Proceedings of ERAD (2002): 255–260 © Copernicus GmbH 2002

ERAD 2002

Radar rainfall estimation: lessons learned from the NASA/TRMM validation program

E. Amitai[1, 2], **D. B. Wolff**[1, 3], **D. A. Marks**[1, 4], **and D. S. Silberstein**[1, 4]

[1]NASA Goddard Space Flight Center, 912.1, Greenbelt, Maryland, 20771, USA
[2]University of Maryland Baltimore County, Joint Center for Earth Systems Technology
[3]Science Systems and Applications, Inc.
[4]George Mason University, Center for Earth Observing and Space Research

Abstract. Evaluation of the Tropical Rainfall Measuring Mission (TRMM) satellite observations is conducted through a comprehensive Ground Validation (GV) program. Standardized instantaneous and monthly rainfall products are routinely generated using quality-controlled ground based radar data from four primary validation sites. As part of the TRMM GV program, effort is being made to evaluate these GV products and to determine the uncertainties of the rainfall estimates. This paper describes: (1) the GV climatological processing and product development effort and the products available to the public; (2) the product evaluation efforts for the Melbourne, Florida site, which are based on comparisons of radar to rain gauge data using several techniques; and (3) the limiting factors in the evaluation process that reflect current limitations in radar rainfall estimation. Lessons learned and possible improvements from this five-year mission are summarized.

1 Introduction

The Tropical Rainfall Measuring Mission (TRMM) Ground Validation (GV) program is an ongoing effort to provide reliable rainfall estimates that can be used to help validate the various TRMM satellite rainfall products over key locations around the globe. The focus for the NASA TRMM Office is to provide the best radar estimates possible from four primary sites: Melbourne, Florida (WSR-88D); Kwajalein Atoll, Republic of the Marshall Islands (WSR-93D-S band); Houston, Texas (WSR-88D); and Darwin, Australia (C-band polarimetric). In addition to routinely providing standardized instantaneous and monthly rainfall products, effort is being made to evaluate these GV products and to determine the uncertainties of the rainfall estimates. This paper provides an overview of the GV program, including a description of current algorithms and data availability. It presents the product evaluation efforts for the Melbourne, Florida site and its limitations, and summarizes the lessons learned and possible improvements.

Correspondence to: E. Amitai (eyal@radar.gsfc.nasa.gov)

2 TRMM ground validation data processing, algorithms and products

The GV data processing and climatological product development system is principally automated; however, the quality control (QC) of the reflectivity data does require some manual tuning when anomalous propagation or other anomalous non-meteorological echoes are present. The QC algorithm removes non-meteorological radar echoes using a combination of eight adjustable height and reflectivity thresholds. The GV system is described more fully in Marks et al. (2000) – a summary of a keynote presentation made in the First ERAD conference, which also provides a data flow chart of the system. The radar estimates are based on using a power law $Z_e = AR^b$ with a fixed exponent b of 1.4. The A coefficient is tuned to a network of QC-ed gauges in such a way that the total monthly rainfall, as estimated from the radar Z_e pixels of 2×2 km^2 above the gauges, is matched to the combined gauge accumulations (ΣG).

$$A = [\Sigma(Z_e^{1/1.4})/\Sigma G]^{1.4} \quad (1)$$

An automated quality control (AQC) procedure to filter unreliable rain gauge data upon comparison to radar data is used prior to generation of the $Z_e - R$ relations, as described in Amitai (2000). Separate $Z_e - R$ relations are derived for different range intervals of the radar. Interpolated radar reflectivities within 99 km of the radar are taken from a 1.5-km CAPPI, while reflectivities from 100–150 km are taken from a 3-km CAPPI. Separate relations are also generated for convective and stratiform rain as defined by the Steiner et al. (1995) classification scheme. The derived relations are then applied to QC-ed $Z_e s$ to obtain instantaneous surface rain rate (R) maps. The latter are then integrated to produce monthly products. Climatological GV products, statistics, and descriptions can be found at: http://trmm-fc.gsfc.nasa.gov/trmm_gv/index.html.

