Abstract

Characterization of the error associated to satellite rainfall estimates is a necessary component of deterministic and probabilistic frameworks involving space-born passive and active microwave measurements for applications ranging from water budget studies to forecasting natural hazards related to extreme rainfall events. We focus here on the error structure of NASA's Tropical Rainfall Measurement Mission (TRMM) Precipitation Radar (PR) quantitative precipitation estimation (QPE) at ground. The problem is addressed by comparison of PR QPEs with reference values derived from ground-based measurements using NOAA/NSSL ground radar-based National Mosaic and QPE system (NMQ/Q2). A preliminary investigation of this subject has been carried out at the PR estimation scale (instantaneous and 5 km) using a three-month data sample in the southern part of US. The primary contribution of this study is the presentation of the detailed steps required to derive a trustworthy reference rainfall dataset from Q2 at the PR pixel resolution. It relies on a bias correction and a radar quality index, both of which provide a basis to filter out the less trustworthy Q2 values. Several aspects of PR errors are revealed and quantified including sensitivity to the processing steps with the reference rainfall, comparisons of rainfall detectability and rainfall rate distributions, spatial representativeness of error, and separation of systematic biases and random errors. The methodology and framework developed herein applies more generally to rainfall rate estimates from other sensors onboard low-earth orbiting satellites such as microwave imagers and dual-wavelength radars such as with the Global Precipitation Measurement (GPM) mission.

Key words: satellite-based rain estimation, radar, QPE, conditional bias, random error

Toward a Framework for Systematic Error Modeling of NASA Spaceborne Radar with NOAA/NSSL Ground Radar-based National Mosaic QPE

Pierre-Emmanuel Kirstetter[1,2,4], Y. Hong[1,4], J.J. Gourley[2], S. Chen[1,4], Z. Flamig[2,3], J. Zhang[2], K. Howard[2], M. Schwaller[5], W. Peterson[6], E. Amitai[5,7]

[1]School of Civil Engineering and Environmental Sciences, University of Oklahoma

[2]NOAA/National Severe Storms Laboratory, Norman OK 73072

[3]Cooperative Institute for Mesoscale Meteorological Studies, Norman OK 73072

[4]Atmospheric Radar Research Center, National Weather Center, Norman OK 73072

[5]NASA Goddard Space Flight Center, Greenbelt, MD 20771

[6]NASA Wallops Flight Facility, Wallops Island, VA 23337

[7]Chapman University, Orange CA

submitted for publication in *Journal of Hydrometeorology*

14 October, 2011

Corresponding author:

Professor Yang Hong

National Weather Center, 120 David L. Boren Blvd,

Atmospheric Radar Research Center Rm. 4610, Norman OK 73072-7303, US

e-mail: yanghong@ou.edu; http://hydro.ou.edu

1. Introduction

Reliable quantitative information on the spatial distribution of rainfall is essential for hydrologic and climatic applications, which range from real-time flood forecasting to evaluation of regional and global atmospheric model simulations. Given their quasi-global coverage, satellite-based quantitative rainfall estimates are becoming widely used for such purposes. Converting satellite measurements into quantitative precipitation estimates poses challenges. The link between the observations and surface rain rates depends on the calibration and operating protocol of the instrument itself, the spatial heterogeneity of the rain fields (e.g., co-existence of convective and stratiform precipitation within a single instrumental field-of-view and vertical heterogeneity of rainfall), the indirect nature of the measurement, and the retrieval algorithm used. 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/), 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). The error characterization is needed for data assimilation and climate analysis (Stephens and Kummerow 2007) and more specifically over land in hydrological modelling of natural hazards and budgeting water resources (Grimes and Diop 2003; Lebel *et al.* 2009).

In this study, we focus primarily on the TRMM Precipitation Radar (PR) quantitative precipitation estimation (QPE) at ground. The methodology presented herein would equally apply to all satellite precipitation products, in particular those onboard low-earth orbiting satellites. The TRMM-PR is currently the only active instrument measuring rainfall from a satellite platform conjointly with a radiometer (TMI). PR rainfall estimates are often considered as a reference for TMI-based rainfall estimates (e.g., Yang *et al.* 2006; Wolff and

Fisher 2008). It impacts rain estimates from polar-orbiting passive microwave measurements and a number of satellite-based high-resolution precipitation products (Ebert *et al.* 2007; Bergès *et al.* 2009; Ushio *et al.* 2006). Given the variety of potential sources of error in PR-based QPE and the impact of correction algorithms, the only practical solution is to evaluate PR QPE with respect to an external, independent reference rainfall data set. The reference is derived from high-resolution ground validation measurements using NOAA/NSSL ground radar-based National Mosaic and QPE system (NMQ; Zhang *et al.* 2011). These products yield instantaneous rainfall rate products over vast regions including regions of the conterminous US covered by NEXRAD data. While a number of studies have investigated the quality of PR estimates in various regions of the world (e.g., Adeyama and Nakamura 2003; Wolff and Fisher 2008; 2009; Amitai *et al.* 2009; 2011), our aim is to perform a systematic and comprehensive evaluation for regions over the southern conterminous US (CONUS).

We will characterize errors in PR estimates at the pixel measurement scale in order to minimize additional uncertainties caused by resampling. Systematic and stochastic errors of PR estimates will be documented in terms of bias and spatial structure. 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. Rainfall estimates from low-earth orbiting satellites suffer from their poor temporal sampling (Wolff and Fisher 2008; 2009; Lin and Hou 2008). Hence representative samples of direct comparisons between instantaneous coincident measurements from ground and space are difficult to achieve without a sufficient number of overpasses. This study uses three months (March-May 2011) of satellite overpasses over the lower CONUS. The data are pixel-matched in both time and space, and statistics are provided for comparing reference rain intensities to satellite-based estimates.

where a_i denotes a satellite pixel and N is the number of pixels covering the domain of interest. The reference data $R_{ref}(A)$ used to evaluate the satellite estimates should spatially match the corresponding true rainfall averaged over the same area A.

a) original ground-based products

The NOAA/NSSL National Mosaic and Quantitative Precipitation Estimation system (NMQ/Q2) (http://nmq.ou.edu; Zhang et al. 2011) is a set of experimental radar products comprising high resolution (0.01°, 5 min) instantaneous rainfall rate mosaics available over the CONUS. The NMQ system combines information from all ground-based radars comprising the National Weather Service's Weather Surveillance Radar – 1988 Doppler (WSR-88D) network (NEXRAD) network, mosaics reflectivity data onto a common 3D grid, estimates surface rainfall accumulations and types to arrive at accurate ground-based estimates of rainfall (Zhang *et al.* 2005; Lakshmanan *et al.* 2007; Vasiloff *et al.* 2007; Kitzmiller *et al.* 2010). Figure 1 shows an example of the CONUS coverage of Q2 rainfall at 0725 UTC on 11 April 2011 highlighting several rainy systems associated with orography in the West and a wide frontal system in the central part of the domain.

At hourly time step Q2 adjusts radar estimates with automated rain gauge networks using a spatially variable bias multiplicative factor. A radar quality index (RQI) is produced at the (0.01°, 5min) resolution. While the true quality of the Q2 QPEs varies in space and time due to a number of complicating factors (e.g., measurements errors, non-precipitation echoes, uncertainties in Z–R relationships, variability in the vertical profile of reflectivity), the RQI represents the radar QPE uncertainty associated with VPRs (Zhang *et al.* 2011). The RQI field is composed of a static part relative to the radar beam sampling characteristics such as percent blockage, beam height and width and a dynamic part accounting for the freezing level height. The static part is illustrated in Fig. 1, where the reduced radar coverage

The quasi-instantaneous matching is performed at the scale of the PR measurement scale (4.5 x 4.5 km).

The PR data and steps required to refine the Q2 ground-based rainfall to arrive at the reference rainfall used for comparisons are presented in section 2. Section 3 assesses the ability of PR rain retrievals to represent the rainfall variability derived from the reference data in terms of rainfall detectability, sensitivity, and spatial structure. Section 4 provides an empirical error model of the PR estimates versus reference rainfall and segregates systematic and random error. The paper is closed with concluding remarks in section 5.

