Evaluation of Radar Rainfall Products: Lessons Learned from the NASA TRMM Validation Program in Florida

EYAL AMITAI AND DAVID A. MARKS

Science and Exploration Directorate, NASA Goddard Space Flight Center, Greenbelt, Maryland, and Center for Earth Observing and Space Research, George Mason University, Fairfax, Virginia

DAVID B. WOLFF

Science and Exploration Directorate, NASA Goddard Space Flight Center, Greenbelt, and Science Systems and Applications, Inc., Lanham, Maryland

DAVID S. SILBERSTEIN

Science and Exploration Directorate, NASA Goddard Space Flight Center, Greenbelt, Maryland, and Center for Earth Observing and Space Research, George Mason University, Fairfax, Virginia

BRAD L. FISHER

Science and Exploration Directorate, NASA Goddard Space Flight Center, Greenbelt, and Science Systems and Applications, Inc., Lanham, Maryland

JASON L. PIPPITT

Science and Exploration Directorate, NASA Goddard Space Flight Center, Greenbelt, Maryland, and Center for Earth Observing and Space Research, George Mason University, Fairfax, Virginia

(Manuscript received 21 February 2005, in final form 12 December 2005)

ABSTRACT

Evaluation of the Tropical Rainfall Measuring Mission (TRMM) satellite observations is conducted through a comprehensive ground validation (GV) program. Since the launch of TRMM in late 1997, standardized instantaneous and monthly rainfall products are routinely generated using quality-controlled ground-based radar data adjusted to the gauge accumulations from four primary sites. As part of the NASA TRMM GV program, effort is being made to evaluate these GV products. This paper describes the product evaluation effort for the Melbourne, Florida, site. This effort allows us to evaluate the radar rainfall estimates, to improve the algorithms in order to develop better GV products for comparison with the satellite products, and to recognize the major limiting factors in evaluating the estimates that reflect current limitations in radar rainfall estimation. Lessons learned and suggested improvements from this 8-yr mission are summarized in the context of improving planning for future precipitation missions, for example, the Global Precipitation Measurement (GPM).

1. Introduction

Data assimilation in forecast numerical models and many hydrologic applications requires satellite observations of precipitation. However, providing values of precipitation is not sufficient unless they are accompanied by their associated uncertainty estimates. Having satellite validation datasets with defined uncertainty levels is crucial to providing better rainfall estimates from space. The evaluation of the National Aeronautics and Space Administration/Japan Aerospace Exploration Agency (NASA/JAXA) Tropical Rainfall Measuring Mission (TRMM) satellite observations is conducted through a comprehensive NASA ground valida-

Corresponding author address: Dr. Eyal Amitai, NASA Goddard Space Flight Center, Code 613.1, Greenbelt, MD 20771.
E-mail: eyal@radar.gsfc.nasa.gov

© 2006 American Meteorological Society

tion (GV) program. Since the launch of TRMM in late 1997, standardized instantaneous and monthly rainfall products are routinely generated using quality-controlled (QC) ground-based radar data adjusted to gauge accumulations from four primary sites: Melbourne, Florida; Kwajalein Atoll, Republic of the Marshall Islands; Houston, Texas; and Darwin, Australia (Wolff et al. 2005). In addition to routinely providing the GV rainfall products to the scientific community and to the public, effort is being made to evaluate these GV products. This paper describes the effort of evaluating the GV products for the Florida site. In addition to "common use" comparisons between point gauge accumulations and radar estimates over the gauges, the uncertainties are evaluated 1) by utilizing a dense network of independent gauges in which more than one gauge exists within a radar pixel; 2) by conducting sensitivity tests of the distribution of the radar-derived rain rates to different gauge-adjusted observed reflectivity–rain rate (Z_e–R) relations; 3) by testing the sensitivity of the estimates to the algorithm used to select the quality gauges for Z_e–R generation; and 4) by testing the sensitivity of the estimates to the application of a single Z_e–R relation as compared to dual convective/stratiform Z_e–R relations. These procedures allow us to evaluate the estimates, to improve the algorithms to have better GV products for comparison with the satellite products, and to recognize the limiting factors in the evaluation process that reflect current limitations in radar rainfall estimation using conventional radar systems, that is, nonpolarimetric or dual-wavelength systems. Lessons learned and suggested improvements from this 8-yr mission are summarized.

2. Product evaluation effort for the Melbourne, Florida, site

a. Radar and rain gauge data, algorithms, and products

This site is centered around the Next-Generation Weather Radar (NEXRAD) Weather Surveillance Radar-1988 Doppler (WSR-88D) located on the eastern coast of central Florida in Melbourne (MELB). Anomalous propagation is common over land in the northwest and southwest quadrants and is more prevalent during the late evening and early morning hours. Clear-air echo is enhanced during dry periods and along density-driven boundaries (e.g., sea breeze, land breeze, fronts, drylines). The precipitation is dominated by sea-breeze-induced and airmass convection during warm months, with some frontal induced systems during cooler months (Wolff et al. 2005).

The gauge network consists of approximately 100 tipping-bucket gauges at distances of less than 150 km from the radar (the number can vary slightly from month to month). Additionally, during August–September 1998, a high-density rain gauge network (DRGN) of 20 tipping-bucket gauges within an approximately 30-km^2 region was operated 40 km west of the Melbourne radar to provide validation of the ground radar rainfall estimates. In the TRMM GV products, rain gauge data are used to adjust the Z_e–R relations. The removal of the mean bias of the radar rainfall estimates at the gauge site is referred to as a rain gauge adjustment. This DRGN network, unlike all other networks, is not used for generating Z_e–R relations and, therefore, can be viewed as an "independent" validation network.

The NASA TRMM GV data processing and climatological product development system is principally automated; however, the QC of the reflectivity data does require some manual tuning when anomalous propagation or other anomalous nonmeteorological echoes are present. The QC algorithm removes nonmeteorological radar echoes using a combination of eight adjustable parameters consisting of three echo height and five reflectivity thresholds (Kulie et al. 1999). This algorithm is based on modification of the technique developed by Rosenfeld et al. (1995). The QC of the gauge data is performed using the gauge software package (GSP) developed by the NASA TRMM Satellite Validation Office (TSVO), which applies several levels of automated quality control. The description of the GSP and how it treats single-, double-, and multiple-tip events is available at NASA (2003). Once the data have been properly quality controlled, the radar and gauge data are merged so that gauge-adjusted Z_e–R relations can be determined. Prior to the determination of the Z_e–R relations an automated quality control (AQC) procedure to filter unreliable rain gauge data upon comparison to radar data is used, as described in Amitai (2000). To mitigate range effects on the results, separate Z_e–R relations are derived for different range intervals of the radar (15–50, 50–98, and 98–150 km). Interpolated radar reflectivities within 98 km of the radar are taken from a 1.5-km constant altitude plan position indicator (CAPPI), while reflectivities from 98 to 150 km are taken from a 3-km CAPPI.

Following the direction of the prelaunch TRMM GV Science Team, gauge-adjusted power-law $Z_e = AR^b$ relations with a fixed b of 1.4 were implemented. The A coefficient is tuned to a network of quality-controlled gauges in such a way that the total monthly rainfall, as estimated from the radar Z_e pixels of 2 km $\times$ 2 km over

TABLE 1. Dependent monthly validation for each month of 1998 at Melbourne, FL. Shown are the radar area-averaged rain rate in mm day^{-1} ($\langle R \rangle$), radar-to-gauge ratio (R/G), normalized mean absolute difference (NMAD), and regression correlation coefficient (r) values between the monthly radar estimates over the gauges and the gauge accumulations, based on gauge-adjusted $Z_e = AR^{1.4}$ relations (with and without rain-type classification) and WPMM. The number of QC gauges (N) and their total accumulations (G) for each month are also included. The best performance statistics for each month are highlighted in bold (i.e., lowest $|R/G - 1|$, highest r, and lowest NMAD).

