Comparison of TRMM 2A25 Products, Version 6 and Version 7, with NOAA/NSSL Ground Radar–Based National Mosaic QPE

PIERRE-EMMANUEL KIRSTETTER,* Y. HONG,* J. J. GOURLEY, M. SCHWALLER, W. PETERSEN, AND J. ZHANG

* School of Civil Engineering and Environmental Sciences, University of Oklahoma, Norman, Oklahoma
+ NOAA/National Severe Storms Laboratory, Norman, Oklahoma
# Atmospheric Radar Research Center, National Weather Center, Norman, Oklahoma
@ NASA Goddard Space Flight Center, Greenbelt, Maryland
& NASA Wallops Flight Facility, Wallops Island, Virginia

(Manuscript received 29 February 2012, in final form 16 August 2012)

ABSTRACT

Characterization of the error associated with satellite rainfall estimates is a necessary component of deterministic and probabilistic frameworks involving spaceborne passive and active microwave measurements for applications ranging from water budget studies to forecasting natural hazards related to extreme rainfall events. The authors focus here on the relative error structure of Tropical Rainfall Measurement Mission (TRMM) precipitation radar (PR) quantitative precipitation estimation (QPE) at the ground by comparison of 2A25 products with reference values derived from NOAA/NSSL’s ground radar–based National Mosaic and QPE system (NMQ/Q2). The primary contribution of this study is to compare the new 2A25, version 7 (V7), products that were recently released as a replacement of version 6 (V6). Moreover, the authors supply uncertainty estimates of the rainfall products so that they may be used in a quantitative manner for applications like hydrologic modeling. This new version is considered superior over land areas and will likely be the final version for TRMM PR rainfall estimates. Several aspects of the two versions are compared and quantified, including rainfall rate distributions, systematic biases, and random errors. All analyses indicate that V7 is in closer agreement with the reference rainfall compared to V6.

1. Introduction

Given their quasi-global coverage, satellite-based quantitative rainfall estimates are becoming widely used for hydrologic and climatic applications. Characterizing the error structure of satellite rainfall products is recognized as a major issue for the usefulness of the estimates (Yang et al. 2006; Zeweldi and Gebremichael 2009; Sapiano and Arkin 2009; Wolff and Fisher 2009) as underlined by the Program to Evaluate High-Resolution Precipitation Products (Turk et al. 2008) led by the International Precipitation Working Group (IPWG; see http://www.isac.cnr.it/~ipwg/). In this study, we focus on the Tropical Rainfall Measurement Mission (TRMM) precipitation radar (PR) quantitative precipitation estimation (QPE) product.

The TRMM PR is currently the only active instrument dedicated to the measurement of rainfall from a satellite platform conjointly with a radiometer [TRMM Microwave Imager (TMI)]. PR measurements are considered as the starting point for subsequent algorithms that use microwave measurements from low-earth-orbiting satellites and for combined end products that utilize data from geostationary satellites (e.g., Yang et al. 2006; Wolff and Fisher 2008; Ebert 2007; Bergès et al. 2010; Ushio et al. 2006). A number of studies have investigated the quality of PR estimates in various regions of the world (e.g., Adeyewa and Nakamura 2003; Lin and Hou 2008; Michaelides 2008; Wolff and Fisher 2008, 2009). Over the United States, Amitai et al. (2009, 2012) have compared the PR with the National Oceanic and Atmospheric Administration/National Severe Storms Laboratory (NOAA/NSSL)’s ground radar–based

Corresponding author address: Professor Yang Hong, National Weather Center, 120 David L. Boren Blvd., Atmospheric Radar Research Center, Suite 4610, Norman, OK 73072-7303. E-mail: yanghong@ou.edu

DOI: 10.1175/JHM-D-12-030.1

© 2013 American Meteorological Society
National Mosaic and QPE system (NMQ/Q2). Our aim is to perform a systematic and comprehensive comparison of the new PR 2A25, version 7 (V7), products that were recently released as a replacement for version 6 (V6) over the southern conterminous United States (CONUS). This will likely be the final version of TRMM PR rainfall estimates, and a major outcome of this study is to supply uncertainty estimates for users of the data. This new version is considered superior over land areas compared to the previous versions because of changes to the vertical profile of hydrometeor characteristics, which affects the reflectivity-to-rainfall rate \((Z-R)\) relationship and attenuation correction. Finally, a correction for non-uniform beam-filling (NUBF) effects was reintroduced.