2. Data sources

One of the first challenges encountered is the lack of knowledge about the true averaged rainfall for the spatial domains considered. One wants to compare instantaneous satellite rainfall estimates $R(A)$ with reference rainfall $R_{ref}(A)$ for a spatial domain A (which may be a satellite mesh, watershed, etc.) to characterize the accuracy of the satellite QPEs. The true (and unknown) area-averaged rainfall accumulation, denoted $R_{true}(A)$ is written as:

$$R_{true}(A) = \frac{1}{A}\iint R(\underline{x})d\underline{x} \qquad (1)$$

where $\underline{x}$ is the location vector. The reference rainfall $R_{ref}(A)$ is a proxy of $R_{true}(A)$. The final products of the satellite data processing are gridded rainfall fields. Satellite QPEs may then be written as:

$$R(A) = \frac{1}{N}\sum_{i=1}^{N} R(a_i) \qquad (2)$$

difference of the temporal resolution between the hourly adjustment factors applied downscale to 5-min Q2 rainrates. Nevertheless, they provide the best possible reference at the scale of PR in terms of sampling conditions and unbiased estimates.

b) Q2-based reference rainfall

In the current study, all significant rain fields observed coincidentally by TRMM overpasses and the NEXRAD radar network from March to May 2011 are collected. The Q2 products closest in time to the TRMM satellite local overpass schedule time are used. To compute the reference rainfall, a block-Q2 rainfall pixel is computed to match each PR pixel in case of TRMM overpasses in a similar manner to Kirstetter et al. (2010, 2011).

Although the quantitative interpretation of the weather radar signal in terms of rainfall may be complex, radars enable a reliable evaluation of area-averaged rainfall estimates. The spatial variability of rainfall at small scales and the resolution difference between radar and PR footprint (as much as 2 orders of magnitude in area) may cause significant discrepancies in the statistical sampling properties and adds statistical noise in the comparison (see e.g. Ciach and Krajewski (1999) for a similar issue when comparing point-measurement raingauge to area-rainfall radar data). In order to estimate PR-pixel-averaged ground rainfall accumulation (and the associated sampling errors), a weighted mean estimator is considered to determine the reference rainfall $R_{ref}(A)$ over the PR footprint A from Q2 products. As the representativeness of the rainfall sampled by PR is related to the characteristics of the radar beam, the weighting function is given by the PR beam pattern inside a PR footprint. The reference rainfall is therefore:

$$R_{ref}(A) = \frac{1}{\sum_{i=1}^{n} \omega_i} \sum_{i=1}^{n} \omega_i Q2(a_i) \quad \text{with} \quad \omega_i = \int_{\theta_{mesh}(a_i)} f^2(\theta, \theta_0)\, d\theta \qquad (3)$$

in the western part of US results in lower RQI values. The dynamic part causes the RQI values to decrease in cool season months when the freezing level is lower and the radar samples the melting layer and the ice phase at closer range and to increase in the warm season when the freezing level is at higher altitudes. This is illustrated in Fig. 1 where the freezing level is lower behind a cold frontal system, which deteriorates the already limited coverage in the western part of the CONUS.

The original Q2 products utilized in this study are (i) the radar-only instantaneous rain-rate national mosaic updated every 5min, (ii) the radar-only rain-rate national mosaic at hourly time step, (iii) the hourly raingauge-corrected national mosaic product and (iv) the RQI. The primary Q2 product used for comparison with PR is the radar-only instantaneous rainrate mosaic. Current Q2 radar products do not include an instantaneous gauge-adjusted rainrate mosaic. For this study and similarly to Amitai *et al.* (2009; 2011), a second reference rainfall was derived from the bias-corrected Q2 product by using the hourly gauge-adjusted and the hourly radar-only products. Pixel-by-pixel hourly (multiplicative) adjustment factors are calculated and applied to the radar-only instantaneous product. Extreme adjustment factors (outside the [0.1-10] range) are discarded and no comparison is performed with PR for the corresponding Q2 values. Thus, the gauge-adjustment also serves as a data quality control procedure. A subsequent reference is derived from the bias-corrected Q2 product filtered using the RQI index. Only the rainrates associated with the best RQI values (i.e., equal to 1) were retained. This selection insures that only Q2 estimates presenting the best measurements conditions (i.e., no beam blockage and radar beam below the melting level of rainfall) are retained.

One should note these incremental improvements of the Q2 products may not screen out all possible errors in ground-based radar estimates. In particular, the gauge-adjustment may suffer from representativeness errors from scarce raingauge network density and from the

to the PR resolution preserves the statistical characteristics of the PR product including the total rainfall amount, the total rainy area, and the PDF shapes. All of these properties may therefore be compared to the reference at once.

Figure 1 shows an example of continuous mapping of the weighted mean estimator for the reference rainfall $R_{ref}(A)$. The estimator is a smoother of the original Q2 rain field. The maximum of the rainfall rate decreases from 145 mm h^{-1} to 130 mm h^{-1}. The total rainfall area increases, mainly at the edges of the rainfield. In order to avoid a contamination of the PR-reference comparison by the uncertainty on the ground reference, the reference pixels were segregated into "robust" ($R_{ref} > \sigma_{footprint}$) and "non robust" ($R_{ref} < \sigma_{footprint}$) estimators. This procedure illustrated in Fig. 1 filters out the reference values at the edges of the rain fields. Non-robust reference values are discarded for quantitative comparison. The robustness check is applied to the three Q2 products considered for reference (native Q2, bias corrected Q2, RQI+bias corrected Q2). As an example for the "RQI+bias corrected Q2" the averaged relative error (σ_{ref}/R_{ref}) of the reference decreases from 832% to 16%. The ratio of the mean error to the standard deviation of the reference ($\sigma_{ref}/\sigma(R_{ref})$) decreases slightly from 5.6% to 5.4%. This method of reference selection therefore increases the reliability and representativeness of the block-Q2 values that constitute our ground reference.

c) Precipitation Radar (PR) based rainfall

The PR measures reflectivity profiles at Ku band. Surface rain rates are estimated over the southern US up to a latitude of 37°N (Fig. 1). Artifacts such as contamination by surface backscatter, attenuation and extinction of the signal, brightband effects and accuracy of the Z-R relationship (Wolff and Fisher, 2008) must be accounted for. In the present study, the

where notations have been simplified for the sake of convenience. $Q2$ denotes the Q2 rain rate product for the mesh a_i. The value $R_{ref}(A)$ depends on the number n of Q2 meshes inside the PR footprint; the weights ω are derived from the two-way normalized power-gain function of the radar antenna f (assumed to be Gaussian) and the beamwidth θ_0; each ω_i is computed over the domain θ_{mesh} corresponding to the Q2 mesh a_i. It is assumed the PR footprint remains constant (circle of approximately 5km) whatever the radar beam off-nadir inclination angle. Additional research may be needed to take into account the deformation of the footprint with off-nadir angle (Takahashi et al, 2006).

Two weighted standard errors are computed with the reference rainfall. The first one is the weighted sample standard deviation, which represents the variability of the Q2 rainfall (at native resolution) inside the PR footprint:

$$\sigma_{footprint} = \sqrt{\frac{V_1}{V_1^2 - V_2}\sum_{i=1}^{n}\omega_i\left(Q2(a_i) - R_{ref}(A)\right)^2} \quad \text{with} \quad V_1 = \sum_{i=1}^{n}\omega_i \quad V_2 = \sum_{i=1}^{n}\omega_i^2 \qquad (4)$$

It is used to select the PR-reference pairs for which the $R_{ref}(A)$ is trustworthy. The second one is the standard deviation relative to the weighted mean $R_{ref}(A)$:

$$\sigma_{ref} = \sqrt{\frac{V_2}{V_1^2}\sum_{i=1}^{n}\left(Q2(a_i) - R_{ref}(A)\right)^2} \qquad (5)$$

It allows us to assess the $R_{ref}(A)$ estimation quality.