			Bulk-adjusted $Z_e = AR^{1.4}$ with rain classification				Bulk-adjusted $Z_e = AR^{1.4}$ with no classification				WPMM with no classification			
Month	N	G (mm)	$\langle R \rangle$ (mm day^{-1})	R/G	r	NMAD	$\langle R \rangle$ (mm day^{-1})	R/G	r	NMAD	$\langle R \rangle$ (mm day^{-1})	R/G	r	NMAD
Jan 1998	50	3563	3.60	**1.00**	0.82	0.17	3.63	1.01	0.83	0.16	3.74	**1.00**	**0.86**	**0.15**
Feb 1998	73	11 849	6.57	0.99	**0.79**	**0.11**	6.68	**1.00**	**0.79**	**0.11**	6.71	1.01	**0.79**	**0.11**
Mar 1998	67	5348	5.00	0.99	**0.82**	**0.17**	5.06	1.01	**0.82**	**0.17**	4.96	**1.00**	**0.82**	0.18
Apr 1998	56	2696	1.32	0.97	**0.96**	0.15	1.28	**0.99**	0.95	0.15	1.27	0.98	**0.96**	**0.14**
May 1998	60	2640	1.34	0.99	0.82	0.28	1.37	1.01	0.81	0.29	1.34	**1.00**	**0.89**	**0.23**
Jun 1998	61	2404	1.31	0.96	0.88	0.26	1.35	**0.99**	**0.89**	0.26	1.37	**0.99**	0.88	**0.25**
Jul 1998	61	7438	3.83	**1.00**	**0.92**	**0.13**	3.85	**1.00**	**0.92**	**0.13**	3.98	**1.00**	0.91	0.14
Aug 1998	78	9227	4.10	0.96	**0.86**	0.17	4.22	**0.98**	**0.86**	0.17	4.08	**0.98**	**0.86**	**0.16**
Sep 1998	76	13 716	6.48	0.97	0.80	**0.14**	6.64	**1.00**	0.80	**0.14**	6.58	0.99	**0.81**	**0.14**
Oct 1998	55	2966	2.09	0.98	0.81	0.29	2.14	1.01	0.81	0.30	2.03	**1.00**	**0.84**	**0.27**
Nov 1998	58	4800	3.50	0.98	0.93	0.13	3.62	0.99	0.92	0.14	3.49	**1.00**	**0.96**	**0.10**
Dec 1998	38	723	0.94	0.96	**0.83**	**0.21**	1.02	**1.00**	0.79	0.25	0.99	0.96	0.82	0.22

the gauges, is matched to the combined gauge accumulations (ΣG):

$$A = \left[\sum \left(Z_e^{1/b}\right) \Big/ \sum G\right]^b. \quad (1)$$

The 7-min-averaged gauge-measured rain rates centered on the time of each radar scan are taken to calculate the gauge accumulations (ΣG). The accumulations are simply the summation of the radar and gauge rates regardless of the time interval between radar scans. Separate relations are generated for convective and stratiform rain as defined by the Steiner et al. (1995) classification scheme.

Based on results from the evaluation process, some of which are presented here, the TSVO has switched to version number 5 (V5) in which the gauge adjustment scheme is based on the Window Probability Matching Method (WPMM; Rosenfeld et al. 1994) rather than a power law (V1–V4). As in the power-law adjustment procedure, WPMM ensures unbiased total radar rainfall accumulation over the gauges, but it is achieved by matching probabilities of R measured by the gauges to probabilities of Z_e observed over the gauges. The matching is performed unconditionally on rain. As in V4 only the reflectivity value over the gauge and the 7-min averaged gauge rain rate centered on the time of the radar scan are taken to construct the probability density functions (PDFs). Then, the R and Z_e having the same cumulative probability are matched. Separate Z_e–R relations are still derived for different distances (range intervals) of the radar, but not for different rain types.

The derived Z_e–R relations are then applied to quality-controlled Z_es from the same CAPPI level from which the relations were generated to obtain instantaneous surface rain-rate maps (TRMM Standard Product TSP 2A-53). The latter are then integrated to obtain monthly rainfall accumulation products (TRMM Standard Product TSP 3A-54). Beginning in the summer of 2003, V5 products are delivered to the NASA Goddard Earth Sciences Distributed Active Archive Center (GES/DAAC) for distribution. The GV system is described more fully in Marks et al. (2000) and Wolff et al. (2005). Climatological GV rainfall statistics, and descriptions can be found online at http://trmm-fc.gsfc.nasa.gov/trmm_gv/gv_products/GVproducts.html; GV rainfall products are available from http://daac.gsfc.nasa.gov/data/datapool/TRMM_DP/.

b. *Evaluation of monthly estimates*

1) DEPENDENT GAUGES

Table 1 presents monthly comparisons between the performance of the WPMM and the bulk adjustment power-law-based Z_e–R relations with a fixed $b = 1.4$ (with and without rain-type classification) using all dependent gauges within 150 km during 1998. The radar accumulations were computed using a maximum 10-min integration period, and gauge data have been filtered for gaps greater than 10 min. As noted previously, according to the bulk adjustment scheme for Z_e–R development, gauge rates are taken only at times of radar scan and the accumulations for both gauge and radar

are simply the summation of their rain rates regardless of the time interval between the radar scans. Therefore, the *R/G* ratio in Table 1 is very close to 1.0 as expected with dependent gauges. Note that the maximum 10-min integration period is a modification applied to V5 and to all three methods presented in Table 1. However, in the official V4 products, which are based on bulk adjustment with rain classification, the maximum integration time for the monthly products was 75 min and resulted with *R/G* ranging from 0.64 to 1.28. The rationale for reducing the integration time is to achieve the best estimates for the periods the radar is operating rather than to estimate the *monthly* rainfall. Based on parameters, such as the fraction of time the radar was operating during the month and the fraction of rain the gauges were measuring at times the radar was operating, modification can be applied to the "monthly" rainfall accumulation product TSP 3A-54 to estimate the *monthly* rain accumulation. The first type of modification (i.e., to divide each pixel in TSP 3A-54 by the fraction of time the radar operated during the month) is similar to the way the monthly products based on the TRMM satellite radar observations are generated, that is, the mean *R* obtained at times when the radar operates is applied to the entire month. The second type of modification matches the radar estimate over the gauges to the monthly gauge accumulations. It is up to the user to choose which type of modification to apply.

Upon rain-type classification, *R*s measured by the gauge at times when no Z_es were observed over the gauge are not used for Z_e–*R* development. This is due to the fact that the classification scheme is applied to the reflectivity dataset, while the classification of the gauge data is determined, not by analyzing the gauge rates, but solely by the matching procedure. As a result, the radar underestimates the rain as expressed by the relatively low *R/G* values (Table 1, fifth minor column). These cases were found to contribute up to 4% of the total rain amount and are assumed to be a result of wind sorting, data mismatch, and issues related to the gauge data interpolation scheme. By switching to V5 these cases are taken into account since the data are no longer classified to convective/stratiform rain type and all gauge data are used for Z_e–*R* development.

The normalized mean absolute difference (NMAD) is defined as

$$\mathrm{NMAD} = \sum |R_i - G_i| \Big/ \sum G_i, \qquad (2)$$

where R_i and G_i are the monthly radar and gauge accumulations at gauge *i*, respectively. The NMAD on a monthly time scale ranged between 0.10 and 0.27 with mean and median values of 0.17 and 0.16, respectively (for WPMM). In Table 1, the best performance statistics for each month are highlighted in bold (i.e., lowest $|R/G - 1|$, highest *r*, lowest NMAD). On a month-by-month basis, very little variation in the statistics is evident between the different methods. The advantage of using the WPMM over adjusted power low will be presented in section 2c.

