Enhancing the consistency of spaceborne and ground-based radar comparisons by using quality filters

Irene Crisologo[1], Robert Warren[2], Kai Mühlbauer[3], and Maik Heistermann[1]

[1]Institute of Earth and Environmental Sciences, University of Potsdam, Germany
[2]School of Earth, Atmosphere and Environment, Monash University, Australia
[3]Meteorological Institute, University of Bonn, Germany

Correspondence: Irene Crisologo (icrisologo@gmail.com)

Abstract.

Coinciding monsoon and typhoon seasons in the Philippines cause torrential rainfall, and associated hazards such as flooding and landslides. While early warning systems require accurate radar-based rainfall estimates, low-density rain gauge networks in the Philippines make it challenging to monitor the calibration of the ground-based radars (GRs). As an alternative, we explore the potential of spaceborne radar (SR) observations from the Ku-band precipitation radars on board the TRMM and GPM satellites as a reference to quantify the calibration bias of an S-band GR in the Philippines. To this end, the 3D volume-matching algorithm proposed by Schwaller and Morris (2009) is implemented and applied to five years (2012-2016) of observations. We further extend the procedure by a framework to take into account the data quality of each ground radar bin. Through these methods, we are able to assign a quality index to each matching SR-GR volume, and thus compute the GR calibration bias as a quality-weighted average of reflectivity differences in any sample of matching GR-SR volumes. We exemplify the idea of quality-weighted averaging by using the beam blockage fraction as a basis of a quality index. As a result, we can increase the consistency of SR and GR observations, and thus the precision of calibration bias estimates. The remaining scatter between GR and SR reflectivity, as well as the variability of bias estimates between overpass events indicate, however, that other error sources are not yet fully addressed. Still, our study provides a framework to introduce any other quality variables that are considered relevant in a specific context. The code that implements our analysis is based on the open source software library *wradlib*, and is, together with the data, publicly available to monitor radar calibration, or to scrutinize long series of archived radar data back to December 1997, when TRMM became operational.

1 Introduction

Weather radars are essential tools in providing high quality information about precipitation with high spatial and temporal resolution in three dimensions. However, several uncertainties deteriorate the accuracy of rainfall products, with calibration contributing the most amount (Houze Jr et al., 2004), while also varying in time (Wang and Wolff, 2009). While adjusting

ground radars (GR) by comparison with a network of rain gauges is a widely used method, it suffers from representativeness issues. Furthermore, comparison with rain gauge observations accumulates uncertainties along the entire rainfall estimation chain (e.g. including the uncertain transformation from reflectivity to rainfall rate), and thus does not provide a direct reference for the measurement of reflectivity. The comparison with space-borne precipitation radars (SR) (such as those on-board the

5 Tropical Rainfall Measuring Mission (TRMM; 1997-2014) and the subsequent Global Precipitation Measurement mission (GPM; 2014-present)) has been steadily gaining popularity, since both ground radars and space-borne precipitation radars provide a volume-integrated measurement of reflectivity, thus allowing for a direct comparison of the observations in three dimensions (Anagnostou et al., 2001; Gabella et al., 2006, 2011; Warren et al., 2017). In addition, as the spaceborne radars are and have been constantly monitored (with their calibration accuracy proven to be consistently within 1 dB; Kawanishi et al.

10 (2000); Takahashi et al. (2003)), they have been suggested as a suitable reference for ground radar calibration (Anagnostou et al., 2001; Islam et al., 2012; Liao et al., 2001; Schumacher and Houze Jr, 2003).

Comparing measurements of space-borne radars with ground radars was originally suggested by Schumacher and Houze (2000), but the first method to match SR and GR reflectivity measurements was developed by Anagnostou et al. (2001). In their method, SR and GR measurements are resampled to a common three-dimensional grid. Liao et al. (2001) developed a similar

15 resampling method. Such 3D-resampling methods have been used in comparing SR and GR for both SR validation and GR bias determination (Bringi et al., 2012; Gabella et al., 2006, 2011; Park et al., 2015; Zhong et al., 2017; Wang and Wolff, 2009). Another method was suggested by Bolen and Chandrasekar (2003) and later on further developed by Schwaller and Morris (2011), where the SR-GR matching is based on the geometric intersection of SR and GR beams. This geometry matching algorithm confines the comparison to those locations where both instruments have actual observations, without interpolation or

20 extrapolation. The method has also been used in a number of studies comparing SR and GR reflectivities (Chandrasekar et al., 2003; Chen and Chandrasekar, 2016; Islam et al., 2012; Kim et al., 2014; Wen et al., 2011). A sensitivity study by Morris and Schwaller (2011) found that method to give more precise estimates of calibration bias as compared to grid-based methods.