This bulk adjustment scheme, in general, was implemented following the direction of the pre-launch TRMM GV Science Team. However, modifications are taking place from time to time. For example, in previous versions the $Z_e - R$ relations were derived from radar and gauge data within 99 km range only, but applied to $Z_e s$ from the lowest beam to create rainfall products up to 150 km range. In the proposed operational version the $Z_e s$ for generating the $Z_e - R$ relations are taken from the same CAPPI level to which the relations are then applied. There are many problems associated with the bulk adjustment procedure and current investigation is underway to determine if a more effective approach will improve the quality of the GV products. The first such test is the implementation of the Window Probability Matching Method (WPMM, Rosenfeld et al., 1994). The improvement due to rain type classification is also being investigated.

3 Product evaluation effort for the Melbourne, Florida site

3.1 The Melbourne, Florida GV site: Radar and rain gauge data

Central Florida is one of four primary GV sites selected for TRMM. The site is centered around the NEXRAD Weather Surveillance Radar-1988 Doppler (WSR-88D) located in Melbourne. The rain gauge network consists of approximately 100 tipping buckets at distances less than 150 km from the radar. Additionally, a dense rain gauge network (DRGN) of 20 tipping buckets within an approximately 10 km^2 area was installed 35 km west of the radar to provide validation of the radar estimates during August–September 1998. The DRGN is not used for generating $Z_e - R$ relations, and therefore, can be viewed as an “independent” validation network.

3.2 Evaluation of monthly rainfall estimates

Figure 1 shows an example of the differences between the radar estimates over the gauge and the gauge accumulations at times when the radar is operating, using the independent gauge network. In general, the normalized mean absolute error (MAE) on a monthly time scale was found to be below 20%. The difference between the point gauge measurement and the true averaged rain within the radar pixel, caused by the natural variability of rain and gauge instrumental errors, might explain a major fraction of the MAE. Figure 1 includes several gauges located within the same radar pixel of 2 × 2 km^2. These gauges are marked by rectangles. The difference in gauge accumulations within each group suggests that MAE represents just an upper limit to the accuracy of the radar estimates. The accuracy may be higher but a denser gauge network is required for verification.

An alternative method, which also ensures unbiased total radar rainfall accumulation above the gauges, is the WPMM. This method is based on matching unconditional probabilities of R measured by the gauges, to probabilities of Z_e observed above the gauges. In Table 1, monthly comparisons between the performance of the WPMM- and the bulk adjustment power law-based $Z_e - R$ relations (with and without rain type classification) are presented using all the dependent gauges within 150 km. The radar accumulations both in Fig. 1 and in Table 1, were computed using a maximum 10-minute integration period, and gauge data has been filtered for gaps greater than 10 minutes. Therefore, the R/G ratio in Table 1 is very close to 1.0 as expected with dependent gauges. Note that according to the bulk adjustment scheme for $Z_e - R$ development, gauge rates are taken only at times of radar scan and the accumulations are simply the summation of the gauge and radar rates regardless of the time between the radar scans. Upon rain type classification, Rs measured by the gauge at times when no Z_e was observed above (due to data mismatching) are not used for $Z_e - R$ development. Thus, radar underestimates the rain as expressed by the relatively low R/G values. On a month-by-month basis, very little variation in all three statistics is evident between the different methods. It is, however, worth mentioning that for a fixed $b = 1.4$, A was found, in general, to be larger for stratiform than for convective rain. Allowing A and b to vary, a minimum MAE value was also associated with a larger A-value during stratiform rain for almost all months analyzed. However, a systematic connection between the type of rain and the value of b has not been established. WPMM-based $Z_e - R$ relations, on the other hand, demonstrate systematic variation between convective and stratiform rain (Amitai, 2000).