Figure 1 presents distribution of "error" estimates based on the V5 monthly products for the 79 months during December 1997–June 2004. Figure 1a presents the distribution of the 79 monthly NMAD values as calculated by Eq. (2). Mean and median values are both 0.18; the standard deviation equals 0.046. The other three panels present the distribution of the monthly radar–gauge accumulation differences based on 6153 values, defined as (b) $R - G$, (c) $(R - G)/(R + G)$, and (d) by the relative difference (RD) as defined by Eq. (3). These parameters were calculated for each gauge and for each month separately. Notice that we choose to replace the commonly used statistical term *error* (e.g., mean absolute *error*, relative *error*) with *difference* (i.e., mean absolute *difference*, relative *difference*):

$$\mathrm{RD} = |R - G|/G. \qquad (3)$$

The overall total rain amount from all 6153 gauge months combined is 585 760 mm. Total $R/G = 1.008$ and linear correlation coefficient $r = 0.95$. The normalized mean absolute value [defined by Eq. (2)] for all 6153 monthly $R - G$ values presented in Fig. 1b is 0.17. The standard deviation of the $R - G$ values (24 mm) relative to mean monthly gauge measured rain amount ($\langle G \rangle = 95.2$ mm) is 0.25. (This parameter is sometimes called the *fractional standard error*. Since in this case it is calculated for bias-free estimates [$\Sigma R = \Sigma G$], it is equal to the *normalized root-mean-square error* [NRMSE]. The fact that the NMAD is not equal to approximately 80% [square root of $2/\Pi$] of the NRMSE suggests that $R - G$ values in Fig. 1b are not normally distributed.) Additional useful information could be obtained from the distribution of the normalized $R - G$ values. The distribution of $(R - G)/G$ will have a very large positive skewness due to the nature of ratio nonnegative distributions such as *R/G*. Therefore, the statistic $(R - G)/(R + G)$ was chosen. While its mean is 0.0 as expected, its standard deviation value is 0.13. Interestingly, the frequency of RD reaches its highest value near zero. The mean and median RD values are 0.20 and 0.15, respectively. It is worth noting that the tails of the distribution in Figs. 1c and 1d are mainly associated with low rain accumulations. For example, for a distribution based on the 1538 gauge months that are associated with the highest rain accumulations (the top 25% *G*), the standard deviation of $(R - G)/(R + G)$ will

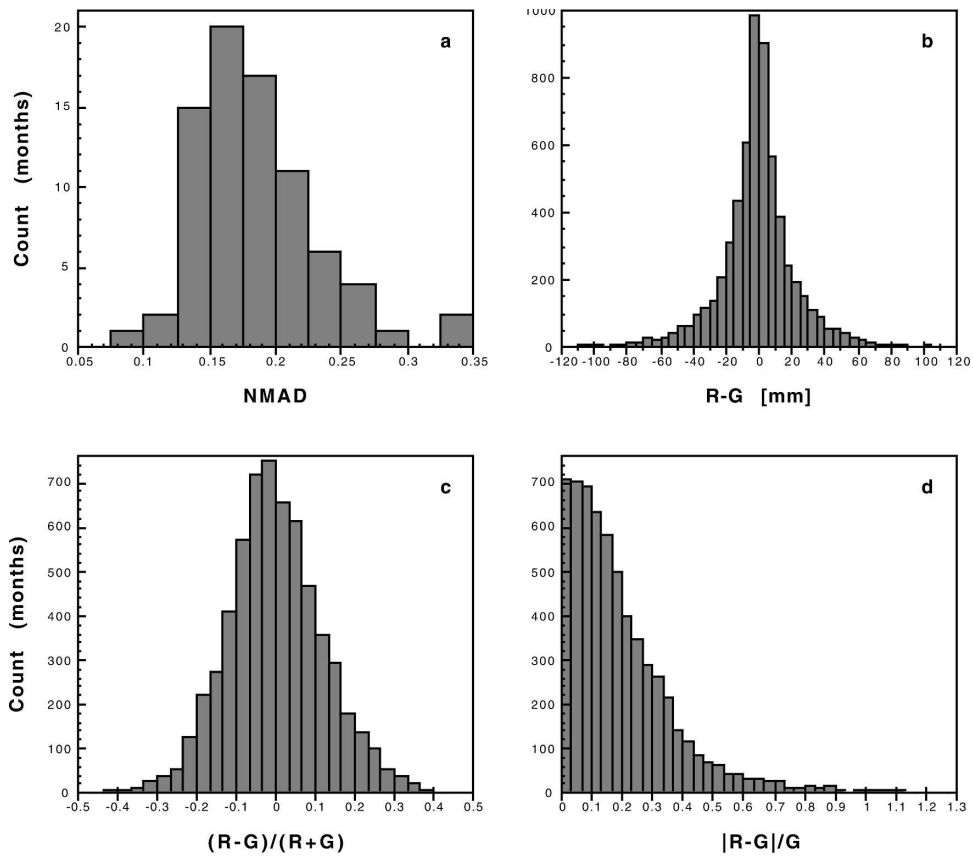

FIG. 1. The distribution of error estimates for the V5 monthly products for each of the months during December 1997–June 2004. (a) The distribution of the 79 monthly radar–gauge NMAD values as calculated by Eq. (2). (b)–(d) The distribution of the monthly radar (R)–gauge (G) accumulation differences based on 6153 values, defined as R-G in mm month^{-1} in (b); $(R - G)/(R + G)$ in (c); and $|R - G|/G$ in (d).

drop to 0.09, and the mean and median RD values will drop to 0.15 and 0.12, respectively.

While there are many other statistics that could be used to evaluate the radar accumulations, we maintained consistency by providing one metric, the NMAD, because its simple meaning is easily grasped. It was calculated in several variations to demonstrate that regardless of the exact way we manipulate the data, the average difference between the gauge and the radar estimates over the gauge location (2 km $\times$ 2 km) during a month is approximately 15%–20%. Other statistics could be used, such as RMS and the logarithm of the ratio, but just getting more statistics would not help to make this point.

It is worth mentioning that for a fixed $b = 1.4$, A was found, in general, to be larger for stratiform than for convective rain (V4). Allowing b to vary with A calculated according to Eq. (1), a minimum NMAD 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). Note that none of the official TRMM GV product's versions allows for a variable b value.

2) INDEPENDENT GAUGES

For MELB, August 1998 was a unique month in that a true independent validation of the TSP 3A-54 was possible using the DRGN gauge network. Independent validation of this specific monthly rain map is accomplished by validating against 15 gauges that were not used in the operational Z_e–R development. Figure 2 shows the differences between the radar estimates over the gauges and the gauge accumulations at times when the radar was operating. The radar overestimated the rain by 5% relative to the gauge accumulations. The NMAD on a monthly time scale was found to be 8%. These monthly statistics fall within acceptable bounds

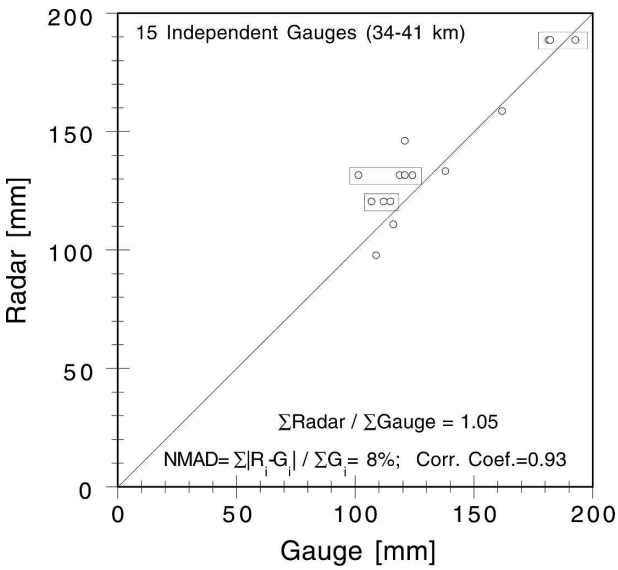

FIG. 2. The relations between the gauge and the radar-integrated rain rates over the Florida independent gauge network (DRGN) during August 1998. Gauges located within the same radar pixel of 2 km × 2 km are marked by rectangles. The radar estimates (TSP 3A-54, version 5) are based on WPMM Z_e–R relations using 21 quality-controlled gauges within 15–50 km from the radar.

and are consistent with the independent findings of Habib and Krajewski (2002). The calculated NMAD might vary from month to month. Its value as presented here does not represent the uncertainties for all monthly products, but rather is used to make the point that the radar–gauge differences might be of the same order of the difference in gauge accumulations within each radar pixel. The figure includes several gauges located within the same radar pixel of 2 km × 2 km. These gauges are marked by rectangles. The difference in gauge accumulations within each group is of the same order of the NMAD. Therefore, 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 instrument errors, might explain a major fraction of the NMAD. The NMAD 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.