Due to different viewing geometries, ground radars - in contrast to space-borne precipitation radars - are affected by processes that can deteriorate reflectivity measurements. Observational errors with regard to atmospheric properties such as reflectivity

25 are, for example, caused by ground clutter or partial beam blocking. Persistent systematic errors in the observation of reflectivity by ground radars are particularly problematic: the intrinsic assumption of the bias estimation is that the only systematic source of error is radar calibration. It is therefore particularly important to address such systematic observation errors.

In this study, we demonstrate that requirement with the example of partial beam blocking. The analysis is entirely based on algorithms implemented in the open source software library *wradlib* (Heistermann et al., 2013b), including a technique to

30 infer partial beam blocking by simulating the interference of the radar beam with terrain surface based on a digital elevation model. Together, that approach might become a reference for weather services around the world who are struggling to create unbiased radar observations from many years of archived single-polarized radar data, or to consistently monitor the bias of their radar observations. We demonstrate the approach in a case study with five years of data from the single-polarized S-band radar near the city of Subic, Philippines, which had been shown in previous studies to suffer from substantial miscalibration

(Abon et al., 2016; Heistermann et al., 2013a). The approach, however, is limited to within ±35 degrees latitude for TRMM and ±65 degrees latitude for GPM.

2 Methods

2.1 Data

2.1.1 Ground Radar

The Philippine Atmospheric, Geophysical, and Astronomical Services Administration (PAGASA) maintains a nationwide network of ten weather radars, eight of which are single-polarization S-band radars and two are dual-polarization C-band radars. Subic radar, which covers the greater Metropolitan Manila area, has the most extensive set of archived data. The radar coverage includes areas that receive some of the highest mean annual rainfall in the country.

The Subic radar sits on top of a hill at 532 m.a.s.l. in the municipality of Bataan, near the border to Zambales (location: 14.82 °N, 120.36 °E) (see Figure 1). To its south stands Mt. Natib (1253 m.a.s.l.) and to its north runs the Zambales Mountain Range (highest peak stands at 2037 m.a.s.l.). To the west is the Redondo Peninsula in the southern part of the Zambales province, where some mountains are also situated. Almost half of the coverage of the Subic radar is water, with Manila Bay to its southeast and the West Philippine Sea to the west. Technical specifications of the radar are summarized in Table 1. Data from April 2012 to December 2016 were obtained from PAGASA. Throughout the five years the scan strategy remained the same, except for 2015 when it was limited to only three elevation angles per volume due to hardware issues. The standard scanning strategy was re-implemented in 2016.

2.1.2 Space-Borne Precipitation Radar

Precipitation radar data were gathered from TRMM 2A23 and 2A25 version 7 products for overpass events intersecting with the Subic ground radar coverage between 1 June 2012 to 30 September 2014, and GPM 2AKu version 6 products from 1 June 2014 to 31 December 2016, covering the monsoon season of June to August which overlaps with the typhoon season of June to November. On that basis, we collected 225, 378, and 295 TRMM overpasses for 2012, 2013, and 2014, respectively. For GPM, there were 178, 224, and 211 for 2014, 2015, and 2016. The data were downloaded from NASA's Precipitation Processing System (PPS) through the STORM web interface (https://storm.pps.eosdis.nasa.gov/storm/). The parameters of TRMM/GPM extracted for the analysis are the same as Warren et al. (2017; their Table 3).