The methodology and framework followed here are described in a previous paper dedicated to the evaluation of 2A25, V6 (Kirstetter et al. 2012). The PR QPE product was assessed with respect to an independent reference rainfall dataset derived from high-resolution measurements using NOAA/NSSL’s NMQ/Q2 (Zhang et al. 2011). These products yield instantaneous rain-rate products over vast regions including the CONUS. A systematic and comprehensive evaluation for regions over the southern CONUS was performed by characterizing errors in PR estimates by matching quasi-instantaneous data from Q2 at the \(\sim 5\)-km-pixel measurement scale of PR in order to minimize uncertainties caused by resampling. The study used 3 months (March–May 2011) of satellite overpasses over the lower CONUS (up to \(36^\circ\)N). Despite the seemingly short period for evaluation, the use of gridded Q2 data for reference provided a large sample size totaling 392 713 nonzero PR reference pairs. The exact same reference dataset that was used to evaluate V6 is used in this study for V7.

The PR and Q2 reference data are briefly described in section 2. In section 3 we assess the differences in the probability density functions (PDFs) of rain rate for 2A25, V6 and V7, and their ability to represent rainfall variability. A quantitative comparison of empirical error models for V6 and V7 estimates versus reference rainfall is provided in section 4. The paper is closed with concluding remarks in section 5.

2. Data sources

a. Q2-based reference rainfall

All significant rain fields observed coincidentally by TRMM overpasses and the Next Generation Weather Radar (NEXRAD) network from March to May 2011 are collected. The NMQ/Q2 products closest in time to the TRMM satellite local-overpass schedule time are used.

The NOAA/NSSL National Mosaic and Quantitative Precipitation Estimation system (http://nmq.ou.edu; Zhang et al. 2011) combines information from all ground-based radars comprising the Weather Surveillance Radar-1988 Doppler (WSR-88D) network to derive experimental radar-based products comprising high-resolution \((0.01^\circ, 5\) min) instantaneous rainfall rate mosaics available over the CONUS (Zhang et al. 2005; Lakshmanan et al. 2007; Vasiloff et al. 2007; Kitzmiller et al. 2010). At hourly time steps, Q2 adjusts radar estimates with automated rain gauge networks using a spatially variable bias multiplicative factor. A radar quality index (RQI) is produced to represent the radar QPE uncertainty associated with reflectivity changes with height and near the melting layer (Zhang et al. 2012). One should note that it is not possible to “validate” the PR estimates in a strict sense because independent rainfall estimates with no uncertainty do not exist. Many errors affect the estimation of rainfall from ground-based radars, such as nonweather echoes, NUBF, range dependency due to vertical profile of reflectivity (VPR) variability, conversion of \(Z\) to \(R\), and calibration of the radar signal [see Villarini and Krajewski (2010) for a recent review]. While several procedures are already in place within the Q2 system to correct for these errors, additional postprocessing steps were taken to refine the reference dataset as much as possible. The original Q2 products utilized in this study are (i) the radar-only instantaneous rain-rate national mosaic updated every 5 min, (ii) the radar-only rain-rate national mosaic at hourly time steps, (iii) the hourly rain gauge–corrected national mosaic product, and (iv) the RQI. The reference rainfall is derived from an instantaneous bias-corrected Q2 product. Instantaneous Q2 products are adjusted using a spatially variable multiplicative bias field to minimize the aforementioned errors: pixel-by-pixel ratios between the hourly gauge-adjusted products and the hourly radar-only products are calculated and applied as multiplicative adjustment factors to the radar-only 5-min product. Extreme adjustment factors (outside the 0.1–10 range) are discarded so the gauge adjustment also serves as a data quality control procedure. To eliminate overestimation in the bright band and mitigate range dependency caused by VPR effects, a filtering is finally applied using the RQI index so only Q2 estimates representing the best measurement conditions (i.e., no beam blockage and radar beam below the melting level of rainfall) are retained. One must keep in mind these improvements may not screen out all possible errors in ground-based radar estimates.

