A space-time stochastic model of rainfall for satellite remote-sensing studies. *J. Geophys. Res.,* **92,** 9631–9644.

——, 1990: Sampling errors for satellite-derived tropical rainfall: Monte Carlo study using a space-time stochastic model. *J. Geophys. Res.,* **95,** 2195–2205.

Braud, I., J. D. Creutin, and C. Barancourt, 1993: The relation between the mean areal rainfall and the fractional area where it rains above a given threshold. *J. Appl. Meteor.,* **32,** 193–202.

——, P. Crochet, and J. D. Creutin, 1994: A method for estimating mean areal rainfall using moving trend functions of the intensities. *J. Appl. Meteor.,* **33,** 1551–1561.

Chiu, L. S., 1988: Estimating areal rainfall from rain area. *Tropical Rainfall Measurements,* J. S. Theon and N. Fugono, Eds., A. Deepak Publishing, 361–367.

Crow, E. L., and K. Shimizu, Eds., 1988: *Lognormal Distributions: Theory and Applications.* Marcel Dekker, 387 pp.

Doneaud, A. A., P. L. Smith, S. A. Dennis, and S. Sengupta, 1981: A simple method for estimating convective rain over an area. *Water. Resour. Res.,* **17,** 1676–1682.

——, S. I. Niscov, D. L. Priegnitz, and P. L. Smith, 1984: The area-time integral as an indicator for convective rain volumes. *J. Climate Appl. Meteor.,* **23,** 555–561.

Kedem, B., and H. Pavlopoulos, 1991: On the threshold method for rainfall estimation: Choosing the optimal threshold level. *J. Amer. Stat. Assoc.,* **86,** 626–633.

——, and D. A. Short, 1993: On a curious linear relationship between rainfall averages. *Resenhas IME-USP,* **1,** 217–232.

——, L. S. Chiu, and Z. Karni, 1990: An analysis of the threshold method for measuring area-averaged rainfall. *J. Appl. Meteor.,* **29,** 3–20.

Kitabatake, N., and M. Obayashi, 1991: A study comparing radar–AMeDAS precipitation chart with observed data of rain gauge network of Tokyo metropolis (in Japanese). *J. Meteor. Res.,* **43,** 285–310.

Krajewski, W. F., M. L. Morrissey, J. A. Smith, and D. T. Rexroth, 1992: The accuracy of the area-threshold method: A model-based simulation study. *J. Appl. Meteor.,* **31,** 1396–1406.

Makihara, Y., 1993: The accuracy of radar–AMeDAS composite (in Japanese). *Kisyou,* **37,** 4–8.

——, 1996: A method for improving radar estimates of precipitation by comparing data from radars and raingauges. *J. Meteor. Soc. Japan,* **74,** 459–480.

——, N. Kitabatake, and M. Obayashi, 1995: Recent developments in algorithms for the JMA nowcasting system. Part I: Radar echo composition and radar–AMeDAS precipitation analysis. *Geophys. Mag. Series 2,* **1,** 171–204.

——, N. Uekiyo, A. Tabata, and Y. Abe, 1996: Accuracy of Radar-AMeDAS precipitation. *IEICE Trans. Trans. Commun.,* **E79-B,** 751–762.

Meneghini, R., and J. A. Jones, 1993: An approach to estimate the areal rain-rate distribution from spaceborne radar by the use of multiple thresholds. *J. Appl. Meteor.,* **32,** 386–398.

Morrissey, M. L., 1994: The effect of data resolution on the area threshold method. *J. Appl. Meteor.,* **33,** 1263–1270.

——, W. F. Krajewski, and M. J. McPhaden, 1994: Estimating rainfall in the Tropics using the fractional time raining. *J. Appl. Meteor.,* **33,** 387–393.

Obayashi, M., 1991: Data processing in the system for very short-range forecast of precipitation (in Japanese). *Sokko-Jihou,* **58** (3), 189–207.

Okamoto, K., T. Kozu, K. Nakamura, and T. Ihara, 1988: Tropical rainfall measuring mission rain radar. *Tropical Rainfall Measurements,* J. S. Theon and N. Fugono, Eds., A. Deepak Publishing, 213–219.

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

Rosenfeld, D., D. Atlas, and D. A. Short, 1990: The estimation of convective rainfall by area integrals. 2: The height–area rainfall threshold (HART) method. *J. Geophys. Res.,* **95,** 2161–2176.

Shimizu, K., and K. Kayano, 1994: A further study of the estimation

of area rain-rate moments by the threshold method. *J. Meteor. Soc. Japan,* **72,** 833–839.

——, and D. A. Short, and B. Kedem, 1993: Single- and double-threshold methods for estimating the variance of area rain rate. *J. Meteor. Soc. Japan,* **71,** 673–683.

Short, D. A., K. Shimizu, and B. Kedem, 1993a: Optimal thresholds for the estimation of area rain-rate moments by the threshold method. *J. Appl. Meteor.,* **32,** 182–192.

——, D. B. Wolff, D. Rosenfeld, and D. Atlas, 1993b: A study of the threshold method utilizing raingage data. *J. Appl. Meteor.,* **32,** 1379–1387.

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

Takase, K., and coauthors, 1988: Operational precipitation observation system in the Japan Meteorological Agency. *Tropical Rainfall Measurements,* J. S. Theon and N. Fugono, Eds., A. Deepak Publishing, 407–413.

Takemura, Y., K. Takase, and Y. Makihara, 1984: Operational precipitation observation system using digitized radar and rain gauges. *Proc. Nowcasting-II Symp.,* Norrkoing, Sweden, ESA SP-208, 411–416.

Yamamoto, A., 1991: Verification of monthly precipitation estimated by JMA's precipitation observation system using digitized radar and rain gauges(AMeDAS) (in Japanese). *J. Meteor. Res.,* **43,** 311–322.

Yamasaki, H., and T. T. Wilheit, 1990: The electrically scanning microwave radiometer and the special sensor microwave/imager (in Japanese). *Rev. Communications Res. Laboratory*, **36,** 83–92.