2.2 Partial Beam Blocking

In an ideal situation, SR and GR should have the same measurements for the same volume of the atmosphere, as they are measuring the same target. However, observational differences may arise due to different view geometries, different operating frequencies, different environmental conditions of each instrument, and different processes along the propagation path of the beam. As pointed out before, we focus on beam blockage as an index of GR data quality.
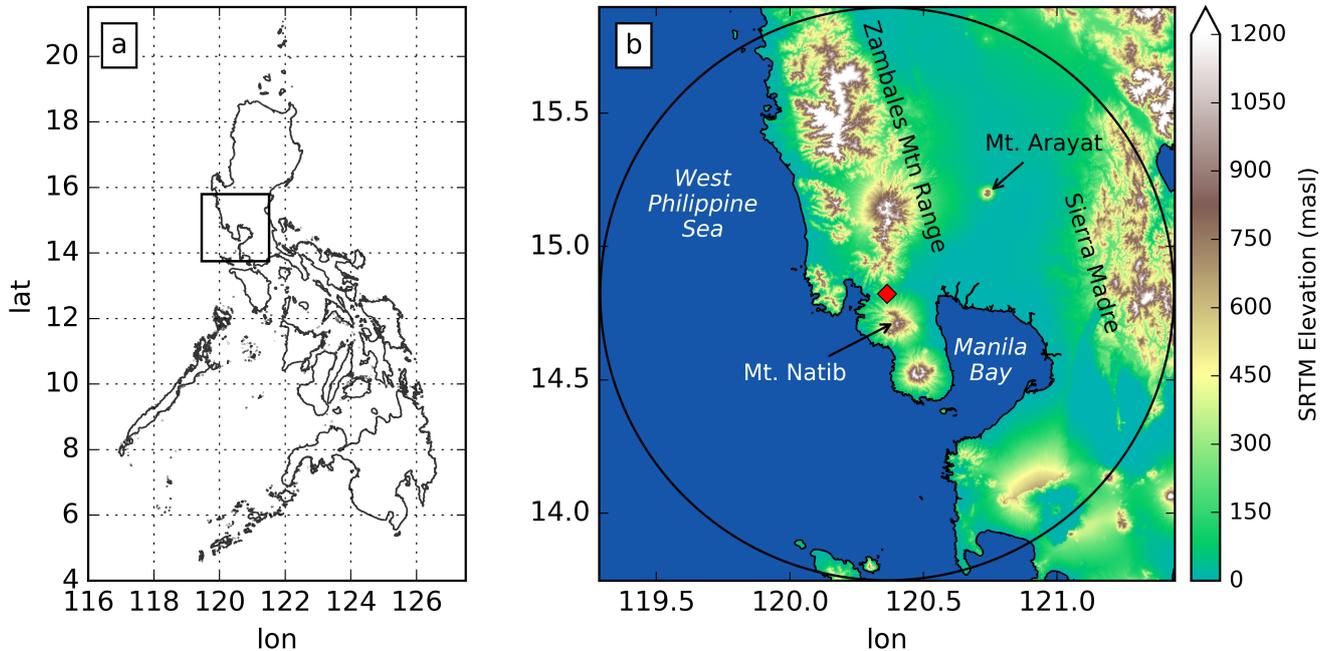
Figure 1. (a) A map of the Philippines showing the region of study and (b) the 120km coverage of the Subic radar (location marked with red diamond) with the SRTM Digital Elevation Model of the surrounding area.

In regions of complex topography, ground radars are typically affected by the effects of beam blockage, induced by the interaction of the beam with the terrain surface or other solid obstacles along the beam propagation path, resulting into a weakening or even loss of the signal.

To quantify that process within the Subic radar coverage, a beam blockage map is generated following the algorithm proposed

5   by Bech et al. (2003). It assesses the extent of occultation using a digital elevation model (DEM). The DEM used in this study is from the Shuttle Radar Topography Mission (SRTM) data, with 1-arc-second (approximately 30-meter) resolution. The DEM was resampled to the coordinates of the radar bin centroids, using spline interpolation, in order to match the polar resolution of the radar data (500 m in range and 1° in azimuth, extending to a maximum range of 120 km from the radar site; see Figure 1). A beam blockage map is generated for all available elevation angles.

10   The function `wradlib.qual.beam_block_frac()` which is an implementation of this algorithm was used to calculate the beam blockage fraction for each bin and each antenna pointing angle. The function `wradlib.qual.cum_beam_block_frac` was used to compute the cumulative beam blockage along each ray. A cumulative beam blockage fraction (BBF) of 1.0 corresponds to full occlusion, and a value of 0.0 to perfect visibility.

Table 1. Characteristics of the Subic radar and its volume scan strategy. The numbers in parentheses correspond to scans in 2015, where the scanning strategy was different due to hardware issues.