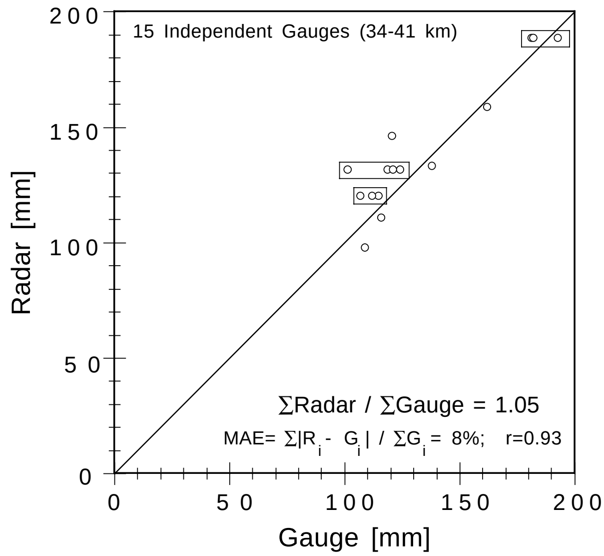

Fig. 1. The relations between the gauge integrated rain rates and the radar integrated rain rates over the gauges of the Florida independent network (DRGN) during August 1998. Range dependent power law $Z_e - R$ relations adjusted to 79 QC-ed gauges were applied.

Table 1. Comparison of area averaged rain rate (R), radar-to-gauge bias (R/G), and normalized mean absolute error (MAE) between the monthly radar estimates and the gauge accumulations based on bulk-adjustment (with and without rain type classification) and WPMM for 5 months at Melbourne, Florida

Month	Bulk-Adj. With Rain Classification			Bulk-Adj. No Classification			WPMM No Classification		
	R (mm/day)	R/G	MAE	R (mm/day)	R/G	MAE	R (mm/day)	R/G	MAE
07-1998	3.81	0.99	0.13	3.84	1.01	0.13	3.95	1.00	0.15
08-1998	4.16	0.97	0.17	4.28	0.99	0.16	4.10	0.99	0.16
09-1998	6.16	0.95	0.15	6.34	0.98	0.15	6.39	0.98	0.15
08-1999	5.05	0.98	0.16	5.14	1.00	0.16	5.08	1.00	0.17
09-1999	7.34	0.98	0.14	7.48	1.00	0.14	7.52	0.99	0.14

3.3 Evaluation of instantaneous rainfall estimates

While monthly estimates are relatively simple to evaluate, the evaluation of the instantaneous products is much more of a challenge. At the small scale there are many sources of uncertainties. Scatter plots of point comparisons between radar and rain gauges are extremely noisy for several reasons (e.g. sample volume discrepancies, timing and navigation mismatches, variability of $Z-R$ relations, quality of "instantaneous" gauge data), and are therefore of little use for evaluating the estimates. An alternative method based on the analysis of the distribution of R as derived from gauge intensities and from reflectivities above the gauge network is used.

In Fig. 2a and b probability density functions (PDFs) and cumulative distribution functions (CDFs) of the radar estimated Rs above the gauges are generated for different gauge adjusted power laws. Several b-values are tested, while A in each trial is tuned to the monthly gauge accumulations. All QC-ed gauges within 99 km from the radar site are used (63 gauges). The shape of the distribution is found to be very sensitive to b and A in the adjusted $Z_e = AR^b$. Note that the official TRMM GV system does not allow for a variable b-value.