Since the independent gauge data available are quite limited, the TSVO also tested a "quasi-independent" approach. Marks et al. (2003) presented in addition to a dependent validation approach, a quasi-independent approach to evaluate the monthly estimates at the Melbourne site. They used a Monte Carlo testing procedure to randomly select and withhold 10% of the gauges from a particular month from the WPMM Z_e–R development, applied the new Z_e–R lookup tables, and compared the resulting monthly rainfall accumulations with these withheld gauges. They arrived to almost identical statistics as for the case of using independent gauges. NMAD values ranged from 0.08 to 0.28 (see their Table 1). This is perhaps not surprising since we are dealing with many gauges, and the adjustment is performed by matching the radar estimates with the rain amounts from all the gauges combined and not with each individual gauge. Technically, this method does not evaluate the official monthly rainfall product; however, significant changes to the Z_e–R distributions have not been noted due to the small percentage of gauges withheld.

3) SENSITIVITY OF THE ESTIMATES TO THE AQC ALGORITHM

Monthly radar accumulation estimates 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), and to whether a single Z_e–R relation or dual (stratiform/convective) Z_e–R relations are applied (Table 1). However, they are very sensitive to the quality of the gauge data used for adjustment (Steiner et al. 1999). Independent QC of radar and rain gauges alone is not sufficient. Although collocated rain gauges will improve reliability, proper QC of rain gauge data upon comparison with radar data is essential. As described in section 2a, prior to the determination of the Z_e–R relations an AQC procedure (Amitai 2000) to filter unreliable rain gauge data upon comparison to radar data is used. The main filtering parameters (P1–P4) that are calculated for each gauge separately for a given month are as follows.

- The monthly fraction of rain depth, as measured by the radar (P1), at times when the gauge measured no rain (for every minute in the 15-min time window centered at the time of the scan). This value is calculated only when all 3 × 3 reflectivity pixels around the gauge exceed a given threshold in order to ensure that the rain is not local and therefore should be represented by the gauge.
- The monthly fraction of rain depth, as measured by the gauge (P2), at times when the radar did not observe precipitation echoes over the gauge.
- The correlation coefficient between the radar pixel over the gauge and the 7-min-averaged gauge-measured rain rate centered at the time of the radar scan, for all pairs during the month (P3).
- Monthly normalized *R/G* rain accumulation bias (P4). The normalization is performed with respect to

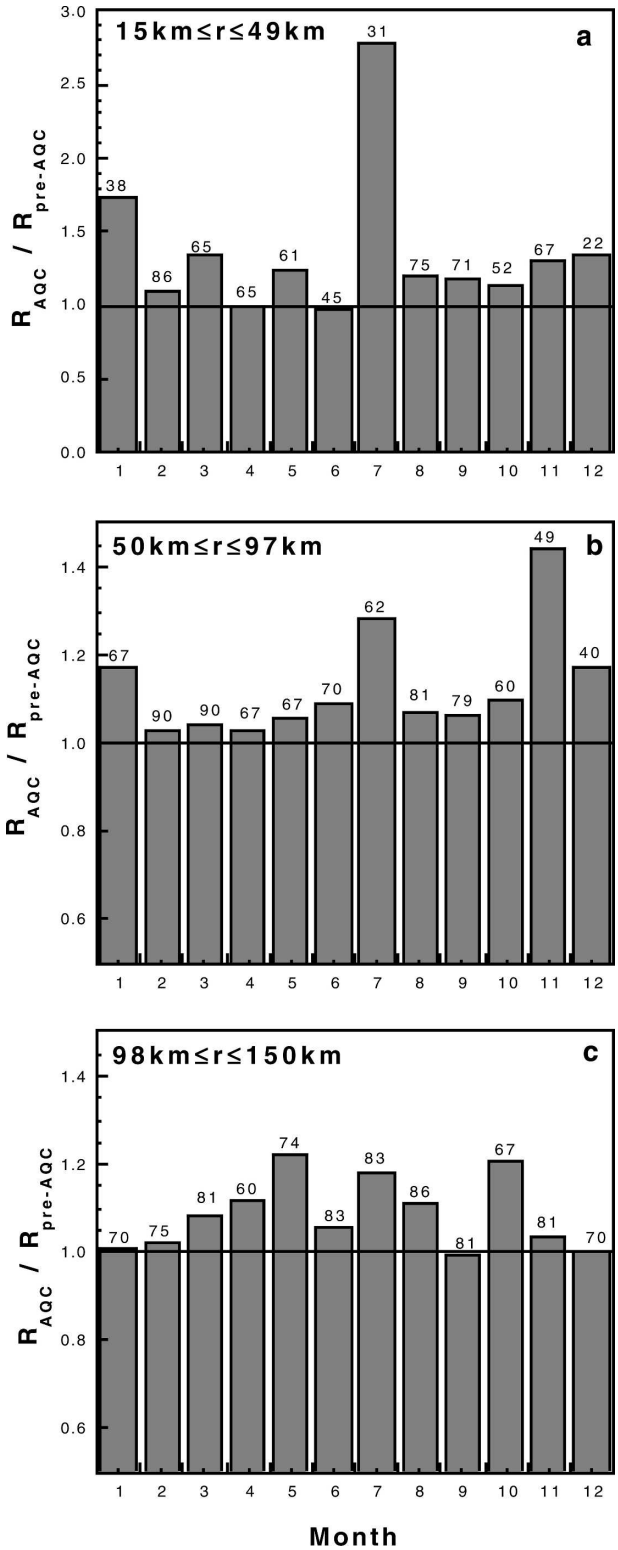

FIG. 3. The effect of the AQC algorithm on the radar rainfall estimates. The ratio of the monthly radar estimates derived from the post-AQC dataset to the estimates derived from the pre-AQC dataset, for each month during 1998 at Melbourne, FL. The number above each column represents the percentage of gauges approved for Z_e–R development by the AQC algorithm. Different panels represent different range internals from the radar.

the overall R/G bias calculated from all gauges combined that met each of the first three criteria.

The "monthly" parameters mentioned above represent the total accumulation derived from each of the 7-min accumulations centered at the time of the radar scan—they are not aimed at representing the actual monthly values. Moreover, data at times in which both the gauges and the radar observed no rain were excluded. All parameters had to be within a defined range of values in order to accept the monthly gauge data for generating the Z_e–R relations. The threshold/range value for each of the four parameters listed above are P1 < 0.4, P2 < 0.2, P3 > 0.15, and $0.5 <$ P4 < 2.0. If the individual gauge failed to meet any one of these criteria, its data and the collocated radar reflectivities were removed for the entire month. The use of radar information in addition to gauge information to decide which dataset to use for Z_e–R development perhaps compromises the independence of the gauge data to some degree. However, most gauge networks and those being used in this study do not have dual/triple gauges at each location to improve reliability of the gauge data. Moreover, the objective of the AQC algorithm is to detect gauges that for some period did not work properly or completely stopped working, and also to detect whether the radar had "bad" or missing data over the gauge.