	Subic Radar
Polarization	Single-Pol
Position (lat/lon)	$14.82°N\ 120.36°E$
Altitude	532 (m.a.s.l.)
Maximum Range	120 km (150 km)
Azimuth resolution	1 °
Beam width	0.95 °
Gate length	500 m (250 m)
Number of elevation angles	14 (3)
Elevation angles (°)	0.5, 1.5, 2.4, 3.4, 4.3, 5.3, 6.2, 7.5, 8.7, 10, 12, 14, 16.7, 19.5 (0.0, 1.0, 2.0)
Volume cycle interval	9 minutes
Data available since	April 2012
Peak power	850 kW
Wavelength	10.7 cm

The quality index based on beam blockage fraction is then computed as follows, following the equation from Zhang et al. (2011):

$$Q_{BBF} = \begin{cases} 1 & BBF \leq 0.1 \\ 1 - \frac{BBF - 0.1}{0.4} & 0.1 < BBF \leq 0.5 \\ 0 & BBF > 0.5 \end{cases} \tag{1}$$

With this index, all the values with near perfect visibility have the highest quality, while those below 50% visibility have

5   zero quality.

2.3 SR-GR Volume Matching

For each SR gate, several meta-data parameters were extracted, such as the corresponding ray's bright band height and width, gate coordinates in three dimensions (longitude and latitude of each ray's Earth intercept and range gate index), time of overpass, precipitation type (*stratiform*, *convective*, or *other*), and rain indicators (*rain certain* or *no-rain*). The parallax-corrected

10   altitude (above mean sea level) and horizontal location (with respect to the GR) of each gate were determined as outlined in the appendix of Warren et al. (2017). From the bright band height/width and the altitude of each SR gate, the bright band membership of each gate was calculated by grouping all rays in an overpass and computing the mean brightband height and
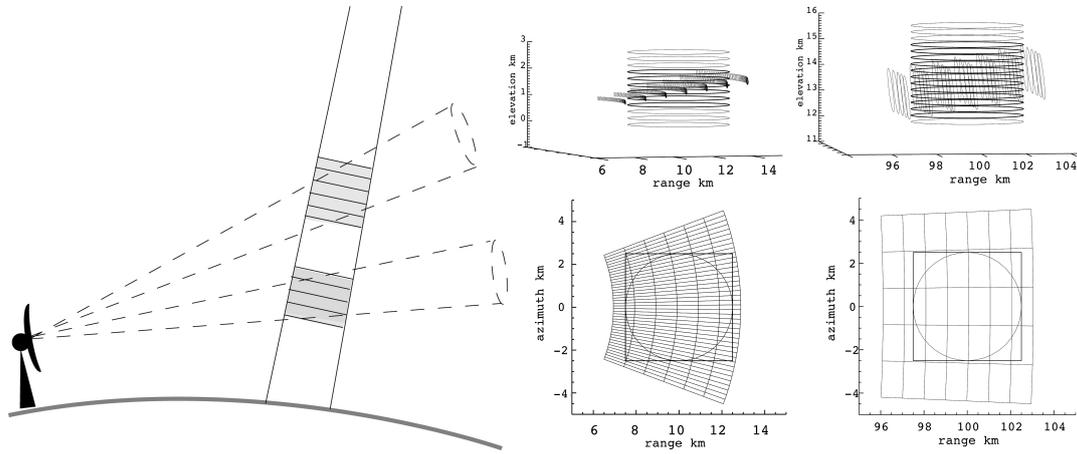
Figure 2. Diagram illustrating the geometric intersection. Left panel shows a single SR beam intersection ground radar beams of two different elevation angles. From Schwaller and Morris (2011) ©American Meteorological Society. Used with permission.

width, using the `wradlib.qual.get_bb_ratio` function. A ratio value of less than zero indicates that the gate is below bright band, greater than one indicates that the gate is above the bright band, and a value between zero and one means that the gate is within the bright band. The gates below and above the brightband were considered in the comparison. To ensure that there are sufficient bins with actual rain included in the comparison, overpasses with less than 100 gates flagged as rain certain

5 were discarded.

For each SR overpass, the GR sweep with the scan time closest to the overpass time within a 5-min window was selected. Both the SR and GR data were then geo-referenced into a common azimuthal equidistant projection centered on the location of the ground radar.

In order to minimize systematic differences in comparing the SR and GR reflectivites caused by the different measuring

10 frequencies, the SR reflectivities are converted from Ku to S Band (using the function `wrl.trafo.KuBandToS`) following the formula from Cao et al. (2013):

$$Z(S) = Z(Ku) + \sum_{i=0}^{4} a_i [Z(Ku)]^i \qquad (2)$$

where the a_i are the coefficients for dry snow and dry hail, rain, and in between at varying melting stages (Table 1 of Cao et al. (2013)).