Matching PR and $R_{ref}(A)$ estimates only exist at locations where both the PR and ground radars have taken actual observations. This technique averages the minimum number of Q2 meshes needed to produce spatially coincident sample $R_{ref}(A)$ estimates. The advantages of the current technique over gridded approaches are that there is no interpolation, extrapolation, smoothing or oversampling of PR data. Bringing the Q2 product

surface rain rate at each PR footprint location is a standard TRMM product (2A25 v6) described in Iguchi *et al.* (2000). The scan geometry and sampling rate of the PR lead to footprints spaced approximately 4.3 km cross and along-track (5.1 after TRMM boost), over a 215-km-wide swath. The minimum theoretical detectable rain rate by the PR is fixed by its sensitivity and is about 17 dBZ, or ~0.5 mm h^{-1}.

d) Comparison samples

Several factors including rainfall intermittency, discrete temporal sampling of TRMM and censoring of reference values for required quality reduce the number of comparison samples for reference and PR estimates over the comparison period. Table 1 provides the number of these samples for the reference values, inclusive and exclusive of non-rainy pixels. The comparison sample sizes in Table 1 are primarily driven by the number of rain events and the overpass frequency of TRMM, then by the censoring of reference values. The quality control in the bias adjustment discarded 26% of original Q2 values and an additional 34% were filtered using RQI. Note that after two levels of processing and censoring, the comparison sample size for the "RQI+bias corrected Q2" remains significant at 393 347. This is credited to the large number of samples offered by the high-resolution, gridded Q2 product.

To assess the representativeness of our spatially and temporally limited samples, we compared the statistics of the reference rainfall resampled to the PR-pixel resolution with respect to the whole reference dataset (CONUS-wide below 38°N which don't necessary match a TRMM overpass). Figure 3 shows quantile-quantile plots between (i) the whole reference data set (x-axis) and (ii) the subset of pixels that matched to PR-pixel resolution for the different reference datasets. Table 2 provides values of the conditional mean and standard deviation. The PR-resampled reference rainfall distribution does not show a clear

deviation from the 1:1 degree line compared to the whole distribution. The reference distributions are fairly stable given the different censoring levels with the mean of the PR-resampled distributions being within 7% of the one for the whole dataset. We may therefore consider each reference dataset to be quite representative of the corresponding whole rainfall distribution.

Similarly we compared the different PR datasets to assess the impact of censoring on their representativeness. Figure 4 shows quantile-quantile plots between (i) the complete ("Native") PR data set (x-axis) and (ii) the censored subsets according to the "Bias corrected" and "Bias+RQI corrected" samples. Table 3 provides values of the conditional mean and standard deviation for each set. The different PR rainfall distributions do not show a clear deviation from the 1:1 degree line compared to the "Native" PR rainfall distribution. The means and standard deviations of the "Bias corrected"-censored and "Bias+RQI corrected"-censored distributions are less than 1% and around 10% higher respectively. We may therefore consider the representativeness of each PR dataset, following censoring steps, to be quite comparable to each other.

3. Rainfall data analysis

This section reviews the ability of PR rain retrievals to represent the rainfall variability derived from the Q2 data. First, contingency tables provide information on the reference rainfall reliability and on the influence of PR sensitivity to detect rainfall occurrence. The PDF of rainfall estimates provide in-depth information on the sensor's global ability to capture rain regimes given the influence of its sensitivity and the attenuation of the radar signal. Another feature to compare is the spatial structure of rainfall fields.

volume. Note the lowest values (less than 8%) are obtained with the "Bias+RQI corrected Q2" reference.

The impact of the reference rainfall on the contingency scores is shown in Figure 5. Contingency values are used to compute probability of detection (POD), false alarm rates (FAR), and critical success index (CSI). Scores are generally better for the robust reference than for the non-robust one. Within this category CSI shows a general increase with sequential Q2 data quality steps, while the FAR shows the lowest values with additional processing of the Q2 reference. A general convergence between the Q2 reference and PR estimates is therefore acknowledged as a function of the reference accuracy.

Considering that 80% of the reference rain rates that are not detected by the PR are lower than 0.3 $mm.h^{-1}$, the sensitivity of PR is close to this value. The misses are probably associated with high intermittency and/or the "rain/no rain" limits of rain fields. These features are missed by the PR because the rainrates are close to the detection threshold. This suggests that the PR can indeed capture the main rain regions but loses the weaker echoes (Schumacher and Houze 2000), probably due to its sensitivity.

b) Probability distributions by occurrence and rain volume

Hereafter, the PR rain estimates are the conditional ones (positive rainfall) coincident and collocated with nonzero reference estimates. Two PDFs for PR versus reference rainfall are computed and shown in Fig. 6: (i) the PDF by occurrence (PDF_c) and (ii) the PDF by rain volume (PDF_v) (Wolff and Fisher, 2009; Amitai *et al.*, 2009-2011). The PDF_c provides statistical information on the rain rate distribution and highlights the estimate's sensitivity as a function of rainrate; it is computed as a ratio between the number of the rainrates inside each bin and the total number of rainrates. The PDF_v represents the relative contribution of each rainrate bin to the total rainfall volume; it is computed as a ratio between the sum of the

a) Contingency tables

Table 4 shows the contingency tables for PR rain/no rain occurrence relative to the references with percentile of hits (H; both Q2 and PR detect rain), misses (M; PR does not detect rain while Q2 does), false alarms (F; PR detects rain while Q2 does not), and correct rejections (C; both Q2 and PR do not detect rain). The reference data are separated into three sub-samples: the non-robust set ($R_{ref} < \sigma_{footprint}$, see section 2b), the robust set ($R_{ref} > \sigma_{footprint}$) and the "whole" Q2 set. Reference null values are considered as robust. All coincident and collocated PR values are considered and sorted according to the reference samples. Table 5 provides the mean rainfall values according to the same contingency tables with PR on the left hand side of the "/" sign and the reference on the right hand side.

The false detections (M + F) of PR are mainly associated with the non-robust reference data; more than 80% for all "non-robust sets", while around 50% are improperly classified when using the "valid" reference data set. The Misses (M) are the main contributors to the false detection population (i.e., approximately 85% for the "whole" data set). These Misses of PR are coincident with low reference values (less than 0.15 mm h^{-1} for the non-robust set for all references; see Table 5). By comparison, the correct detections (H + C) of PR are mainly associated with the robust reference set from 45% to 52%. For the same robust reference sets, the Hits of PR are coincident with the higher reference values with mean rainfall rates more than 6 mm h^{-1}. One should note for all references that (i) the mean PR (F) values are significantly lower than the PR (H) values, and (ii) the mean reference (M) values are significantly lower than the mean reference (H) values. Finally both mean reference and PR values are higher for the robust Q2 set than for the non-robust Q2 set. Table 6 shows the discarded rain volumes in question; the Misses of PR represent less than 12% of the reference rainfall volume, while False Alarms represent less than 16% of the PR rainfall

simple function, we use a normalized variogram, which represents the spatial correlation of the rain field (Journel and Huijbregts 1978; Lebel *et al.* 1987; Kirstetter *et al.* 2010; 2011). An appropriate model is fit to the empirical normalized variogram. Among the set of classical models, the exponential model was found most suitable. It is expressed as:

$$\gamma(h) = C_0 + (C - C_0)\left(1 - \exp\left(-\frac{h}{d}\right)\right) \quad (6)$$

where the three parameters are the nugget (C_0), the sill (C) and the variogram range parameter (d). The exponential model reaches its sill asymptotically as $h \to \infty$. The "effective range" corresponds to the mean decorrelation distance of the estimates. It is the distance where the variogram reaches 95% of its maximum and corresponds to $3d$ for the exponential model. The nugget parameter can be used to describe a possible discontinuity of the variogram at the origin which may be due to (i) the process variability at scales poorly resolved by the observation system and/or (ii) measurement errors. In the following, these parameters are used to characterize the structure of rainfall.

Spatially normalized variograms of references and PR estimates are displayed in Fig. 7. Table 7 summarizes the parameters of these variograms. The variogram ranges of PR are quite similar to the three references' (approximately 18km). The nugget values, however, are more distinct. While it is ~ 32% for the Q2 references, it is significantly higher for PR (approximately 45% of the sill). These decorrelations of spatial structure at short interdistances suggest the resolution of the PR measurements may be limited when sampling the variability of small, disorganized rainfall structures associated with localized convection. The smaller reference nugget is an indication of the better sampling of the rain field by the reference rainfall, an issue previously discussed in section 2b. The comparatively higher nugget with PR may be caused by the rain intermittency, contamination by surface backscatter, attenuation of the signal, brightband effects or inaccuracy of the Z-R

rainrates inside each bin and the total sum of rainrates. It is therefore an important characteristic of the instantaneous products from the perspective of building merged rainfall accumulations; it enables a comparison of PDFs based on estimates derived from instruments characterized by different detection limits (in particular at weak intensities).