Due to the spatial variability of R, the distribution of the gauge rates does not necessarily represent the true distribution of R at the scale of a radar pixel. CDFs derived from the intensities measured by the same 63 gauges are compared for different time averaged R, simulating different domain sizes (Fig. 2c). As the space-time domain increases, the PDF trends toward lower R-values, thus overestimating probabilities of weak intensities while probabilities of strong intensities are underestimated. The spatial smoothing of the radar beam over an area of S km^2 is analogous to the smoothing of the rain gauge measurements in time T according to $T = 1.3S^{1/2}/V$, where V (km h^{-1}) is the horizontal velocity of the rainfall system (Zawadzki, 1975). For $S = 4$ km^2 and V of 20–25 km h^{-1}, the smoothing of the gauge rates at $T \approx 7$ minutes simulates the radar pixel size. For the given month investigated here, $b = 1.3$ seems roughly to correspond to a 7-minute averaged rain rate. The PDF curve based on a 7-minute averaged gauge rain rate (dashed line in Fig. 2a) lies between the two PDF curves based on $b = 1.2$ and $b = 1.4$, for most of the dynamic range of R.

Comparing panel (b) with panel (c) reveals that the changes in the PDF as a result of the smoothing are small compared to those caused by the change of b and A. The PDFs are based on a dynamic range of b-values that is commonly used, and on a range of time intervals that is expected to include the rain rate distribution function that represents the area beneath the radar-sampled volume (2×2 km^2). While the average horizontal velocity may vary, the averaging times range used here is in accordance with the optimal range of ΔT found by Habib and Krajewski (2002) for the best correlation between the instantaneous radar estimates and the gauge observations for the same data set. While $b = 1.3$, as mentioned previously, seems roughly to correspond to a 7-minute averaged R, $b = 1.4$ already overestimates the probabilities of the weak intensities and underestimates the probabilities of the strong intensities, relative to any time averaging gauge-based PDF tested in this study; i.e., 1–15 minutes. However, among the gauge adjusted power laws which ensure unbiased total accumulations, $b = 1.5$ yields the minimum MAE between the monthly radar estimates and the gauge accumulations. The minimum MAE, 17.6%, increases only to 17.8% and 18.2% for $b = 1.4$ and $b = 1.3$, respectively (Fig. 3). While adjusted power laws may yield satisfactory estimates of rainfall accumulations, for a wide range of b-values, as seen in Fig. 3, the rain rates might be systematically biased, as seen in Fig. 2a. An exponent value $b = 1.5$, which was selected objectively using the conventional optimization criterion, yields a PDF that is far from representing any of the gauge-based PDFs. Note that the discrepancies in the MAE value presented in Fig. 3 for $b = 1.4$ and those shown in Table 1 are due to different integration schemes and range from the radar being used.

The PDF's shape is found to be less sensitive to different gauge time averaged Rs than to different gauge adjusted b- and A-values (as seen by comparing Fig. 2b with Fig. 2c). This means that the uncertainties in the derived PDFs are likely to be smaller if using the gauge rates compared to those based on applying an adjusted power law. The above results