15 The actual volume matching algorithm closely follows the work of Schwaller and Morris (2011), where SR reflectivity is spatially and temporally matched with GR reflectivity without interpolation. The general concept is highlighted by Figure 2: Each matching sample consists of bins from only *one SR ray* and *one GR sweep*. From the SR ray, those bins were selected that intersect with the vertical extent of a specific GR sweep at the SR ray location (as a result of GR beam width and distance from radar). From each GR sweep, those bins were selected that intersect with the horizontal footprint of the SR ray at the corre-

Table 2. Filtering criteria for the matching workflow.

Criteria	Condition
Minimum number of pixels in overpass tagged as 'rain'	100
Bright band membership	below and above
Range limits (min - max)	15 km - 115 km
Minimum fraction of bins above minimum SR sensitivity	0.7
Minimum fraction of bins above minimum GR sensitivity	0.7
Maximum time difference between SR and GR	5 min
Minimum PR reflectivity considered	18 dBZ
Minimum GR reflectivity considered	0 dBZ

sponding altitude. The SR and GR reflectivity of each matched volume was computed as the average reflectivity of the intersecting SR and GR bins, where the averaging is done in linear units (mm^6/m^3). The entire workflow of this volume matching procedure is based on wradlib functions and available at `https://github.com/wradlib/radargpm-beamblockage`.

The nominal minimum sensitivity of both TRMM PR and GPM KuPR is 18 dBZ, so only values above this level were
5 considered in the calculation of average SR reflectivity in the matched volume. In addition, the fraction of SR gates within a matched volume above that threshold was also recorded. On the other hand, all GR bins are included in the calculation of average GR reflectivity, after setting the bins with reflectivities below 0.0 dBZ to 0.0 dBZ, as also suggested by Schwaller and Morris (2011). The filtering criteria applied in the workflow are summarized in Table 2.

2.4 Estimating GR calibration bias

10 Beam blockage and the corresponding GR quality maps were computed for each GR bin (cf. Section 2.2.) For each matched SR-GR volume, the data quality was then based on the minimum quality of the GR bins in that volume.

To analyze the effect of data quality on the estimation of GR calibration bias, we compared two estimation approaches: a simple mean bias that does not take into account beam blockage, and a weighted mean bias that considers the quality value of each sample as weights. The corresponding standard deviation and weighted standard deviation were calculated as well. The
15 overall process is summarized in Figure 3. This way, we provide an overview of the variability of our bias estimates over time.

3 Results and Discussion

3.1 Beam Blockage Map

Figure 4 shows the beam blockage map for the two lowest elevation angles of each scanning strategy. Figure 4a and c are for 0.0° and 1.0°, which are the two lowest elevation angles in 2015, while Figure 4b and d are for 0.5° and 1.5°, which are the
20 two lowest elevation angles for the rest of the dataset.
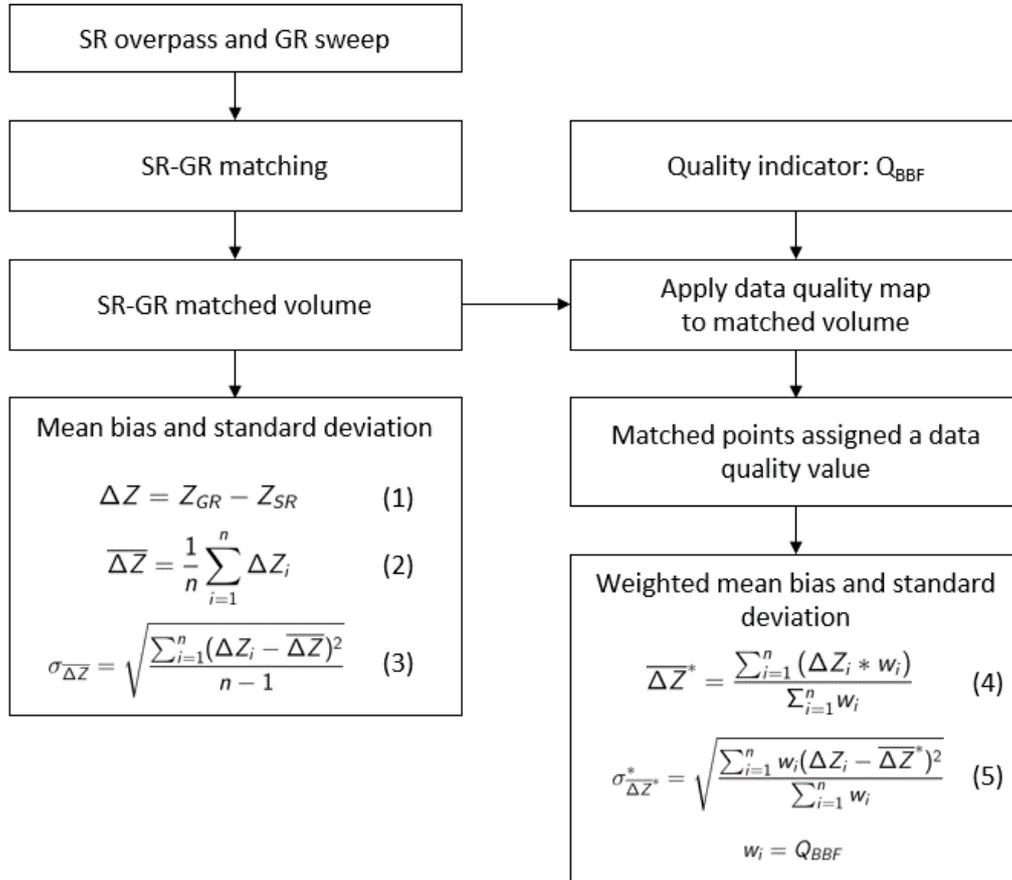
Figure 3. Flowchart describing the processing steps to calculate the mean bias and the weighted mean bias between ground radar data and satellite radar data. The results of each step are shown in Section 3