The rainrates of PR exhibit similar PDF_c for all references. Compared to references' PDF_c, PR tends to overestimate light rain rates (~[0.3-0.5] mm h^{-1}). But, PR demonstrates poor detection of the lightest rain rates (below ~0.3 mm h^{-1}) compared to the two bias corrected references. This is coherent with the concept of rain area "edges" that might be only partially detected by PR, resulting in misses associated with low rain rates (see previous section). PR PDF_c presents similar features with references for rain rates > ~1 mm h^{-1}. One may note the improved convergence between PR and reference rainfall PDF_c in the rain rates interval [0.5-1.0] mm h^{-1} with the sequential Q2 data quality steps.

Despite the low occurrence of relatively high rain rates (> 10 mm.h^{-1}), their contribution to the total rainfall volume is significant (greater than 60 %). As a consequence, the mode of PDF_v for PR is shifted toward lower rainrates (~18 mm h^{-1}) compared to the reference's mode (~60 mm h^{-1}), in agreement with the results found in Amitai *et al.* (2006, 2009). This is attributed to high rainfall rates (> 10 mm.h^{-1}), which are evidently underestimated by PR because insufficient correction due to attenuation losses as suggested by Wolff and Fisher (2008) for the 2A25 version 6.

c) Space structure of estimated rainfall fields

For hydrological applications, the total amount of water over a basin as well as the location and spatial correlation within the catchment might be important. It is therefore relevant to assess the ability of space-based estimates to retrieve the spatial structure of rainfall fields as seen by the reference. In order to describe the structure by a relatively

relationship. An interesting feature is that both sensors present a slightly decreasing nugget with the sequential Q2 data quality steps. This feature could be attributed to the censoring of the reference, which filters out complicated sampling situations for the ground-based radars.

4. Quantitative error modelling

a) Correlations and biases

Scatterplots of PR versus reference rainfall are presented for the three sets of Q2 reference in Fig. 8. Classical performance criteria of satellite-based rainfall estimation compared to reference values are listed in Table 8: correlation coefficient and mean relative error (MRE), expressed in percentage and defined as MRE = (PR_mean - Ref_mean)/Ref_mean. The comparisons between the PR and reference estimates are assessed on a point-to-point basis. A rainy pixel is included in the statistics if both PR and the reference are nonzero to emphasize the PR ability to quantify precipitation when it is raining. This is particularly significant given the significant misses of PR.

The two sensors present coherent mean and standard deviation values as long as the representativeness of the comparison samples are kept in mind. As expected, the means of the three PR sets are quite similar. In all cases the PR underestimates the reference mean values by ~17%. This is once again attributed to the significant underestimation of the higher rainrates in the 2A25-v6 products, presumably due to attenuation losses. The variations of the reference mean for the three sets explain in large part the variations in the apparent bias of PR relative to the reference. The native reference set is affected by (i) a global overestimation of rain rates, which could be due to the inaccuracy of the Z-R relationship and (ii) an underestimation of rain rates linked to partial beam blockage and vertical profile of reflectivity effects (i.e., overshooting above the melting layer by the radar beam far from

the radar). The gauge-based bias correction of the native Q2 product decreases the mean reference values, so the negative bias of PR is apparently improved. The additional RQI filtering removes the underestimation of Q2 at far range so the bias of PR is degraded. The reference shows higher standard deviation than the PR in coherence with the PDF features presented in section 3b.

The correlation coefficients between PR and Q2 reference estimates are moderate (around 0.6). One could note the best correlation between the two sensors is achieved with the "Bias+RQI corrected" reference. The differences between the two products on a point-to-point comparison basis can be attributed to sample volume discrepancies, timing and navigation mismatches and the uncertainties in the respective rainfall estimates. The significantly greater nugget in the PR variogram than in the reference variogram is also an indication of the greater level of noise in the PR rain field spatial structure, which may limit the correlation between the two series on a point-to-point comparison.

b) error model

The departures of PR estimates from the references are analyzed in this section on a point-to-point basis. 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.*, 2011). There is a fundamental issue in segregating the proportion of the scatter due to purely random error and the proportion due to conditional biases of the PR estimates that may be either positive or negative, producing additional scatter.

With the true rainfall being unknown, the residuals are defined as the difference $\varepsilon = (R - R_{ref})$ between the reference rainfall (R_{ref}) and the satellite estimates (R). Only pairs for which R_{ref} and R are both nonzero are considered in the calculations in order to emphasize the PR ability to quantify precipitation where it is raining. The sets of ε

distributions are studied using the generalized additive models for location, scale and shape (GAMLSS, Rigby and Stasinopoulos 2005) technique. As a preliminary step, R_{ref} is considered as the main driving (explanatory) variable conditioning the departures of PR estimates from references.

Generalized linear models for location, scale and shape aim at modeling the parameters of a response variable's distribution. Two main assumptions are made: (1) the response variable ε is a random variable following a known parametric distribution with density $f(\varepsilon \mid \mu, \sigma)$ conditional on the parameters (μ, σ); (2) the observations ε are mutually independent given the parameter vectors (μ, σ). Each parameter is modeled as a function of R_{ref} (the explanatory variable) using monotonic (linear/non-linear or smooth) link functions. More details are provided by Rigby and Stasinopoulos (2001, 2005), Akantziliotou, Rigby, and Stasinopoulos (2002) and Stasinopoulos and Rigby (2007). A wide variety of distributional forms are available, but for the sake of simplicity and to distinguish between systematic and random errors, a number of conditional densities with the first two moments (the location μ - mean, to be linked on to systematic errors - and the scale σ - standard deviation representative of random errors) as parameters are considered here. For a given conditional distribution of the response variable, the conditional quantiles can be expressed as a function of the location and scale. GAMLSS is best fitted using the package gamlss in R (Stasinopoulos and Rigby 2007). The rainfall trends for each parameter are fitted using locally weighted scatterplot smoothing (loess), which are more flexible than polynomials or fractional polynomials for modelling complex nonlinear relationships. It is a polynomial curve determined by R_{ref}, which is fitted locally by weighted polynomial regression, giving more weight to points near the point whose response is being estimated and less weight to points further away (see Cleveland, Grosse, and Shyu 1993).

Several two-parameter density functions (log normal, normal, reverse gumbel, logistic, gamma, etc.) have been tested to fit the data. The distributions of residuals (not shown here) were generally found to be unimodal and asymmetric. The goodness-of-fit on the whole dataset has been checked by investigating the Akaike information criteria (AIC) for the each semi-parametric density fits. The reverse Gumbel distribution $\left(f(\varepsilon)\frac{1}{\sigma}\left[-\left(\frac{\varepsilon-\mu}{\sigma}\right)-\exp\left\{-\left(\frac{\varepsilon-\mu}{\sigma}\right)\right\}\right]\right)$, where μ is the mean and σ the standard deviation of the residual population) was found to be the most appropriate. Figure 9 shows the residuals as a function of R_{ref} as well as the fitted GAM model for PR in the representative case of the "Bias+RQI corrected" reference. The conditional PDF of residuals ε present a high conditional shift versus the 0 line and a high conditional spread. Note that for R_{ref} > ~50mm.h^{-1}, the model is quite undetermined because of the lack of observed residuals. All models show that PR present a tendency to overestimate light rain rates (the median of residuals is positive) and underestimate higher rain rates (negative median of residuals): e.g. PR underestimates R_{ref} = 20 mm.h^{-1} rain rates with an occurrence of 70% and with a representative bias of -7 mm.h^{-1} and underestimates R_{ref} = 40 mm.h^{-1} with an occurrence of 92% and with a representative bias of -24 mm.h^{-1}. This is likely to be once again due to the insufficient correction of PR attenuation for heavier rain rates.