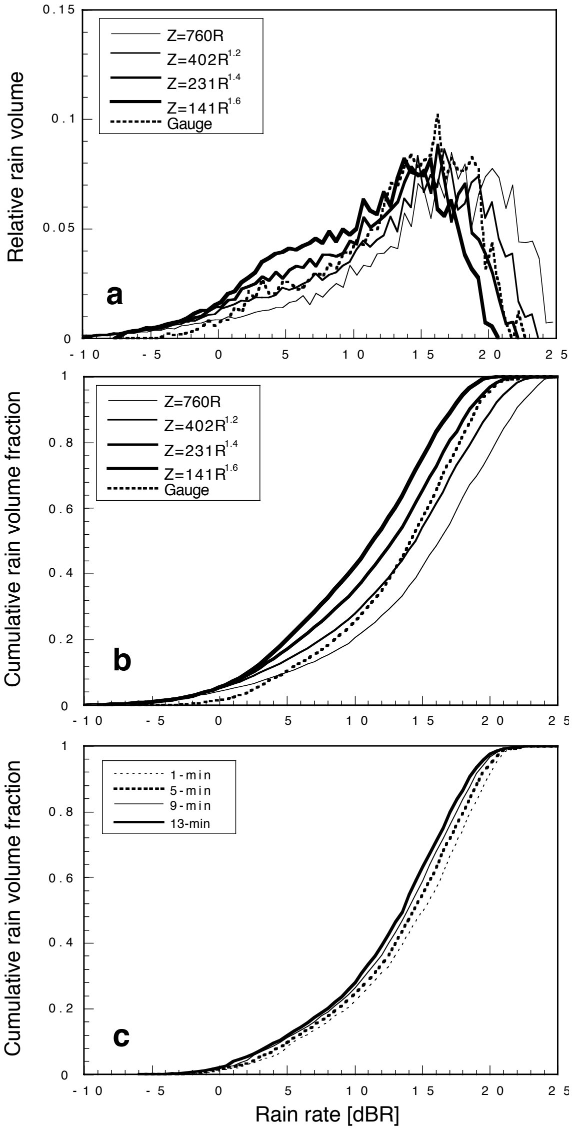

Fig. 2. Distribution of rain volume by rain rate for the Melbourne, Florida area during August 1998. (a) The PDFs are derived from 2 ×2 km^2 reflectivity pixels above 63 QC-ed gauges for different gauge adjusted power laws (solid curves), and from 7-minute averaged gauge rates measured by the same 63 gauges centered at the radar scan time (broken curve). The A coefficient in $Z_e = AR^b$ is tuned to the gauges in such a way that the total monthly rainfall, as estimated from $Z_e s$ above the gauges, is matched to the accumulation from all gauges combined (6246 mm). (b) The CDFs for the same data set. (c) The CDFs as derived from the 63 gauge intensities for different time averaging. Note, dBR = 10 log(R/ 1 mm h^{-1}) — [R]=[mm h^{-1}].

suggest the advantage of the WPMM over the power law relations when R distributions are of utmost importance. The WPMM method of relating Z_e to R forces the distribution of R, as derived from Z_e above the gauge, to be identical to the distribution of the gauge rates. Thus, any PDF of R based on gauge data can be viewed as a PDF of R derived from $Z_e s$ above the gauge upon application of the WPMM.

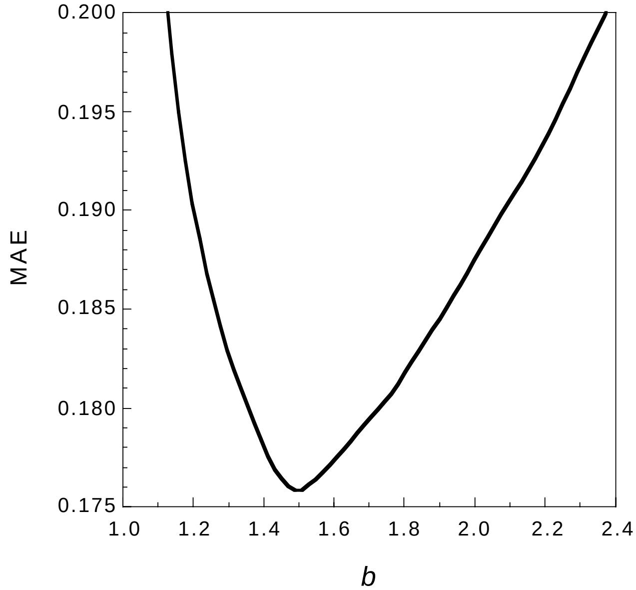

Fig. 3. Sensitivity of radar-gauge mean absolute error (MAE) to different gauge adjusted power laws $Z_e = AR^b$ as represented here by their exponent b-value. The A-value is chosen such that the total accumulation by radar is equal to the total accumulation by gauges (i.e., $\Sigma R_i = \Sigma G_i$, where R_i and G_i are the monthly radar and gauge accumulations at gauge i, respectively). The MAE is defined as $MAE = \Sigma|R_i - G_i|/\Sigma G_i$. The sensitivity test is performed using the Melbourne, Florida, August 1998 data set (63 gauges).