The reference rainfall \(R_{\text{ref}}\) is computed from a block-Q2 rainfall pixel matching each PR pixel. All of the Q2
pixels (rainy and nonrainy) found within an approximate 2.5-km radius around the center of the PR pixel location are considered to compute unconditional mean rain rates for the Q2 at the PR pixel scale. The estimation reference quality is assessed using a standard error $\sigma_{\text{footprint}}$ that is computed alongside the mean reference rainfall value. It represents the variability of the Q2 rainfall (at its native resolution) inside the PR footprint and is used to select the PR-Q2 reference pairs for which the $R_{\text{ref}}$ is trustworthy [see Kirstetter et al. (2012) for more details]. The reference pixels are segregated into “robust” ($R_{\text{ref}} > \sigma_{\text{footprint}}$) and “nonrobust” ($R_{\text{ref}} < \sigma_{\text{footprint}}$) estimators. Nonrobust reference values are discarded for quantitative comparison in this study. The current technique preserves the PR rainfall statistical characteristics (the product remains free of undesirable impacts caused by resampling) and increases the reliability and representativeness of our ground reference.

b. Precipitation radar–based rainfall

The PR measures reflectivity profiles at $K_u$ band. Surface rain rates are estimated over the southern United States up to 36°N [see Fig. 1 of Kirstetter et al. (2012)]. The scan geometry and sampling rate of the PR lead to footprints spaced approximately 5.1 km in the horizontal and along track, over a 245-km-wide swath. The TRMM product used in this work is the PR 2A25 product (versions 6 and 7) described in Iguchi et al. (2000, 2009), which provides 3D reflectivity and 2D rain-rate fields at the ground. The 2A25 algorithm relies on a hybrid attenuation correction method that combines the surface reference technique and Hitschfeld–Bordan method (Iguchi et al. 2000; Meneghini et al. 2000, 2004). Retrieval errors of the algorithm have mainly been attributed to the uncertainty of the assumed drop size distribution (DSD), incorrect physical assumptions (freezing-level height, hydrometeor temperatures), and NUBF effects (Iguchi et al. 2009). Some of the weaknesses previously reported in performance with V6 (i.e., underestimation of rain rates) over land compared to over sea (Wolff and Fisher 2008; Iguchi et al. 2009) are expected to improve as $Z$–$R$ relationships over land were recalibrated and the NUBF correction, which was abandoned in V6, was reintroduced in the new V7 product.

3. Rainfall data analysis

The case of PR having zero rainfall when it is raining (according to the reference) was addressed in Kirstetter et al. (2012, their section 3a) and highlighted the poor detection performances of the PR of light rain rates (Schumacher and Houze 2000). The major changes from version 6 to version 7 address rather the quantitative estimation of rainfall in 2A25 products than the detection of rainfall itself. Accordingly, we did not find any significant differences in performances in rainfall detection from the two versions, so we focus hereafter on specific cases when both PR and reference are nonzero.
a. Probability distributions by occurrence and by rain volume