In case of a non-symmetric density for residuals or in case of extreme values, the median is preferred to the expectation for a better representativeness of the systematic component of the residuals. The systematic error component (i.e. conditional bias) is therefore described by the conditional median of these distributions. For the same reason we consider the interquantile (q90-q10) value to assess the random part of the error. It is computed after having applied the error separation variance correction to the conditional standard deviation

conditional bias is less significant), which could be seen as a sign of a better convergence between PR estimates and this Q2 reference. This is confirmed when considering the random part of error. The "Bias+RQI corrected" curve shows the lowest random errors up to R_{ref} = 4 mm h^{-1} (more than 65% of the reference rain rates are under this value). The random error increases consistently with R_{ref}. It is systematically higher for the "Bias corrected" than for the "Native" reference, a result consistent when applying a bias correction (Ciach *et al.* 2000). It represents a significant part of error, suggesting that other factors than R_{ref} could be considered to evaluate the error of PR rain rate estimates at ground.

5. Conclusions

In preparation for NASA's future Global Precipitation Measurement (GPM) mission, a three-month data sample of TRMM-PR-based rainfall products have been compared to surface rainfall derived from Q2 over the lower conterminous US. The major advantage of the Q2 ground-based reference dataset is its resolution in both time and space commensurate with rainfall estimates derived from sensors onboard low-earth-orbiting satellites. The comparisons have been performed at the PR-pixel resolution. A framework is proposed herein to address methodological issues so as to provide a preliminary version of an error model for satellite QPEs. The error model is empirically derived and is thus prone to be specific to the dataset considered and the PR/Q2 data processing implemented. However, the results show similarities with previous rainfall comparisons over West Africa (Benin and Niger) and thus give credence to the developed framework (Kirstetter *et al.* 2011). Results from the error model presented herein provide insights into the most significant

σ extracted from the GAM model. The error separation variance concept (Ciach and Krajewski, 1999; Teo and Grimes, 2007; Kirstetter *et al.* 2010) makes it possible to evaluate the variance of the PR with respect to the true unknown rainfall. We assume the errors on the reference rainfall and on the PR estimates to be uncorrelated. Introducing the true rainfall R_{true} in the expression of the variance of the residuals between the PR and reference values leads to (see Kirstetter *et al.* 2010 for details):

$$\sigma\left(R^{*}-R_{true}\right) \approx \sqrt{\mathrm{var}(\varepsilon)-\sigma_{ref}^{2}} \qquad (7)$$

Fortunately as can be seen in Figure 10, the reference estimation standard deviations are lower than the standard deviations of the PR–reference residuals, indicating the reference values to be comparatively reliable to evaluate PR. The standard deviation of the PR residuals with respect to the true rainfall is significantly reduced compared with the PR–reference residual standard deviation. One may note the standard deviations increase up to a reference value (~50 mm.h^{-1}) beyond which we believe sampling issues lead to a stabilisation or a decrease of the standard deviations. We therefore apply the modelling up to this limit only. As ~98% of the reference values are under his limit, this choice won't lead to any significant lack of representativeness.

Figure 10 shows the conditional biases and random errors of PR relative to the three Q2 references. The global bias (see previous section and Table 8) of PR results from a balance between overestimation of light rainrates and underestimation of high rainrates. The underestimation is more frequent, inducing a global negative bias. The conditional biases of PR relative to the references are quite similar. Note the "Bias corrected" conditional bias is shifted to the right compared to the "Native" one, so overestimation of light rainrates is more significant and the underestimation of higher rainrates less pronounced, consistently with the reduced negative global bias for this specific reference (see Table 8). Note also the negative slope of the "Bias+RQI corrected" conditional bias is lower than for the two other (the

characteristics of PR rainfall retrieval errors that need to be taken into account when such data are used in applications.

A consistent result noted throughout each analysis was the increased consistency between PR and the Q2 reference following sequential data quality control steps including bias correction using rain gauges and filtering using the radar quality index (RQI) product. This finding, alone, highlights the importance of matching the scales and refining the accuracy of the reference dataset as much as possible before reaching meaningful conclusions about the PR accuracy.

Different error sources were identified and quantified for PR rainrate estimates. The most significant error is most likely due to the attenuation of the PR radar signal. It is not yet known if this error is due to inadequate correction for attenuation losses or complete loss of the signal. Segregating rain from no-rain boundaries is also a driving contributor to the PR rainrate errors, probably linked to the lack of sensitivity in the most inhomogeneous and light parts of the edges of rainy regions. Nevertheless, the variogram analysis showed that the PR adequately represents the spatial structure of the rain fields. The scatterplots revealed PR-estimated rainrates are only moderately correlated (Pearson correlation coefficient of 0.6) to the best reference rainfall on a point-to-point basis.

The statistical model developed here quantifies the relation between instantaneous PR rainfall and the corresponding reference rainfall. It consists of a deterministic additive function and a random uncertainty component, both conditioned on given reference values. The contribution of systematic PR errors is confirmed to be quite large due to the aforementioned signal attenuation issue.

In terms of perspectives, the relative contributions of errors linked to rainfall type and off-nadir angle need to be evaluated, as well as influence of the underlying terrain. The same framework and reference rainfall datasets can be readily applied to rainfall retrievals from

other sensors onboard low-earth-orbiting satellites (i.e., TMI, AMSR-E, SSMI, MADRAS). This framework will also be applied to GPM rainfall estimates following its launch in 2013. Another important issue to study is how the various error sources in PR, which is often used as a calibrator, propagate when merging with geostationary infrared data for a number of satellite-based, high-resolution precipitation products.

Acknowlegments

We are very much indebted to the team responsible for the NMQ/Q2 products, especially Carrie Langston. This work was funded by a post-doctoral grant from the NASA Global Precipitation Measurement mission Ground Validation Management.

References

Adeyewa, Z.D., and K. Nakamura, 2003: Validation of TRMM radar rainfall data over major climatic regions in Africa. *J. Appl. Meteor.*, **42**, 331–347.

Akantziliotou, K., R.A. Rigby, and D.M. Stasinopoulos, 2002: The R implementation of Generalized Additive Models for Location, Scale and Shape. In: Stasinopoulos, M. and Touloumi, G. (eds.), Statistical modelling in Society. Proceedings, *17th International Workshop on statistical modelling*, Chania, Greece, 75–83.

Amitai, E., X. Llort, and D. Sempere-Torres, 2006: Opportunities and challenges for evaluating precipitation estimates during GPM mission. *Meteorol. Zeitschrift.*, **15**(5), 551-557.

Amitai, E., X. Llort, and D. Sempere-Torres, 2009: Comparison of TRMM Radar Rainfall Estimates with NOAA Next-Generation QPE, *Journal of the Meteorological Society of Japan*, **87**A, 109-118.

Amitai, E., W. Petersen, X. Llort, and S. Vasiloff, 2011: Multi-platform comparisons of rain intensity for extreme precipitation events, *IEEE Transactions on Geoscience and Remote Sensing*, in press.

Berges, J.C., F. Chopin, I. Jobard, and R. Roca, 2010: EPSAT-SG: a satellite method for precipitation estimation; its concept and implementation for AMMA experiment. *Annales Geophysicae*, **28**, 289-308.

Ciach, G.J., and W.F. Krajewski, 1999: On the estimation of rainfall error variance. *Advances in Water Resources*, **2**, 585-595.

Ciach J.G., M.L. Morrissey, and W.F. Krajewski, 2000: Conditional bias in radar rainfall estimation. *Journal of Applied Meteorology*, **39**, 1941-1946.

Cleveland, W.S., E. Grosse, and M. Shyu, 1993: Local Regression Models. In: Chambers, J. and Hastie, T. (eds.), Statistical Modelling in Society, 309–376. Chapman and Hall: New York.

Ebert, E., 2007: Methods for verifying satellite precipitation estimates. In: *Measuring Precipitation from space: EURAINSAT and the future.* Levizzani, V., P. Bauer, and F.J. Turk, eds., Springer, 345-356.

Grimes, D.I.F., and M. Diop, 2003: Satellite-based rainfall estimation for river flow forecasting in Africa. I: Rainfall estimates and hydrological forecasts. *Hydrolog. Sci. J.*, **48**, 567–584.