3.4 Limiting factors in the evaluation process

Rain quantities derived from different instruments representing different time and space domains often differ significantly. For example, the sampling area of a typical radar-rainfall product (2 × 2 km^2) and that of a rain gauge differ by eight orders of magnitude. The radar estimates are based on instantaneous observations averaged over an area of a certain size, while the gauge data represent accumulation over a defined time period. As the latter serve often as the “ground truth” to evaluate the radar-based estimates, one has to make sure that the different domain scales do not artificially introduce the differences between the estimates. For example, if rainfall itself varies significantly at scales smaller than that of a radar pixel, clearly a single gauge cannot represent such variability well, and this distorts our evaluation of the radar performance.

Several methodologies for evaluating the accuracy of TRMM GV *monthly* and *instantaneous* products are used. However, the limiting factor in the evaluation process of both types of products is found to be the lack of the existence of a very dense gauge network. Using the DRGN independent data set, it was found on a *monthly* time scale that the average difference between the radar estimates over the gauge and the gauge accumulation is of the same order as the differences between point gauge measurements taken from the same radar pixel. The evaluation of *instantaneous* products was based on comparing the distribution of rain rates as derived from the radar estimates above the gauges to those derived from the gauge intensities. However, due to the spatial variability of R, the distribution of the gauge rates does not necessarily represent the true distribution of R at the scale of a radar pixel.

How well do the gauge-based PDFs of R, which are used as "ground truth" to evaluate and even generate the radar-based PDFs, represent the actual R distribution at the scale of a radar pixel? The difficulty in addressing this issue is that sufficiently dense gauge networks necessary to represent the distribution of R at a radar pixel size are not available at TRMM GV sites. The best the existing methodologies can do, based on the data collected during the TRMM field campaigns, is estimation of the overall radar-rainfall error variance using, for example, the technique proposed by Ciach and Krajewski (1999). It is also possible to simulate an increase in domain size by smoothing the gauge rates in time according to a space-time relation. This approach was demonstrated here to examine the effect of domain size on the gauges-based PDFs. However, these do not provide the full distribution of errors.

Better estimation of the true rain rate distribution at the scale of a radar pixel is very important since: (1) it permits better *evaluation of radar rain rate estimates*. This methodology of using PDFs is particularly important when the rain rates are of concern rather than just the total rain accumulations, or, when the integral properties of rain rates suffer from large uncertainties like those from space-borne radars (e.g., TRMM) with revisit time of the order of hours or days; and (2) it is critical for improved $Z_e - R$ relations. The TRMM GV program is now testing new non-parametric algorithms, in which the $Z_e - R$ relations are based on matching the distribution of the reflectivities to the distribution of the gauge rates, assuming the latter represents the true distribution at the scale of a radar pixel. On the other hand, the distribution of the radar derived Rs based on gauge-adjustable power laws, the method currently used in TRMM GV program, is found to be very sensitive to the $Z_e - R$ relation being applied.

Further investigation is needed to determine how to convert a point PDF to a spatial PDF. Denser gauge networks for verification accuracy of the derived PDFs as well as for direct validation of R estimates should be used. These should have the right spatial configuration to minimize the spatial sampling random error.

4 Lessons learned

Although the Melbourne GV site has much to offer toward reliable GV measurements, including stable radar calibrations, expansive gauge networks, and significant precipitation, there remain significant problems. Many of these problems could be avoided in the future with careful planning. The following provides an abbreviated list of lessons learned in the Melbourne GV program.

1. Simply processing data and generating products is not sufficient. The products must be evaluated and systematic errors and uncertainties must be established.

2. Independent validation via high quality, supplemental, gauge networks is a must. Validation gauges should be scattered throughout the radar domain at several different ranges. High-accuracy reference measurements based on a very dense gauge network, which can provide sufficient gauge sampling to effectively match the radar sampling volume, are highly desirable. This will allow

a). improved evaluation of (TRMM GV) radar rain rate estimates and rainfall depths, and the testing of alternative algorithms for radar rainfall estimation;

b). improved estimation of the true rain rate distribution at the scale of a radar pixel and at different space-time scales which is critical for development of new algorithms for remote sensing rainfall quantities;

c). improved understanding of the natural variability of rain, which is a limiting factor in evaluating radar rainfall estimates at small space-time domains.