Two PDFs for PR versus Q2 reference rainfall are computed and shown in Fig. 1: (i) the PDF by occurrence (PDF$_c$) and (ii) the PDF by rain volume (PDF$_y$) (Wolff and Fisher 2009; Amitai et al. 2009, 2012; Kirstetter et al. 2012). The PDF$_c$ represents the probability of rain-rate occurrence and tends to emphasize lighter rainfall rates. The PDF$_y$ represents the relative contribution of each rain-rate bin to the total rainfall volume. Compared to Q2’s reference PDF$_c$, both 2A25 versions tend to sample more light rain rates ($\sim0.3$–$0.5$ mm h$^{-1}$) and demonstrate poor detection of the lightest rain rates (below $\sim0.3$ mm h$^{-1}$). A possible explanation is the edges of rain areas might be only partially detected by PR because they are associated with low rain rates and intermittency (Kirstetter et al. 2012). The detectability issue is related to the sensitivity of PR and is thus not readily correctable with an update to the processing algorithm. However, it is noted that the mode of V7’s PDF$_c$ is shifted toward higher values than V6’s and is more consistent with the mode of the reference PDF$_c$. In examining the rain-rate distributions by volume, we see the modes of PDF$_y$ for both V6 and V7 are shifted toward lower rain rates compared to the reference’s mode ($\sim60$ mm h$^{-1}$), which agrees with the results found in Amitai et al. (2006, 2009). This has been attributed to high rainfall rates ($>10$ mm h$^{-1}$), which are likely underestimated by PR because of one or more of the following reasons: insufficient correction due to attenuation losses, NUBF effects, and inaccurate conversion from Z to R (Wolff and Fisher 2008). V7 presents a PDF$_y$ in better agreement with the reference than V6. The mode of the PDF$_y$ has increased from 18 to 25 mm h$^{-1}$, indicating a positive impact from the NUBF correction, correction of attenuation, and/or Z–R improvements over land.

b. Correlations and biases

Density-colored scatterplots of PR versus reference rainfall are presented for the two versions of 2A25 in Fig. 2. Better agreement with the reference (i.e., increases) in V7 is evident particularly for reference rainfall values greater than 30 mm h$^{-1}$. In addition, the relative underestimation from V6 at lighter rain rates (<1 mm h$^{-1}$) has now been mitigated in V7. We also provide common comparison metrics in Table 1. A rainy pixel is included in the statistics if both PR and the reference are nonzero. The V6 and V7 estimates are both subjected to the same discrepancies in spatiotemporal matching with the Q2 reference, which is a source for

| TABLE 1. Performance criteria values for PR estimates: mean, standard deviation, mean relative error (MRE), mean square error (MSE), and correlation (R) with respect to references. Only the reliable Q2 data are kept (see section 2b) for references. |
|-----------------|-----------------|-----------------|-----------------|
| PR 2A25 | Reference | Version 6 | Version 7 |
| Mean | 7.27 | 5.60 | 5.97 |
| Std dev | 13.76 | 8.26 | 9.8 |
| MRE | — | $-23\%$ | $-18\%$ |
| MSE | — | 112 | 102 |
| Correlation w.r.t. reference | — | 0.64 | 0.68 |
differences on a point-to-point comparison basis, so their relative differences can be directly attributed to algorithms themselves. PR shows lower mean values than the reference rainfall in both versions. However, regarding the average bias with the mean relative error (MRE), the V7 products present less difference (−18%) than the prior version (−23%). Note that Amitai et al. (2009) found an underestimation of less than 10%, which may be attributed to differences in the comparison methods, that is, including nonzero values in matched pairs and differences in data quality steps impact the MRE. From the mean square error (MSE) used to characterize the random estimation error, the V7 products present fewer discrepancies from the reference (102 mm² h⁻²) than the prior version (112 mm² h⁻²). This shows a positive impact of the new processing (i.e., recalibrated Z–R relationship over land and NUBF correction). The correlation coefficients between both versions of PR rainfall and Q2 reference estimates are moderate, but we note the correlation with V7 has improved slightly. Improving both the systematic part (MRE, 5% improvement) and the random part of error (MSE, 9% improvement) of the 2A25 products is noteworthy. Ciach et al. (2000) show that postprocessing optimization of a rainfall product relative to a reference can be done by improving the bias or the mean square error, but not both. PR, version 7, shows improvements in both bias and MSE, which can be only obtained by a more accurate processing of the radar signal relative to version 6. The correction of the largely underestimated rain rates in going from V6 to V7 (see Fig. 2) certainly contributes to this improvement.