As expected, the degree of beam blockage decreases with increasing antenna elevation, yielding the most pronounced beam blockage at 0.0°. Each blocked sector can be explained by the topography (see Figure 1), with the Zambales Mountain Range causing blockage in the northern sector, Mt. Natib in the southern sector, and the Redondo peninsula mountains in the western sector. The Sierra Madre mountains also cause some partial beam blocking at the far east, and a narrow partial blocking northeast of the station where Mount Arayat is located.

As the elevation angle increases, the beam blockage becomes less pronounced or even disappears. Substantial blockage persists, however, for the higher elevation angles in the northern and southern sectors.
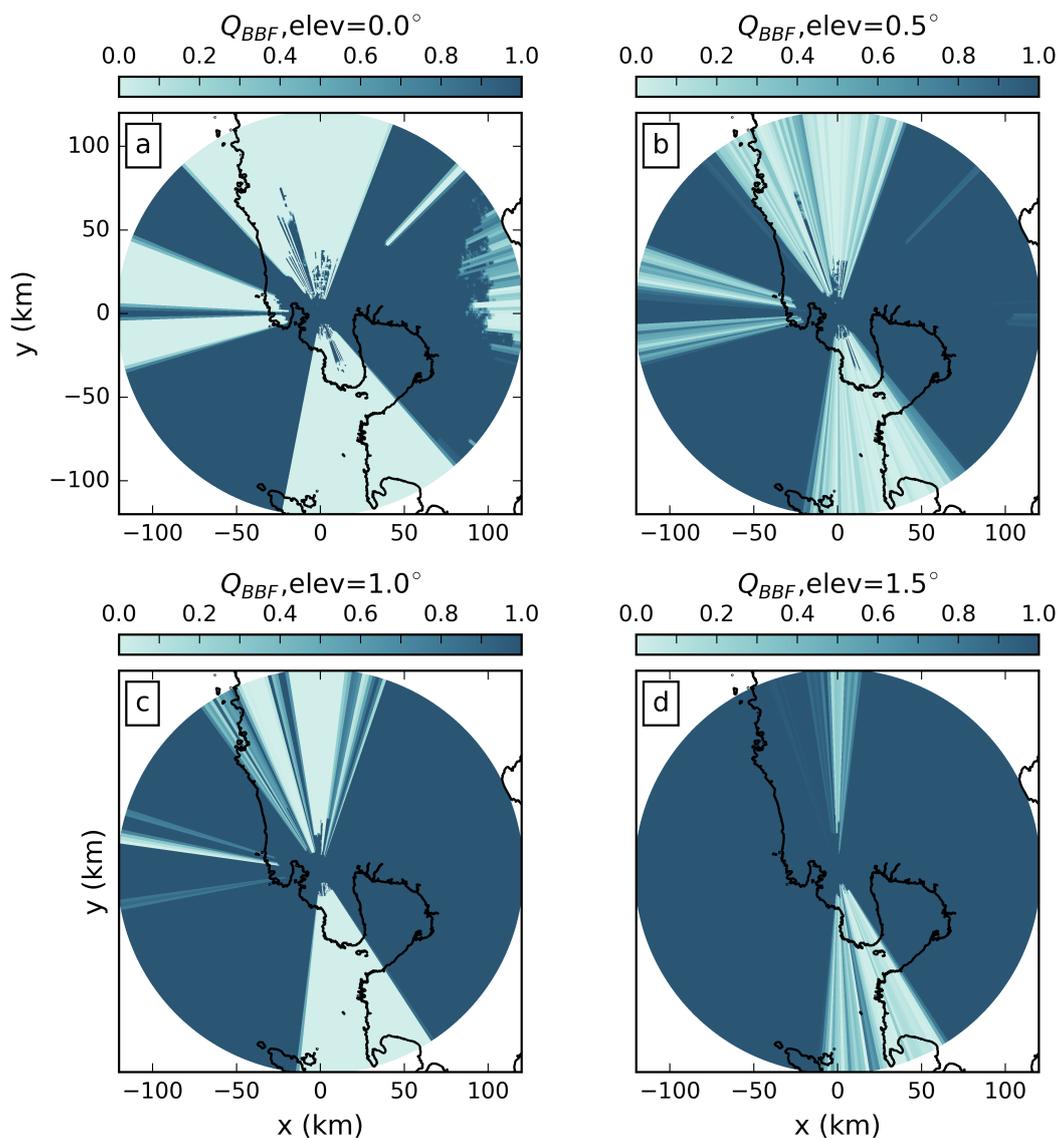
Figure 4. Beam blockage fraction for the Subic radar at (a) 0.0° (b) 0.5° (c) 1.0° and (d) 1.5° elevation angles.