Iguchi, T., T. Kozu, R. Meneghini, J. Awaka, and K. Okamoto, 2000: Rain-profiling algorithm for the TRMM precipitation radar. *J. Appl. Meteor.*, **39**, 2038–2052.

Journel, A., and C. Huijbregts, 1978: *Mining Geostatistics*. London Academic Press, 131 pp.

Kirstetter, P.E., G. Delrieu, B. Boudevillain, and C. Obled, 2010: Toward an error model for radar quantitative precipitation estimation in the Cévennes-Vivarais region, France. *Journal of Hydrology*, **394**(1-2), 28-41. doi:10.1016/j.jhydrol.2010.01.009

Kirstetter, P.E., N. Viltard, and M. Gosset, 2011: Building an error model for instantaneous satellite rainfall estimates. Submitted (accepted) *to Quarterly Journal of the Royal Meteorological Society.*

Lebel, T., G. Bastin, C. Obled, and J.D. Creutin, 1987: On the Accuracy of Areal Rainfall Estimation: A Case Study. *Water Resour. Res.*, **23**(11), 2123–2134.

Lebel, T., C. Cappelaere, S. Galle, N. Hanan, L. Kergoat, S. Levis, B. Vieux, L. Descroix, M. Gosset, and E. Mougin, 2009: AMMA-CATCH studies in the Sahelian region of West-Africa: an overview. *Journal of Hydrology*, **375**(1-2), 3-13. doi:10.1016/j.jhydrol.2009.03.020

Lin, X., and A.Y. Hou, 2008: Evaluation of coincident passive microwave rainfall estimates using TRMM PR and ground measurements as references. *J. Appl. Meteor. Climatol.*, **47**, 3170–3187.

Rigby, R.A., and D.M. Stasinopoulos, 2001: The GAMLSS project: a flexible approach to statistical modelling. In: Klein, B. and Korsholm, L. (eds.), New Trends in Statistical Modeling. *Proceedings of the 16th International Workshop on Statistical Modelling*, Odense, Denmark, 249–256.

Rigby, R.A., and D.M. Stasinopoulos, 2005: Generalized additive models for location, scale and shape (with discussion). *Appl. Statist.*, **54**(3), pp 507-554.

Sapiano, M.R.P., and P.A. Arkin, 2009: An Intercomparison and Validation of High-Resolution Satellite Precipitation Estimates with 3-Hourly Gauge Data. *J. Hydrometeor.*, **10**, 149–166.

Stephens, G.L., and C.D. Kummerow, 2007: The Remote Sensing of Clouds and Precipitation from Space: A Review. *Journal of the Atmospheric Sciences* **64**(11), 3742-3765.

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

Stasinopoulos, D.M., and R.A. Rigby,. 2007: Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software*, **10**(2), 1-64.

Wolff, D.B., and B.L. Fisher, 2008: Comparisons of instantaneous TRMM ground validation and satellite rain-rate estimates at different spatial scales. *J. Appl. Meteor. Climatol.*, **47**, 2215–2237.

Takahashi, N., H. Hanado and T. Iguchi, 2006: Estimation of path-integrated attenuation and its nonuniformity from TRMM/PR range profile data. *IEEE Transactions on Geoscience and Remote Sensing*, **44**(11), 3276-3283.

Table captions

Table 1. Comparison samples for different reference datasets.

Table 2. Conditional mean and standard deviation of "whole" and "PR-resampled" references datasets.

Table 3. Conditional mean and standard deviation of PR estimates for different references.

Table 4. Contingency table for PR relative to the three references. The results are provided for robust / non robust reference data according to a criterion based on the variability of the Q2 rainfall (at native resolution) inside the PR footprint ($R_{ref} > \sigma_{footprint}$).

Table 5. Mean rainfall values associated to the contingency table for PR / references.

Table 6. Discarded rain volumes from PR due to misses relative to references and rain volume implied in the false alarms relative to robust references.

Table 7. Parameters of the normalized variograms (exponential model) for references and PR. The "effective range" values are indicated. The nugget is expressed as a percentage of the normalised sill.

Table 8. Performance criteria values for PR estimates: mean, standard deviation, mean relative error (MRE) and correlation (R) with respect to references. Only the reliable Q2 data are kept (see section 2.b) for references.

Turk, F. J., P. Arkin, E. E. Ebert, and M. R. P. Sapiano, 2008: Evaluating high resolution precipitation products. *Bull. Amer. Meteor. Soc.*, **89**, 1911-1916.

Wolff, D.B., and B.L. Fisher, 2009: Assessing the Relative Performance of Microwave-Based Satellite Rain-Rate Retrievals Using TRMM Ground Validation Data. *Journal of Applied Meteorology and Climatology*, **48**(6), 1069-1099.

Yang, S., W.S. Olson, J.J. Wang, T.L. Bell, E.A. Smith, and C.D. Kummerow, 2006: Precipitation and Latent Heating Distributions from Satellite Passive Microwave Radiometry. Part II: Evaluation of Estimates Using Independent Data. *Journal of Applied Meteorology and Climatology*, **45**(5), 721-739.

Zeweldi, D.A., and M. Gebremichael, 2009: Sub-daily scale validation of satellite-based high-resolution rainfall products. *Atmospheric Research*, **92**(4), 427-433. doi:10.1016/j.atmosres.2009.01.001.

Zhang, J., Y. QI, K. Howard, C. Langston, and B. Kaney, 2011: Radar QPE Uncertainties in a National Network. *Proceedings of the Weather Radar and Hydrology Symposium*, Exeter.

Figure captions

Figure 1: Map of CONUS area with NMQ/Q2 instantaneous rainrates at 0725 UTC on 07 April 2011. The red area shows the good quality radar coverage corresponding to Radar Quality Index equal to 1. The shaded area is not sampled by the TRMM-PR.

Figure 2: Maps of instantaneous rainrates at 0725 UTC on 07 April 2011: the NMQ-Q2 product (top left), the equivalent reference rainfall R_{ref} (top right), the "robust" reference set (bottom left) and the "non robust" reference set (bottom right).

Figure 3: Quantile-quantile plots for reference "PR-sampled" and "Whole" rainfall distribution comparison, for native (a), bias corrected (b) and Bias+RQI corrected (c) references. The positions of 10, 50 and 95 percentiles are showed for each distribution.

Figure 4: Quantile-quantile plots for reference "PR-sampled" and "Whole" rainfall distribution comparison, for native (a), bias corrected (b) and Bias+RQI corrected (c) references. The positions of 10, 50 and 95 percentiles are showed for each distribution.

Figure 5: Critical success index (CSI), probability of detection (POD) and false alarm rates (FAR), and for the three references and partitioned as a function of robustness.

Figure 6: Probability distributions of rain rates the PR rainfall and for the native (a), bias corrected (b) and Bias+RQI corrected (c) reference rainfall. The solid and dashed-dotted lines represent the distribution by volume PDF_v and the distribution by occurrence PDF_c respectively, while the grey and black lines represent the distributions for references and PR respectively. Note that the x-axis is in log-scale.

Figure 7: Spatial variograms for reference (left) and PR (right) for the native (top), bias corrected (middle) and Bias+RQI corrected (bottom) reference. The empirical variograms are plotted with crosses, and the models fitted are represented by the thick black lines.

Figure 8: Scattergraphs of PR versus native (a), bias corrected (b) and Bias+RQI corrected(c) reference rainfall (mm h^{-1}). The first bisectors (solid lines) are displayed.

Figure 9: PR residuals represented versus "Bias+RQI corrected" reference (left) and the GAM model fitted is represented by [5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95] conditional quantile lines (right).

Figure 10: Standard deviation of PR-reference residuals (dashed line), estimated standard deviation of the reference rainfall (dotted line) and standard deviation of PR–true rainfall residuals (solid line) as functions of the "Bias+RQI corrected" reference. The vertical line (50 mm.h^{-1}) indicates the limit of the good sampling conditions.

Figure 11: Conditional bias (median) of residuals (left) and conditional random error (interquantile 90%-10%) of residuals (right) for PR as a function of "Native" (dotted line), "Bias corrected" (dotted line) and "Bias+RQI corrected" references.