Figure 3 provides approximation to the expected bias in the monthly radar rainfall estimates upon application of Z_e–R relations, which were developed after and prior to applying the AQC algorithm. For simplicity, instead of applying both sets of Z_e–R relations to the whole radar data, approximation to the ratio was derived from the A coefficients calculated [according to Eq. (1)] prior to and upon applying the AQC algorithm. In such an approach, only the radar reflectivities over the gauges and the gauge data are required [Eq. (4)]. The values in Fig. 3 are based on an exponent value of $b = 1.4$,

$$\langle R\rangle_{\mathrm{AQC}}/\langle R\rangle_{\mathrm{PRE\text{-}AQC}} = [A_{\mathrm{PRE\text{-}AQC}}/A_{\mathrm{AQC}}]^{1/b} = \left[\sum (Z_e^{1/b})\Big/\sum G\right]_{\mathrm{PRE\text{-}AQC}}\Big/ \left[\sum (Z_e^{1/b})\Big/\sum G\right]_{\mathrm{AQC}}. \quad (4)$$

The values are given for each range interval separately to be consistent with the TRMM GV standard products, which are based on the application of range-

dependent Z_e–R relations according to these three range intervals. As seen in Fig. 3, on average the AQC algorithm increases the monthly rainfall amount by 13%. The net effect was to remove data in which the gauge measured less rainfall than the amount the radar estimated over the gauge, probably since the gauge did not operate well during some periods. The fact that a larger increase of rain amount was found at close ranges as compared to far ranges might indicate that other factors such as ground clutter also played a role (notice the bias in Fig. 3a has a different scale than the other two panels). However, rainfall not reaching the gauges due to evaporation is probably not a factor since evaporation is expected to result with more gauges being removed at the furthest range interval in which the reflectivities are taken from a higher CAPPI. The number of rejected gauges in 1998 was found to be high relative to other years and compared to other locations. The total number of accepted gauges, for example, increases from 733 gauge months for 1998 (Table 1) to 946 gauge months for 1999. These numbers are believed to be closely associated gauge network maintenance, especially at far ranges in which ground clutter is not an issue. The reader is referred to Wolff et al. (2005) for a description of the MELB networks.

Since the radar estimates are strongly affected by the AQC algorithm, we need to make sure the algorithm is meeting expectations. The evaluation of the radar estimates is based on quality-controlled gauges, which are assumed to represent the truth. Therefore, we need to make sure that our error estimate statistics (e.g., NMAD) are valid, and are not a result of removing gauge and radar data from Z_e–R development just because otherwise they may produce large error estimates, but indeed represent bad data. It is quite difficult to evaluate the AQC algorithm if we do not have independent estimates. Rain gauges used for verification are assumed to be well quality controlled. This is a major concern in our evaluation effort. Dense rain gauge networks and dual/triple gauge configuration at a gauge site will assist in the determination of the "good" gauges and in testing the algorithm performance. In addition, information on atmospheric conditions might also be helpful. During the developmental stage of the algorithm, the verification of its performance as well as establishing the threshold values for each of the four parameters was done by using a graphical display in which the time series of gauge and radar accumulations were compared. The graphical display was not intended for determining total accumulations as the radar accumulations were computed using a nominal fixed Z–R relationship. It was used for pattern matching between the radar and gauge to determine geometrical and timing errors and other gauge measurement errors. An example of such a display can be found in Rosenfeld et al. (1995, see their Fig. 5). During the last seven years in which the algorithm has been used operationally, and has been applied to more than 150 months of data from Melbourne and Kwajalein combined, our confidence in its performance is based on 1) the fact that almost all rejected gauges are detected as "bad" by more than one criterion; 2) the fact that almost always the same gauges are rejected upon time shifting the radar data as it takes several minutes for raindrops to fall; 3) testing the sensitivity of the rainfall estimates and error estimates to the threshold values, especially the fourth parameter that might have direct impact on the resulting NMAD; 4) testing the performance of the algorithm at locations in which dual-gauge configurations exist. The reader should refer to Amitai (2000) for more information on the performance of the algorithm. Gabella and Notarpietro (2004) have modified the algorithm for computing the parameters on a daily basis (rather than monthly) in mountainous terrain. After having analyzed the distributions of the parameter values for the meteorological events considered in their study, they chose very similar threshold/range values to those used by TSVO. They modified P3 and calculated the correlation coefficient using gauge–radar data transformed on a logarithmic scale (computed using only radar–gauge data pairs in which both values are nonzero). The threshold value remained 0.15. This modification is based on Habib et al. (2001a), which explains why a log-transformation might be more suitable for correlation information.

c. *Evaluation of instantaneous estimates*

While monthly estimates are relatively simple to evaluate, the evaluation of instantaneous rainfall products is quite a challenge. At the small scale there are many sources of uncertainties. Scatterplots of point comparisons between radar estimates and rain gauge measurements are extremely noisy and are therefore of little use for evaluating the estimates. An alternative method based on analysis of the distribution of R as derived from gauge intensities and from reflectivities over the gauge network is used. Although such a technique still does not evaluate each instantaneous rainfall product alone, it does evaluate rain intensity properties rather than integral properties of rain intensity, that is, rain accumulations. Such an approach is significant for evaluating satellite products.

Determining and reducing the uncertainties in the GV PDFs is very important. The distribution of R is of great interest in many fields. For example, hydrological applications, such as flood forecasting, depend on an

accurate representation of the rainfall—driven by R—that does not infiltrate the soil. Better estimation of the spatial PDF of R is also crucial for better evaluation of rainfall products from space-based radars. Scatterplots of pixel-by-pixel comparisons of space-based R estimates with ground-based radar R estimates are extremely noisy mainly because of sample volume discrepancies, timing, and navigational mismatches, and uncertainties in the observed-radar reflectivity rain-rate Z_e–R relations. Furthermore, comparisons of rainfall over daily, weekly, or even monthly time scales suffer from the temporal sampling errors of the satellite where the revisit time is on the order of hours or days [e.g., TRMM, Global Precipitation Measurement (GPM) mission—a follow-on research satellite to TRMM]. Consequently, comparing space-based radar PDFs with PDFs derived from collocated ground-based radar estimates is attractive for evaluating satellite-based precipitation products, such as those from the TRMM precipitation radar (PR) (Liao et al. 2001; Amitai et al. 2004, 2005).

The PDFs being developed for this study represent the distribution of rain volume by rain rate, as they are constructed according to the relative contribution made by each rain intensity to the total rain volume. The volumetric PDF is defined as the sum of the rain rates for a given 1-dBR interval {dBR = 10 log(R/1 mm h^{-1}); [R] = (mm h^{-1})} divided by the total sum of the rain rates; that is,

$$\mathrm{PDF}(R_i) = R_i P(R_i)\, dR \Big/ \int_o^\infty RP(R)\, dR, \qquad (5)$$

where $P(R)$ is the probability of rain-rate occurrence. Such PDFs have direct hydrological significance and are less sensitive to the instrument rain detection limits than the PDFs of occurrence. This is especially important for comparison of PDFs based on estimates derived from different instruments (e.g., rain gauge, ground- and space-based radar), each characterized by a different detection limit.

Two issues must be addressed when one uses radar-based R PDFs: 1) the effect on the PDF of converting Z_e to R; and 2) the extent to which the gauge-based PDFs, 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. Here we focus on the sensitivity of radar-derived PDFs to gauge adjustment techniques. Monthly mean PDFs of R over a large space domain (up to 100 km from the radar) are used. When using PDFs based on smaller space–time domains other issues should be considered, such as the sensitivity of the PDF of Z_e to range (e.g., Chumchean et al. 2004) and the sensitivity of the PDF to rain type (e.g., Amitai 1999).

For each volumetric PDF, a volumetric cumulative distribution function (CDF) is also generated. While the PDFs describe the relative precipitation amount for a given rain-rate range, the CDFs describe the contribution to the total rain amount by intensities below a given rain rate:

$$\mathrm{CDF}(R_i) = \int_o^{R_i} R_i P(R_i)\, dR \Big/ \int_o^\infty RP(R)\, dR. \qquad (6)$$

In Figs. 4a and 4b PDFs and CDFs of the ground-based radar-estimated Rs over the gauges are generated for different gauge-adjusted power laws $Z_e = AR^b$. Several b values are tested, while A in each trial is tuned to the monthly gauge accumulations. These selected b values represent the dynamic range of values commonly used for rain-rate estimates. The bulk adjustment is performed in such a way that the total monthly rainfall, as estimated from Z_es over all the gauge locations combined, is matched to the accumulation from all gauges (6246 mm). All quality-controlled 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.