To accomplish these objectives, the TRMM Validation Office is planning to have a very high dense gauge network of a typical radar pixel size within the current DRGN site in Florida.

3. Routine radar and gauge calibration is a must. Comparison to TRMM (or Global Precipitation Measurement [GPM] Mission) radar reflectivities provides the best approach.

4. Co-location (2–3) of rain gauges, and redundancy in gauge loggers is a must. Problems associated with tipping bucket gauge measurements are not due to calibration alone, but are also related to logger "dropouts" during which a logger will temporarily (or permanently) cease to function.

5. When using any gauge adjustment technique for radar rainfall estimation, independent QC of the radar and gauge data alone is not sufficient. Although co-located rain gauges will improve reliability, proper QC of rain gauge data upon comparison with radar data is essential as it is a key factor in R/G bias results.

6. Monthly R/G bias and error estimate statistics are mostly insensitive to the gauge adjustment method used to create the $Z_e - R$ relations (i.e., power law or WPMM) or whether the data were subject to rain type classification. Any QC-ed gauge adjustable $Z_e - R$ relations may yield satisfactory estimates of rainfall accumulations over large space-time domain.

7. The shape of the distribution of radar derived Rs is found to be quite sensitive to A and b in the adjusted $Z_e = AR^b$. From PDF analysis, an optimal b (and A) can be inferred. However, the optimal A and b are strongly dependent on the definition of the term optimal: Whether similarity to the gauge-based PDF is desired, or whether minimizing the radar-gauge rain accumulations over relatively longer periods is desired.

8. The shape of the gauge-based PDF is found to be less sensitive to time/space variations than the radar-based PDF to changes of the adjusted power law $Z_e - R$ parameters. This suggests the possible advantage of using the WPMM for estimation of R and R distributions. This method of relating Z_e to R forces the distribution of R, as derived from Z_e above

the gauges, to be identical to the distribution of the gauge rates. In contrast, an adjusted power law might have A and b values which will generate a PDF that diverges from the distribution of the gauge rates and from the true R distribution.

Acknowledgement. The effort of the first author is supported by NASA Grant NAG5-7909. This work was done as part of the TRMM validation program directed by Richard Lawrence of the NASA Goddard Space Flight Center.

References

Amitai, E., Systematic variation of observed radar reflectivity-rainfall rate relations in the tropics, J. Appl. Meteor. (Special Issue on TRMM), 39, 2198-2208, 2000.

Ciach, G.J., and Krajewski, W.F., On the estimation of radar rainfall error variance, Adv. Water Resour., 22(6), 585-595, 1999.

Habib, E., and Krajewski, W. F., Uncertainty analysis of the TRMM ground validation radar-rainfall products: application to the TEFLUN-B field campaign, J. Appl. Meteor., 41, 558-572, 2002.

Marks, D.A., Kulie, M.S., Robinson, M., Silberstein, D., Wolff, D.B., Ferrier, B.S., Amitai, E., Fisher, B., Wang, J., Augustine, D., and Thiele, O., Climatological processing and product development for the TRMM Ground Validation Program, Phys. Chem. Earth (B), 25, 871-875, 2000.

Rosenfeld, D., Wolff, D.B., and Amitai, E., The window probability matching method for rainfall measurements with radar, J. Appl. Meteor., 33, 682-693, 1994.

Steiner, M., Houze, Jr.R.A., and Yuter, S.E., Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data, J. Appl. Meteor., 34, 1978-2007, 1995.

Zawadzki, I., On radar-raingage comparison, J. Appl. Meteor., 14, 1430-1436, 1975.