3.1.1 SR-GR Matching

There were 183 TRMM and 103 GPM overpasses that overlapped with the 120 km Subic radar range from 2012 to 2016. However these numbers decreased to 74 (for TRMM) and 40 (for GPM) after applying the selection criteria listed in Table 2. In order to get a better idea about the overall workflow, we first exemplify the results for two specific overpass events - one for TRMM, and one for GPM.

3.1.2 Case 1: 08 November 2013

For the TRMM overpass event on November 8, 2013, the top row of Figure 5 shows SR (a) and GR (b) reflectivity as well as the resulting differences (c) for matching samples at an elevation angle of 0.5°. Each circle in the plots represents a matched volume. A corresponding map of Q_{BBF} is shown in (d) while (e) shows a scatter plot of GR versus SR reflectivities, with points coloured according to their Q_{BBF}. The reflectivity difference map and scatter plot indicate significant variability with differences of up to -10 dB. Large differences can be observed at the edges of the southern sector affected by beam blockage (cf. also Figure 4). Major parts of that sector did not receive any signal due to total beam blockage. At the edges, however, partial beam blockage caused substantially lower GR reflectivity values. As expected, large negative differences of Z_{GR} - Z_{SR} are characterized by low quality.

Consequently, the estimate of the calibration bias substantially depends on the consideration of partial beam blockage (or quality). Ignoring quality (simple mean) yields a bias estimate of -1.9 dB while the quality-weighted average yields a bias estimate of -1.2 dB. Accordingly, the standard deviation is reduced from 3.4 to 2.6 dB, indicating an increased precision of the bias estimate.

This case demonstrates how partial beam blockage affects the estimation of GR calibration bias. It is important to note that beam blockage only becomes an issue once the reflectivity observed by the ground radar exceeds the detection limit. As a consequence, total beam blockage does not interfere with the bias estimation.

The effect becomes even more obvious for the next elevation angle. Figure 6 is equivalent to Figure 5, but for an elevation angle of 1.5°: As the sector of total beam blockage shrinks at that elevation, the impact of partial beam blockage on the estimation of GR calibration bias increases. That might be considered counterintuitive, as one might expect the effect of beam blockage to disappear at higher elevations. For an antenna elevation of 1.5°, however, ignoring beam blockage yields a bias estimate of -2.1 dB (simple mean), while the quality-weighted average yields a bias of -1.4 dB. At the same time, considering quality substantially reduces the standard deviation from 3.4 dB to 2.1 dB.