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. 4c). The range of averaging times used here is in accordance with the optimal range of ΔT for the best correlation between the instantaneous radar estimates and the gauge observations found by Habib and Krajewski (2002) for the same dataset. Thus, the PDFs are based on a range of time intervals that is expected to include that which represents the 2 km × 2 km spatial scale of the radar pixel. As the space–time domain increases, the PDF trends toward lower R values, such that weak intensities have higher probabilities and strong intensities have lower probabilities. This shift in the PDF shape is not new and has been demonstrated using other rain gauge datasets by presenting the CDF of rain-rate occurrence (e.g., Jones and Sims 1978; Calheiros and Zawadzki 1987). In the current study, as mentioned earlier, the PDFs of rain volume are presented, rather than the PDFs of occurrence, since the former are less sensitive to the instrument rain detection limits. However, the main difference is that these mentioned studies and others were not aimed at representing the distribution at a radar pixel size, nor were comparisons

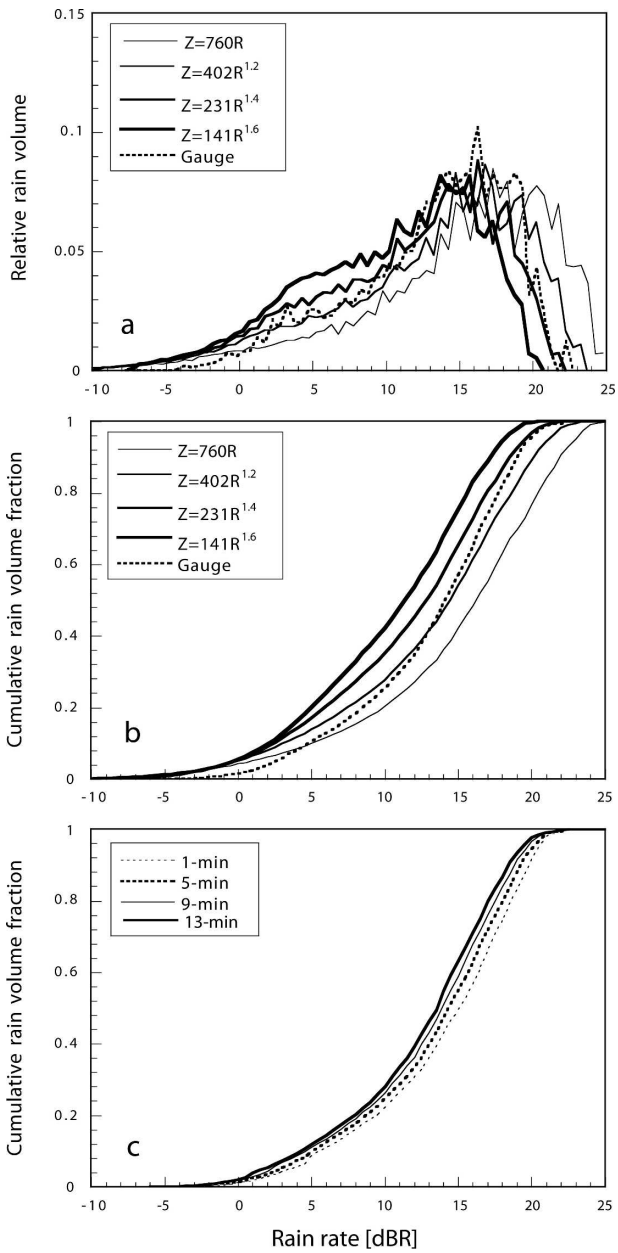

FIG. 4. Distribution of rain volume by rate for the Melbourne, FL, area during August 1998. (a) The PDFs are derived from 2 km × 2 km reflectivity pixels at a 1.5-km height directly over 63 quality-controlled gauges for different gauge-adjusted power laws (solid curves), and from 7-min-averaged gauge rates measured by the same 63 gauges centered at the radar scan time (broken curve). (b) The CDFs for the same dataset. (c) The CDFs as derived from the 63 gauge intensities for different time averaging. Note: $\mathrm{dB}R = 10\log(R/1\ \mathrm{mm\ h^{-1}})$; $[R] = (\mathrm{mm\ h^{-1}})$.

made with derived distribution of radar rates. These major issues are dealt with in the following paragraph.

Comparing Fig. 4b with Fig. 4c reveals that the shape of the PDF is less sensitive to the temporal average of the gauge-derived rain rates than is the choice of the parameter b in the gauge-adjusted power law. This suggests that 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. This is demonstrated here for the radar pixels under which gauges are present, but was found to be the case for independent pixels as well by examining the reproducibility of the gauge PDF. In this study, the data were taken from a summer month in which convective rain dominates. For months when stratiform rain dominates, characterized by less intense rain (i.e., narrow PDF), and the duration of a given rain-rate bin is longer, it is expected that the uncertainties in the estimates of the true distribution of R at the scale of a radar pixel will be further reduced using the gauge-derived PDFs.

The results support the use of the gauge PDFs, and therefore, they support the use of the WPMM for estimation of the rain-rate distribution. The WPMM provides a relationship between Z_e and R by forcing the distribution of R, as derived from Z_e over 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 the Z_e over the gauge after application of the WPMM. In contrast, an adjusted power law might have A and b values that will generate a PDF that diverges from the distribution of the gauge rates and possibly diverge from the true R distribution. It is therefore expected that the new NASA TRMM validation products (V5), based on WPMM, will generate PDFs with improved accuracy.

It is worth mentioning that while the shape of the R distribution is found to be very sensitive to b, an adjusted power law may yield satisfactory estimates of rainfall *accumulations* for a wide range of b values. Among the gauge-adjusted power laws that ensure unbiased total accumulations, $b = 1.5$ yields the minimum NMAD between the monthly radar estimates and the gauge accumulations for the given month investigated (August 1998). The minimum NMAD, 17.6%, increases only to 17.8% and 18.2% for $b = 1.4$ and $b = 1.3$, respectively. However, $b = 1.5$ yields a PDF that is far from representing any of the gauge-based PDFs.

3. Limiting factors in the evaluation effort

Rain quantities derived from different instruments represent different time and space domains that often differ significantly. For example, the sampling area of a typical National Weather Service (NWS) NEXRAD radar-rainfall product (2 km by 2 km) and that of a rain gauge differ by eight orders of magnitude. In time, the radar estimates are based on instantaneous observa-

tions averaged over an area of a certain size, while the gauge data represent accumulation over a defined time period (quasi point). 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 the scales smaller than that of a radar pixel (e.g., Krajewski et al. 2003), a single gauge randomly placed is not capable of adequately representing such variability, 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 dataset it was found on a *monthly* time scale that the average difference between the radar estimates over the gauges and the gauge accumulations 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 over the gauges to those derived from the gauge intensities. However, due to the spatial variability of R and gauge instrument problems, 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 represent the actual rain-rate distribution at the scale of a radar pixel? The difficulty in addressing this question 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 theory on how to model the effect of scale on a PDF is set out in VanMarcke (1983). However, the best the existing methodologies can do, based on the data collected during the TRMM field campaigns, is to estimate the overall radar-rainfall error variance using 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 to examine the effect of domain size on the gauge-based rain-rate PDFs. This approach was used here. However, these do not provide the full PDF of errors. It should also be noted that the tipping-bucket gauges have a threshold sensitivity such that 0.254 mm of rainfall depth has to be exceeded for the gauge to record data. Therefore, there is some uncertainty associated with low rainfall intensities, which is a result of the limited sensitivity of the instrument (e.g., Habib et al. 2001b). Some uncertainty is also associated with other general problems of tipping-bucket gauges, such as the amount of rain being missed by the gauge at times of strong wind (e.g., Nespor and Sevruk 1999; Steiner et al. 2005), and at times of very high rain rates due to the bucket not tipping fast enough (e.g., Marsalek 1981). These problems among others, however, are assumed to have a much smaller effect on the PDF's shape than the threshold sensitivity effect.