TABLE 1. Comparison samples for different reference datasets.

	Including non-rainy (0 mm h^{-1})	Rainy only
Native Q2	35 349 900	984 598
Bias corrected Q2	35 342 653	723 495
Bias+RQI corrected Q2	35 342 653	393 347

TABLE 2. Conditional mean and standard deviation of "whole" and "PR-resampled" references datasets.

	PR resampled dataset		Whole reference data set	
	Mean	St. dev.	Mean	St. dev.
Native Q2	1.64	6.55	1.57	6.30
Bias corrected Q2	2.00	6.99	1.95	6.84
Bias + RQI correctd Q2	1.98	7.26	2.13	7.37

TABLE 3. Conditional mean and standard deviation of PR estimates for different references.

	Mean	Standard deviation
Native Q2	4.21	6.91
Bias corrected Q2	4.24	6.94
Bias+RQI corrected Q2	4.65	7.47

TABLE 4. Contingency table for PR relative to the three references. The results are provided for robust / non robust reference data according to a criterion based on the variability of the Q2 rainfall (at native resolution) inside the PR footprint ($R_{ref} > \sigma_{footprint}$).

PR Estimates		Native Q2		
		> 0.	= 0.	Σ estimates
> 0.	whole set	26 %	6 %	320715
	robust	45 %	15 %	251175
	non robust	12 %	0 %	69540
=0.	whole set	67 %	0. %	664220
	robust	40 %	0. %	167986
	non robust	88 %	0. %	496234
Σ reference	whole set	921758	63177	984935
	robust	355984	63177	419161
	non robust	545774	0	565774

TABLE 5. Mean rainfall values associated to the contingency table for PR / references.

PR Estimates		Native Q2	
		> 0.	= 0.
> 0.	whole set	4.65 / 5.06	2.41 / 0.00
	robust	5.07 / 6.21	2.41 / 0.00
	non robust	3.51 / 1.97	-
=0.	whole set	0.00 / 0.31	-
	robust	0.00 / 0.92	
	non robust	0.00 / 0.10	

PR Estimates		Bias corrected Q2	
		> 0.	= 0.
> 0.	whole set	4.83 / 4.71	2.28 / 0.00
	robust	5.38 / 6.07	2.28 / 0.00
	non robust	3.61 / 1.73	-
=0.	whole set	0.00 / 0.38	-
	robust	0.00 / 1.05	
	non robust	0.00 / 0.14	

PR Estimates		Bias corrected Q2		
		> 0.	= 0.	Σ estimates
> 0.	whole set	34 %	10 %	316941
	robust	48 %	21 %	240223
	non robust	20 %	0 %	76260
=0.	whole set	56 %	0. %	406610
	robust	31 %	0. %	108218
	non robust	80 %	0. %	298392
Σ reference	whole set	650567	72984	723551
	robust	275915	72526	348441
	non robust	374652	0	374652

PR Estimates		Bias+RQI corrected Q2		
		> 0.	= 0.	Σ estimates
> 0.	whole set	32 %	3 %	137978
	robust	52 %	7 %	94533
	non robust	18 %	0 %	43178
=0.	whole set	65 %	0. %	254921
	robust	41 %	0. %	64620
	non robust	81 %	0. %	190301
Σ reference	whole set	380087	12812	392899
	robust	146608	12545	159153
	non robust	233479	0	233479

PR Estimates		Bias+RQI corrected Q2	
		> 0.	= 0.
> 0.	whole set	4.92 / 5.45	2.07 / 0.00
	robust	5.60 / 7.27	2.07 / 0.00
	non robust	3.63 / 2.01	-
=0.	whole set	0.00 / 0.28	-
	robust	0.00 / 0.74	
	non robust	0.00 / 0.12	

TABLE 6. Discarded rain volumes from PR due to misses relative to references and rain volume implied in the false alarms relative to robust references.

	Native Q2	Bias corrected Q2	Bias + RQI correctd Q2
Excess of rain volume implied in False Alarms	13.70 %	15.45 %	5.36 %
Discarded rain volume due to Misses	11.70 %	10.07 %	7.46 %

TABLE 7. Parameters of the normalized variograms (exponential model) for references and PR. The "effective range" values are indicated. The nugget is expressed as a percentage of the normalised sill.

	Reference		PR	
	Nugget (% sill)	Range (km)	Nugget (% sill)	Range (km)
Native	34 %	17	48 %	16
Bias corrected	32 %	15	48 %	16
Bias+RQI corrected	31.5 %	15.5	45 %	15

TABLE 8. Performance criteria values for PR estimates: mean, standard deviation, mean relative error (MRE) and correlation (R) with respect to references. Only the reliable Q2 data are kept (see section 2.b) for references.

PR	Native		Bias corrected		Bias+RQI corrected	
	Reference	PR	Reference	PR	Reference	PR
Mean	6.20	5.07	6.07	5.38	7.27	5.6
standard deviation	12.53	7.80	12.04	8.03	13.76	8.26
MRE / reference (%)	-	-18 %	-	-11 %	-	-23 %
Correlation / reference	-	0.61	-	0.6	-	0.64

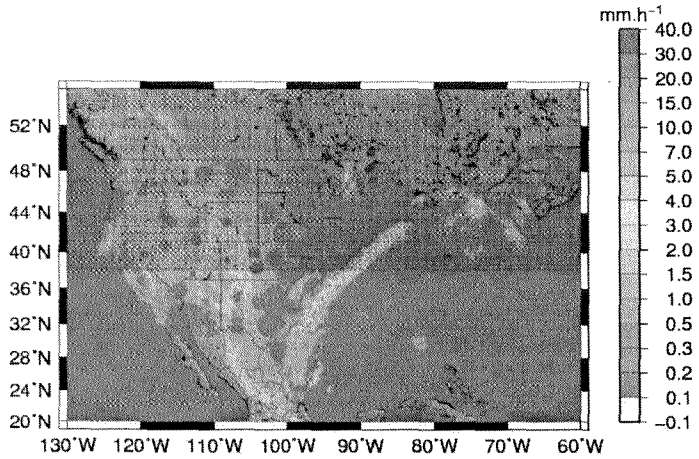

Figure 1: Map of CONUS area with NMQ/Q2 instantaneous rainrates at 0725 UTC on 07 April 2011. The red area shows the good quality radar coverage corresponding to Radar Quality Index equal to 1. The shaded area is not sampled by the TRMM-PR.

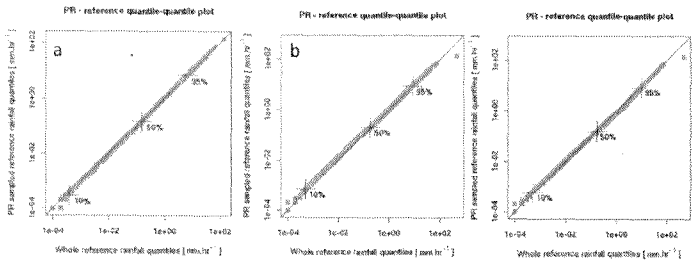

Figure 3: Quantile-quantile plots for reference "PR-sampled" and "Whole" rainfall distribution comparison, for native (a), bias corrected (b) and Bias+RQI corrected (c) references. The positions of 10, 50 and 95 percentiles are showed for each distribution.