c. Error models

The uncertainties associated with satellite estimates of rainfall include systematic errors as well as random
effects from several sources (Yang et al. 2006; Kirstetter et al. 2013). In a similar manner with Kirstetter et al. (2012), the departures of PR estimates from the Q2 reference values are analyzed in this section on a point-to-point basis. With the true rainfall being unknown, the residuals are defined as the difference between the reference rainfall and the satellite estimates:

$$R - R_{\text{ref}}.$$ 

Only pairs for which $R_{\text{ref}}$ and $R$ are both nonzero are considered in the calculations. The sets of $\varepsilon$ distributions are studied using the generalized additive models for location, scale, and shape (GAMLSS) technique (Rigby and Stasinopoulos 2001, 2005; Akantziliotou et al. 2002; Stasinopoulos and Rigby 2007). We consider $R_{\text{ref}}$ the main driving (explanatory) variable conditioning the departures of PR estimates from reference values, and we use the reverse Gumbel distribution $f(\varepsilon) = (1/\sigma)[(\varepsilon - \mu)/\sigma] - \exp[(\varepsilon - \mu)/\sigma]]$ to model the conditional residual distributions, where the location $\mu$ (mean of the residual population) is to be linked to systematic errors and $\sigma$ (the standard deviation) is representative of random errors.

For a given conditional distribution of the response variable $\varepsilon$, the conditional quantiles can be expressed as a function of $R_{\text{ref}}$. Figure 3 shows the residuals as a function of $R_{\text{ref}}$ as well as the fitted GAMLSS model for the two 2A25 versions. The conditional PDFs of residuals present a high conditional shift from the zero line and a high conditional spread. Note that for $R_{\text{ref}} \approx 50 \text{ mm h}^{-1}$, the model is quite undetermined because of the lack of observed residuals. Both 2A25 versions present a tendency to underestimate high rain rates relative to the reference (negative median of residuals); V6 underestimates $R_{\text{ref}} = 20 \text{ mm h}^{-1}$ with an occurrence of 80% and with a representative bias of $-7 \text{ mm h}^{-1}$, while V7 underestimates the same reference value with an occurrence of 75% and with a representative bias of $-6 \text{ mm h}^{-1}$. There is better agreement with V7, but the remaining bias is likely attributed to an inaccurate $Z$–$R$ relationship, NUBF effects, and/or insufficient correction of PR attenuation losses at heavier rain rates.

We consider the conditional median of the residuals to compare the systematic error component for V6 and V7 as well as the interquantile (90%–10%) value to assess the random part of the error. Figure 4 shows the conditional biases and random errors of both versions of 2A25 relative to the Q2 reference dataset. The underestimation with V6 and V7 over a large range of rain rates induces a global negative bias, which was evident in Table 1. The conditional biases of both versions relative to the reference are quite similar but with a slight improvement in V7. The random discrepancies increase consistently with $R_{\text{ref}}$ for both products. The random discrepancies are greater for V7 than V6, suggesting that other factors in addition to $R_{\text{ref}}$ could be considered to properly model the random error of V7 rain-rate estimates.

d. Nonuniform beam filling and rain types

To provide additional feedback to PR QPE algorithm developers, impacts of rain type classification and NUBF on rain rates are investigated. To assess the NUBF, we quantify the inhomogeneity of Q2 precipitation distribution within the PR footprint using $\sigma_{\text{footprint}}$. Figures 5a,b show the residuals as a function of the NUBF and are segregated according to the PR-based rainfall type classification. Q2 standard deviation values ($\sigma_{\text{footprint}}$) are greater for convective than for stratiform rainfall. In fact, convective rainfall generally presents higher rain rates and variability than stratiform rainfall, as expected.
The residuals distribution is also very distinct according to the rainfall types with more spread in the convective case.