While uncertainties in the determination of the ground-based radar PDFs are likely to be reduced upon adjustment with those from gauge data, as demonstrated here (a major motivation for NASA TRMM validation program's product generation based on WPMM Z_e–R relations), further investigation is needed to determine how to convert a point PDF to a spatial PDF. To better represent the true areal distribution and to further reduce the uncertainties in the derived PDFs, additional information, such as horizontal advection based on echo motion calculations, may be required. Denser gauge networks for accuracy verification of the derived PDFs as well as for direct validation of rainfall estimates should be used to further reduce the uncertainties. These should have the right spatial configuration to minimize the spatial sampling random error. To improve the reliability of the data, each of the locations should be equipped with collocated (two or more) rain gauges (e.g., Ciach and Krajewski 1999; Tokay et al. 2003). The dual-gauge design allows for fault detection, recovery of lost data, and reduction of the point measurement error. Such gauge network design will assist to evaluate the AQC algorithm, and in turn, to improve the quality of the merged radar–gauge dataset for Z_e–R development. Dual-gauge configuration is also essential to allow independent validation of the radar estimates by having quality-controlled gauges without using the AQC algorithm, which, as noted earlier, compromises to some degree the independence of the gauge data. Only when the collected dataset is large (several years) is it possible to conduct conditional (on rain) analysis. This is due to the highly skewed nature of the rainfall intensity distribution: the extreme events do not happen very often.

4. Lessons learned

a. A value without an uncertainty value is not much of a value

Is our validation effort being optimally expended? Considerable effort is devoted to data quality control, improvement of algorithms for remote sensing precipitation estimates, basic comparison between ground-based rainfall products and satellite observations, and comparison between observations and numerically

modeled rain fields. However, are we devoting sufficient effort to analysis and modeling of uncertainties and error structure in the precipitation products? According to the TRMM GV experience, we are not. Simply processing data and generating products are not sufficient; we must evaluate the products and determine the systematic errors and uncertainties. Data assimilation and many hydrologic applications require satellite observations of precipitation. However, providing values of precipitation is not sufficient unless it is accompanied by the associated uncertainty estimates. A comprehensive TRMM GV program was established to evaluate the satellite observations. The convergence of the TRMM GV and satellite estimates bodes well for expectations for the proposed GPM program, but it is now understood that providing uncertainties of the estimates is perhaps more important than convergence on its own. While the current GV program includes several procedures to evaluate the GV products (which improve algorithms and products), the validation effort for future missions, such as GPM, should place greater emphasis on error structure characterization and uncertainty determination in near–real time, and understanding the processes that lead to these uncertainties. A protocol should be formulated for standardizing evaluation and assessing uncertainty of estimates.

b. Focus on instantaneous products

In the TRMM GV program considerable effort is devoted to climatological validation and to generating accurate *monthly* products. As mentioned in section 2c, comparisons of such products with the satellite monthly products suffer from the temporal sampling errors of the satellite. Effort should be shifted toward generating more accurate *instantaneous* products and determining their uncertainties. While any gauge-adjustable Z_e–R relations may yield satisfactory estimates of rainfall accumulations over a large space–time domain, the rain rates might be systematically biased. Switching to the new version of GV products (V5), which uses WPMM to allow for better estimation of the true PDF of R, seems to be in the right direction with the potential to improve evaluation of satellite estimates upon comparisons of ground and space-based radar PDFs. However, these Z_e–R relations are still based on *monthly* datasets. Moreover, the quality-controlled effort is a very time-consuming task. Relations based on smaller datasets and upon rain-type classification should be considered.

c. Quality of radar and gauge datasets

Radar calibration stability checking is a must. Routine comparisons to TRMM (or GPM) PR reflectivities, if available, might be the best approach (e.g., Anagnostou et al. 2001). A 1-dB undetected reflectivity shift (which is quite common) corresponds to 15%–25% R uncertainty.

Since we adjust the radar estimates to the gauges, the most significant decision to make could easily be which gauges to use (as demonstrated in Fig. 3), or in other words, how to make sure we are using QC gauge data. Selection of the reflectivity to rainfall conversion algorithm and whether to apply a single Z_e–R relation or stratiform/convective dual relations are not nearly as critical in determining the monthly statistics as is the determination about which gauges to use (as demonstrated in Table 1). This has important implications to future projects, such as the National Oceanic and Atmospheric Administration's (NOAA) National Mosaic and Multisensors Quantitative Precipitation Estimation (NMQ) Project, which will generate gauge-adjusted rainfall products (Seo et al. 2005). Collocation (in sets of two or three) of rain gauges and redundancy in gauge loggers are essential. It is noted that the 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. By tying the collocated gauges to a pair of loggers, many of these dropouts can be recognized and addressed.

Independent quality control of radar and rain gauges alone is not sufficient. Although collocated rain gauges will improve reliability, proper QC of the merged radar–gauge data is essential prior to development of any type of gauge-adjustable Z_e–R relations. The monthly radar accumulation estimates, for example, were found to be more affected by the AQC algorithm rather than by the particular gauge adjustment method used to derive the radar estimates.

d. Dense gauge network, gauge configuration, and gauge type

Lack of existence of a very dense gauge network was found to be the limiting factor in the evaluation process of both instantaneous and monthly products. Dense gauge networks have been designed to study spatial variability of rainfall. However, networks also need to be designed to provide sufficient gauge sampling to match a radar pixel. Such networks will allow for 1) improved estimation of the true rain-rate distribution at the scale of a radar pixel; 2) direct validation of radar rainfall accumulation estimates; and 3) improved reliability of the AQC algorithm, and of the merged radar–gauge datasets.

Improved estimation of the true rain-rate distribution also requires synchronized gauges and gauges that

are able to measure weak intensities. Tipping buckets are not good enough. This is especially important for the validation of the GPM products, which will cover higher latitudes.

The number of gauges and configuration of the dense network that is required depend on a list of factors, such as the application (e.g., monthly validation of radar estimates, better estimation of the true PDF of R, evaluation of the AQC algorithm); the accuracy requirement of that application; the radar sampling resolution; the typical climatological regimes (autocorrelation structure of rainfall associated with the natural variability of rain); the gauge type; the statistics of gauge random and systematic errors (e.g., wind effects); and other factors. Based on TSVO experience gained, for better estimation of the true PDF of R at the Melbourne GV site using the current tipping buckets, we would recommend about 20 gauge sites evenly spaced within a radar pixel of 2 km by 2 km with two gauges at each site. This is supported by studies such as Krajewski (2001), which suggests 17 sampling locations to obtain a 6% level of uncertainty (random error with no bias, i.e., disregarding wind effect on the bias) over a 2 km by 2 km domain. This assumes the correlation function of summer rainfall in Florida as estimated by Krajewski et al. (2003) and Habib and Krajewski (2002) using the 1998 DRGN data. Other tropical locations, such as Brazil and Guam, display similar levels of rainfall variability as measured by the correlation distance (Krajewski et al. 2003). The cost/benefit ratio of additional gauge sites might be large due to gauge systematic errors and the time it takes to complete a single measurement compared to the distance the storm propagates in such a time interval. The TSVO is currently conducting tests on gauge types and more recommendations are expected in the future.

5. Closing remarks

Evaluation of TRMM satellite observations is conducted through a comprehensive GV program. Since the launch of TRMM in late 1997, standardized instantaneous and monthly rainfall products have been routinely generated using quality-controlled ground-based radar data adjusted to the gauge accumulations from four primary sites. For example, 79 months' worth of rainfall products from MELB were available for this study. These products are available to the scientific community and to the general public. As part of the TRMM GV program, efforts are being made to evaluate these products. Insights on the accuracy of these products were presented. In general, the mean NMAD between the gauge accumulations and the radar accumulations over the gauges on a monthly time scale was found to be around 15%–20%. In addition to "common use" comparisons between point gauge accumulations and radar estimates over the gauges, the uncertainties in the rainfall estimates were evaluated by 1) utilizing a dense network of independent gauges in which more than one gauge exists within a radar pixel, 2) conducting sensitivity tests of the distribution of the radar-derived rain rates to different gauge-adjusted Z_e–R relations, 3) testing the sensitivity of the estimates to the algorithm used to select the quality gauges for Z_e–R generation, and 4) testing the sensitivity of the estimates to the application of a single Z_e–R relation as compared to dual convective/stratiform Z_e–R relations. These procedures allowed us to evaluate the rainfall estimates, to improve the algorithms in order to develop better GV products for comparison with the satellite products, and to recognize the major limiting factors in evaluating the estimates that reflect current limitations in radar rainfall estimation using conventional radar systems, that is, nonpolarimetric or dual-wavelength systems. Lessons learned and suggested improvements from this 8-yr mission were summarized.