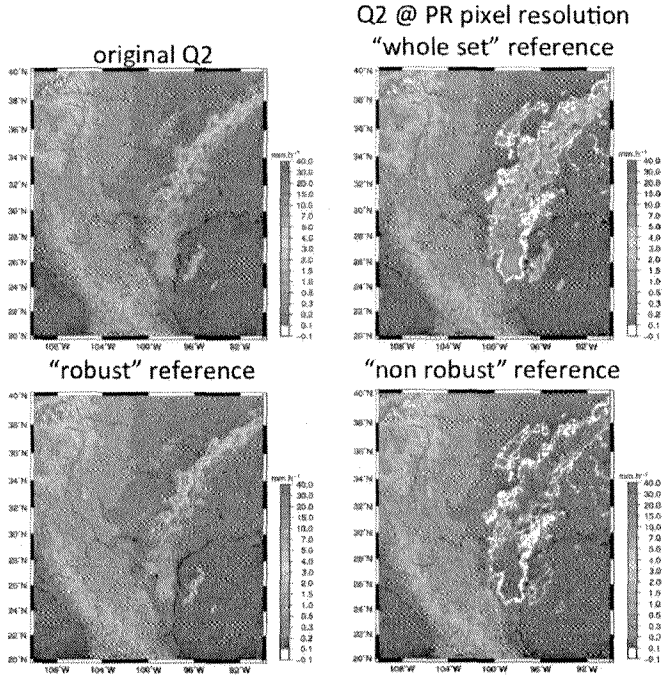

Figure 2: Maps of instantaneous rainrates at 0725 UTC on 07 April 2011: the NMQ-Q2 product (top left), the equivalent reference rainfall R_{ref} (top right), the "robust" reference set (bottom left) and the "non robust" reference set (bottom right).

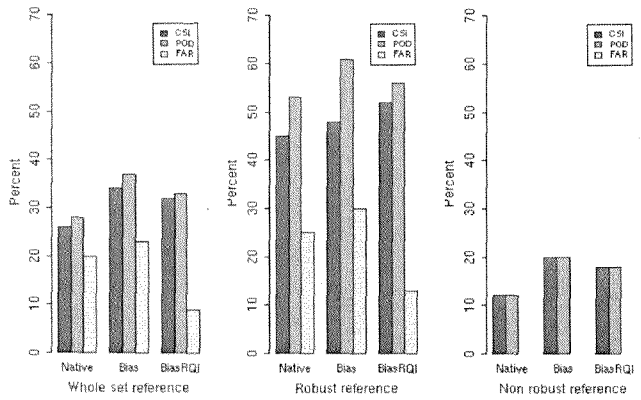

Figure 5: Critical success index (CSI), probability of detection (POD) and false alarm rates (FAR), and for the three references and partitioned as a function of robustness.

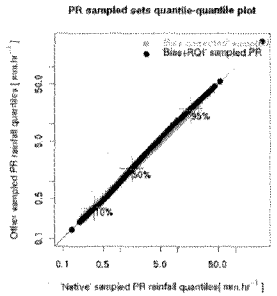

Figure 4: Quantile-quantile plots for reference "PR-sampled" and "Whole" rainfall distribution comparison, for native (a), bias corrected (b) and Bias+RQI corrected (c) references. The positions of 10, 50 and 95 percentiles are showed for each distribution.

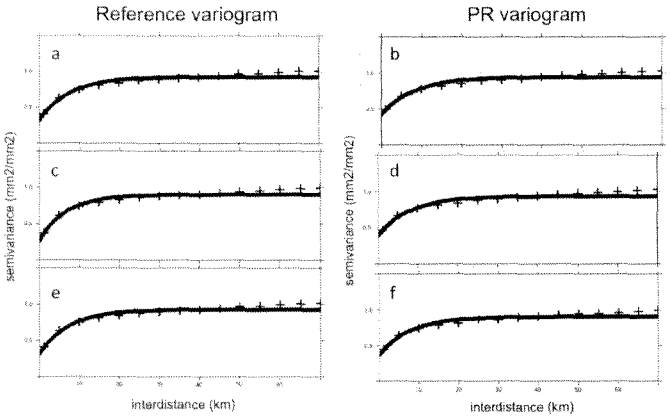

Figure 7: Spatial variograms for reference (left) and PR (right) for the native (top), bias corrected (middle) and Bias+RQI corrected (bottom) reference. The empirical variograms are plotted with crosses, and the models fitted are represented by the thick black lines.

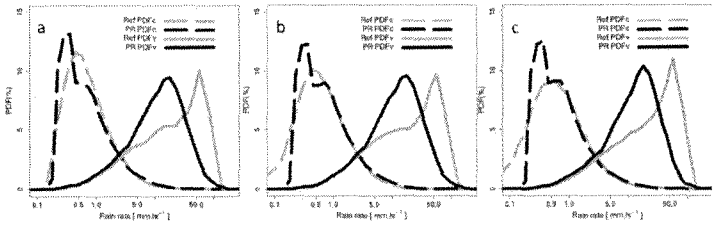

Figure 6: Probability distributions of rain rates the PR rainfall and for the native (a), bias corrected (b) and Bias+RQI corrected (c) reference rainfall. The solid and dashed-dotted lines represent the distribution by volume PDF_v and the distribution by occurrence PDF_c respectively, while the grey and black lines represent the distributions for references and PR respectively. Note that the x-axis is in log-scale.

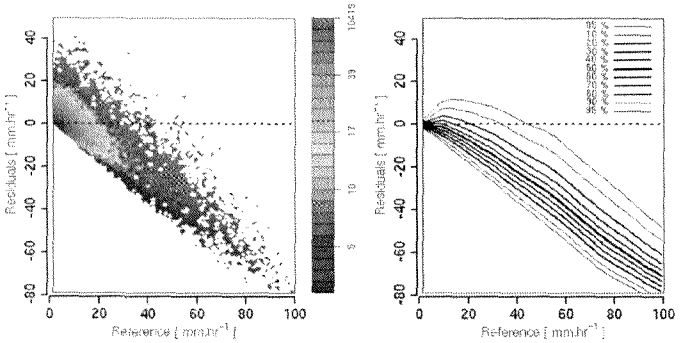

Figure 9: PR residuals represented versus "Bias+RQI corrected" reference (left) and the GAM model fitted is represented by [5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95] conditional quantile lines (right).

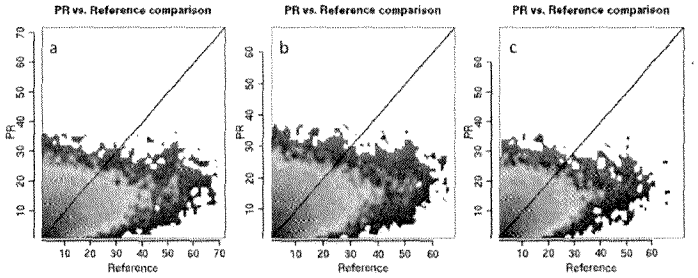

Figure 8: Scatterplots of PR versus native (a), bias corrected (b) and Bias+RQI corrected(c) reference rainfall ($mm.h^{-1}$). The first bisectors (solid lines) are displayed.

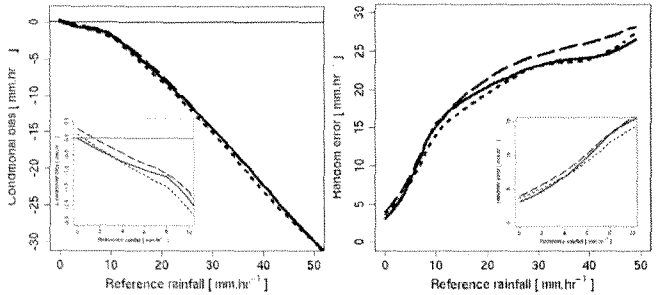

Figure 11: Conditional bias (median) of residuals (left) and conditional random error (interquantile 90%-10%) of residuals (right) for PR as a function of "Native" (dotted line), "Bias corrected" (dotted line) and "Bias+RQI corrected" references.

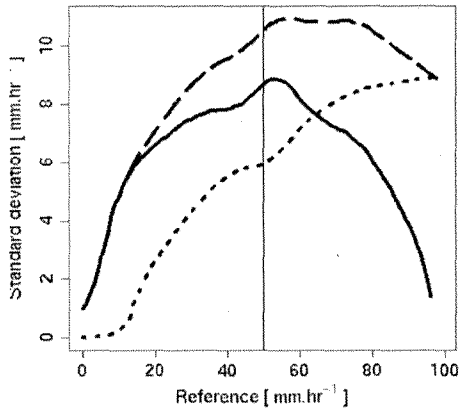

Figure 10: Standard deviation of PR-reference residuals (dashed line), estimated standard deviation of the reference rainfall (dotted line) and standard deviation of PR–true rainfall residuals (solid line) as functions of the "Bias+RQI corrected" reference. The vertical line (50 $mm.h^{-1}$) indicates the limit of the good sampling conditions.