Consistent with a theoretical study performed by Iguchi et al. (2009), increasing the NUBF in the PR beam results in increasing underestimation for the PR estimates relative to the reference rainfall, whatever the rainfall type. Figures 5c,d show the systematic and random parts of error. The error features are confirmed to be very distinct according to the rainfall types. The PR convective systematic biases present a shift toward higher values compared to stratiform biases. Below NUBF = 6 mm h\(^{-1}\) the convective bias is positive, which may be caused by the specific Z–R relationship used in the convective profiling component of the 2A25 algorithm. The random part of the error is consistently greater for convective rainfall. From version 6 to version 7, the systematic bias remains the same for stratiform but decreases significantly for convective rainfall. The random error decreases for stratiform and convective echoes, illustrating the positive impact of the NUBF correction in version 7. The impacts of the NUBF and rainfall type classification, however, remain significant and motivate ongoing and future research.

4. Conclusions

A 3-month dataset of gauge-adjusted, quality-filtered surface rainfall estimates from the NEXRAD-based Q2 has been used to compare and contrast PR-based 2A25...

Fig. 5. Density plots of residuals (PR, V7) vs NUBF estimates for (a) stratiform and (b) convective rainfall. The (c) systematic part and (d) random part of error are represented for stratiform (blue lines) and convective (red lines) types and for 2A25, V6 (dotted lines), and 2A25, V7 (solid lines).
rainfall estimates from the older V6 algorithm and the newly released version (V7), which will be the final version of the TRMM PR rainfall algorithm. A quantification of the uncertainty of these rainfall estimates will be quite useful to users of the data, including hydrologists, which is the principal aim of this study. V7 includes improvements in attenuation correction of the radar signal and a recalibrated $Z-R$ equation for use over land areas, and a correction for NUBF effects was reintroduced. The comparisons have been performed at the PR-pixel resolution over the lower CONUS using a framework proposed in Kirstetter et al. (2012). Our analyses indicate that the bias of the rain-rate estimates from V7 has been improved from a prior underestimation bias of $-23\%$ (from V6) to $-18\%$ relative to our reference. Moreover, this improvement in reducing bias is accompanied by no reduction in the correlation coefficient; simultaneous improvement in both error metrics is quite challenging and was found to be a result of simultaneously correcting overestimation at lighter rain rates ($<10 \text{ mm h}^{-1}$) and underestimation at high rain rates ($>30 \text{ mm h}^{-1}$). The former correction is most likely a result of the recalibration of the $Z-R$ equation over land, while the latter is likely a result from the NUBF correction; NUBF is known to cause underestimation at high rain rates (Iguchi et al. 2009). A statistical error model was developed for both versions of PR algorithms to separate conditional biases and random discrepancies as a function of reference rainfall rate. The PR residuals are confirmed to be quite large. Further research is needed to determine the relative contribution of aforementioned error factors and other potential sources of error like reference accuracy or time/space mismatching. To provide feedback to algorithm designers, a preliminary study on identifiable error sources in PR rainfall estimates like rainfall type classification and NUBF was conducted. It confirmed that the overestimation at lighter rain rates mainly comes from the convective $Z-R$ relationship and underestimation at high rain rates is related to NUBF effects. The positive impact from the recalibration of the $Z-R$ relationships and from the NUBF correction in version 7 is shown and quantified. Future work will evaluate and quantify the relative contributions of PR rainfall estimation errors linked to additional factors such as off-nadir angle, NUBF, attenuation, and influence of the underlying terrain. Efforts are continuing to improve the Q2 algorithm following the upgrade to the NEXRAD network with dual-polarization capability and to quantify the rain-rate uncertainty at hourly and daily time scales. These details will be useful to algorithm designers involved in the to-be-launched Global Precipitation Measurement (GPM) mission.

**Acknowledgments.** We are very much indebted to the team responsible for the NMQ/Q2 products, especially Carrie Langston. We want to thank two anonymous reviewers, whose comments were very useful in improving the manuscript. This work was funded by a postdoctoral grant from the NASA Global Precipitation Measurement mission Ground Validation Management.

**REFERENCES**


Lin, X., and A. Y. Hou, 2008: Evaluation of coincident passive microwave rainfall estimates using TRMM PR and ground