The effort of evaluating the TRMM ground validation radar rainfall products described here is based solely on ground-based observations, that is, rain gauge measurements. Comparison with the TRMM satellite rainfall products is an independent effort beyond the scope of this paper. The reader interested in such comparison is referred to publications such as Bolen and Chandrasekar (2000, 2003), Schumacher and Houze (2000), Anagnostou et al. (2001), Liao et al. (2001), Keenan et al. (2003), Fisher (2004), and Amitai et al. (2004, 2005).

Acknowledgments. This effort is supported by the TRMM Satellite Validation Office at NASA Goddard Space Flight Center headed by Richard Lawrence, through approvals by Drs. Robert Adler (U.S. TRMM Project Scientist) and Ramesh Kakar (Precipitation Program Scientist at NASA HQ).

REFERENCES

Amitai, E., 1999: Relationships between radar properties at high elevations and surface rain rate: Potential use for spaceborne rainfall measurements. *J. Appl. Meteor.,* **38,** 321–333.

——, 2000: Systematic variation of observed radar reflectivity–rainfall rate relations in the Tropics. *J. Appl. Meteor.,* **39,** 2198–2208.

——, J. A. Nystuen, L. Liao, R. Meneghini, and E. Morin, 2004: Uniting space, ground and underwater measurements for improved estimates of rain rate. *IEEE Geosci. Remote Sens. Lett.,* **1,** 35–38.

——, L. Liao, X. Llort, and R. Meneghini, 2005: Accuracy verifi-

cation of spaceborne radar estimates of rain rate. *Atmos. Sci. Lett.,* **6,** 2–6.

Anagnostou, E. N., C. A. Morales, and T. Dinku, 2001: The use of TRMM precipitation radar observations in determining ground radar calibration biases. *J. Atmos. Oceanic Technol.,* **18,** 616–628.

Bolen, S. M., and V. Chandrasekar, 2000: Quantitative cross validation of space-based and ground-based radar observations. *J. Appl. Meteor.,* **39,** 2071–2079.

——, and ——, 2003: Methodology for aligning and comparing spaceborne radar and ground-based radar observations. *J. Atmos. Oceanic Technol.,* **20,** 647–659.

Calheiros, R. V., and I. Zawadzki, 1987: Reflectivity–rain rate relationships for radar hydrology in Brazil. *J. Climate Appl. Meteor.,* **26,** 118–132.

Chumchean, S., A. Seed, and A. Sharma, 2004: Application of scaling in radar reflectivity for correcting range-dependent bias in climatological radar rainfall estimates. *J. Atmos. Oceanic Technol.,* **21,** 1545–1556.

Ciach, G. J., and W. F. Krajewski, 1999: On the estimation of radar rainfall error variance. *Adv. Water Resour.,* **22,** 585–595.

Fisher, L. B., 2004: Climatological validation of TRMM TMI and PR monthly rain products over Oklahoma. *J. Appl. Meteor.,* **43,** 519–535.

Gabella, M., and R. Notarpietro, 2004: Improving operational measurement of precipitation using radar in mountainous terrain—Part I: Methods. *IEEE Geosci. Remote Sens. Lett.,* **1,** 78–83.

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

——, ——, and G. J. Ciach, 2001a: Estimation of rainfall interstation correlation. *J. Hydrometeor.,* **2,** 621–629.

——, ——, and A. Kruger, 2001b: Sampling errors of fine resolution tipping-bucket rain gauge measurements. *J. Hydrol. Eng.,* **6** (2), 159–166.

Jones, D. M. A., and A. L. Sims, 1978: Climatology of instantaneous rainfall rates. *J. Appl. Meteor.,* **17,** 1135–1140.

Keenan, T., E. Ebert, V. Chandrasekar, V. Bringi, and M. Whimpey, 2003: Comparison of TRMM satellite-based rainfall with surface radar and gauge information. Preprints, *31st Int. Conf. on Radar Meteorology,* Seattle, WA, Amer. Meteor. Soc., 383–386.

Krajewski, W. F., 2001: Radar-rainfall estimation: Monitoring the progress. *Proc. Fifth Int. Symp. on Hydrological Applications of Weather Radar,* Kyoto, Japan, International Asasociation of Hydrological Sciences, 1–6.

——, G. J. Ciach, and E. Habib, 2003: An analysis of small-scale rainfall variability in different climatological regimes. *Hydrol. Sci. J.,* **48** (2), 151–162.

Kulie, M. S., M. Robinson, D. A. Marks, B. S. Ferrier, D. Rosenfeld, and D. B. Wolff, 1999: Operational processing of ground validation data for the Tropical Rainfall Measuring Mission. Preprints, *29th Int. Conf. on Radar Meteorology,* Montreal, QC, Canada, Amer. Meteor. Soc., 736–739.

Liao, L., R. Meneghini, and T. Iguchi, 2001: Comparisons of rain rate and reflectivity factor derived from the TRMM precipitation radar and the WSR-88D over the Melbourne, Florida, site. *J. Atmos. Oceanic Technol.,* **18,** 1959–1974.

Marks, D. A., D. S. Silberstein, D. B. Wolff, E. Amitai, J. L. Pippitt, and R. J. Lawrence, 2003: The TRMM ground validation effort at NASA-GSFC: Validating surface rainfall estimates. Preprints, *31st Int. Conf. on Radar Meteorology,* Seattle, WA, Amer. Meteor. Soc., 415–418.

——, and Coauthors, 2000: Climatological processing and product development for the TRMM ground validation program. *Phys. Chem. Earth B,* **25,** 871–875.

Marsalek, J., 1981: Calibration of tipping bucket raingage. *J. Hydrol.,* **53,** 343–354.

NASA, cited 2003: TRMM Satellite Validation Office: Gauge data processing and availability logs, rain interpolation algorithm. [Available online at http://trmm-fc.gsfc.nasa.gov/trmm_gv/index.html.]

Nespor, V., and B. Sevruk, 1999: Estimation of wind-induced error of rainfall gauge measurements using a numerical simulation. *J. Atmos. Oceanic Technol.,* **16,** 450–464.

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

——, E. Amitai, and D. B. Wolff, 1995: Classification of rain regimes by the three-dimensional properties of reflectivity fields. *J. Appl. Meteor.,* **34,** 198–211.

Schumacher, C., and R. A. Houze Jr., 2000: Comparison of radar data from TRMM satellite and Kwajalein oceanic validation site. *J. Appl. Meteor.,* **39,** 2151–2164.

Seo, D.-J., C. R. Kondragunta, D. Kitzmiller, K. Howard, J. Zhang, and S. V. Vasiloff, 2005: The National Mosaic and Multisensors QPE (NMQ) Project—Status and plans for a community testbed for high-resolution multisensor quantitative precipitation estimation (QPE) over the United States. Preprints, *19th Conf. on Hydrology,* San Diego, CA, Amer. Meteor. Soc., CD-ROM, 1.3.

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

——, J. A. Smith, S. J. Burges, C. V. Alonso, and R. W. Darden, 1999: Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation. *Water Resour. Res.,* **35,** 2487–2503.

——, L. C. Sieck, and S. J. Burges, 2005: Revisiting "ground truth"—Assessment of wind and out-of-level collector funnel effects on rain gauge catch. Preprints, *32d Int. Conf. on Radar Meteorology,* Albuquerque, NM, Amer. Meteor. Soc., CD-ROM, P6R.7.

Tokay, A., D. B. Wolff, K. R. Wolff, and P. Bashor, 2003: Rain gauge and disdrometer measurements during the Keys Area Microphysics Project (KAMP). *J. Atmos. Oceanic Technol.,* **20,** 1460–1477.

VanMarcke, E., 1983: *Random Fields: Analysis and Synthesis.* MIT Press, 382 pp. [Web edition published by Rare Book Services, Princeton University, 1998.]

Wolff, D. B., D. A. Marks, E. Amitai, B. L. Fisher, D. S. Silberstein, A. Tokay, J. Wang, and J. L. Pippitt, 2005: Ground validation for the Tropical Rainfall Measuring Mission (TRMM). *J. Atmos. Oceanic Technol.,* **22,** 365–380.