3.1.3 Case 2: 01 October 2015

The second case confirms the findings in the previous section for a GPM overpass on October 1, 2015. That overpass captured an event in the northern and eastern part of the radar coverage where partial beam blockage is dominant, as well as a small part of the southern sector with partial and total beam blockage. Figure 7 shows the results of that overpass in analogy to the previous figures, for an antenna elevation of 0.0 degree. The figure shows a dramatic impact of partial beam blockage, with a
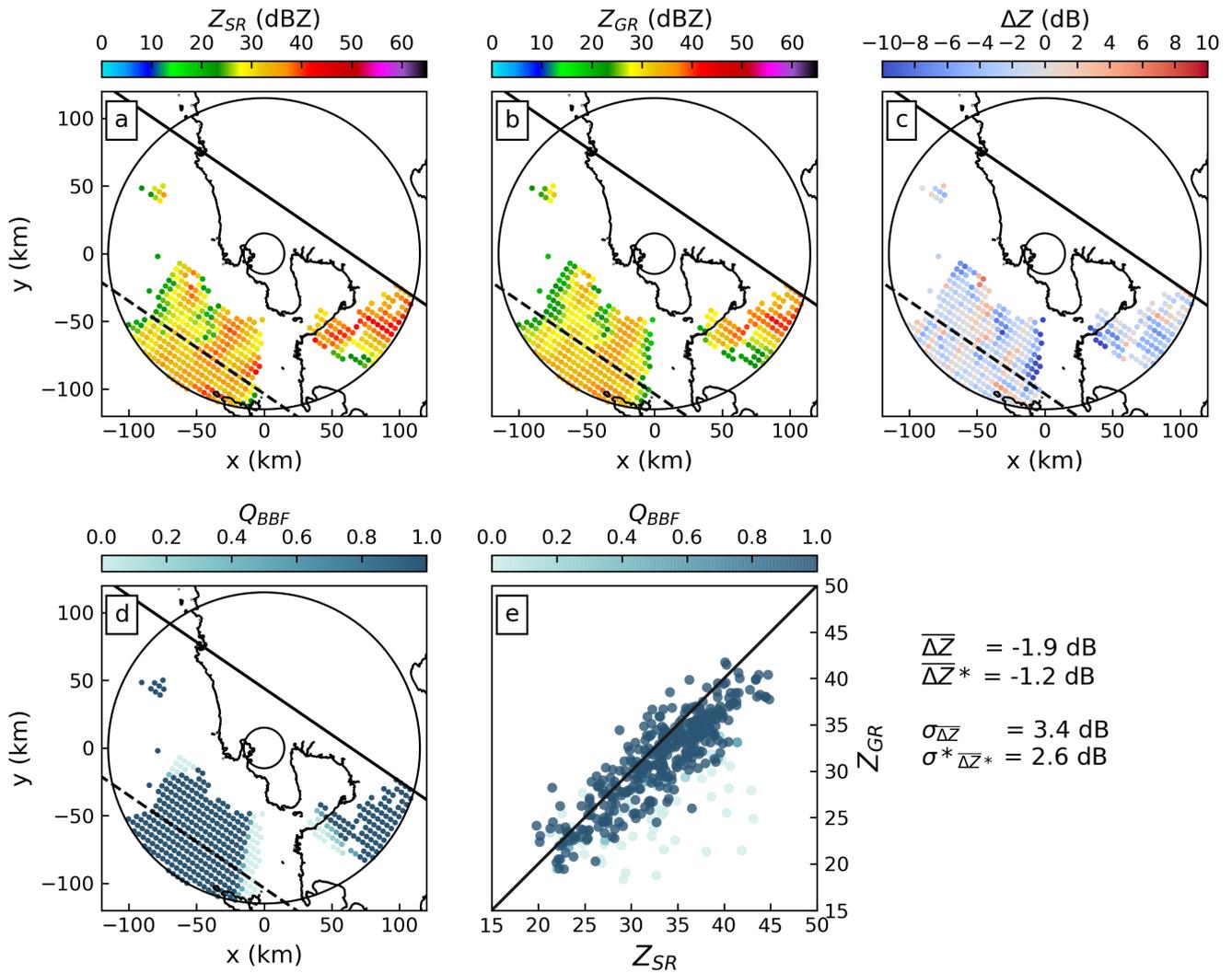
Figure 5. PPI of volume-matched samples from 8 November 2013 at 0.5° elevation angle of (a) SR reflectivity,(b) GR Reflectivity, (c) difference between GR and SR reflectivities, and (d) Q_{BBF}. (e) Scatter plot of Z_{GR} vs Z_{PR} where each point is colored based on the data quality (Q_{BBF}). The solid line in (a)-(d) is the edge of the SR swath, the other edge lies outside the figure. The dotted line denotes the central axis of the swath. The solid circles demarcate the 15 km and 115 km ranges from the radar.
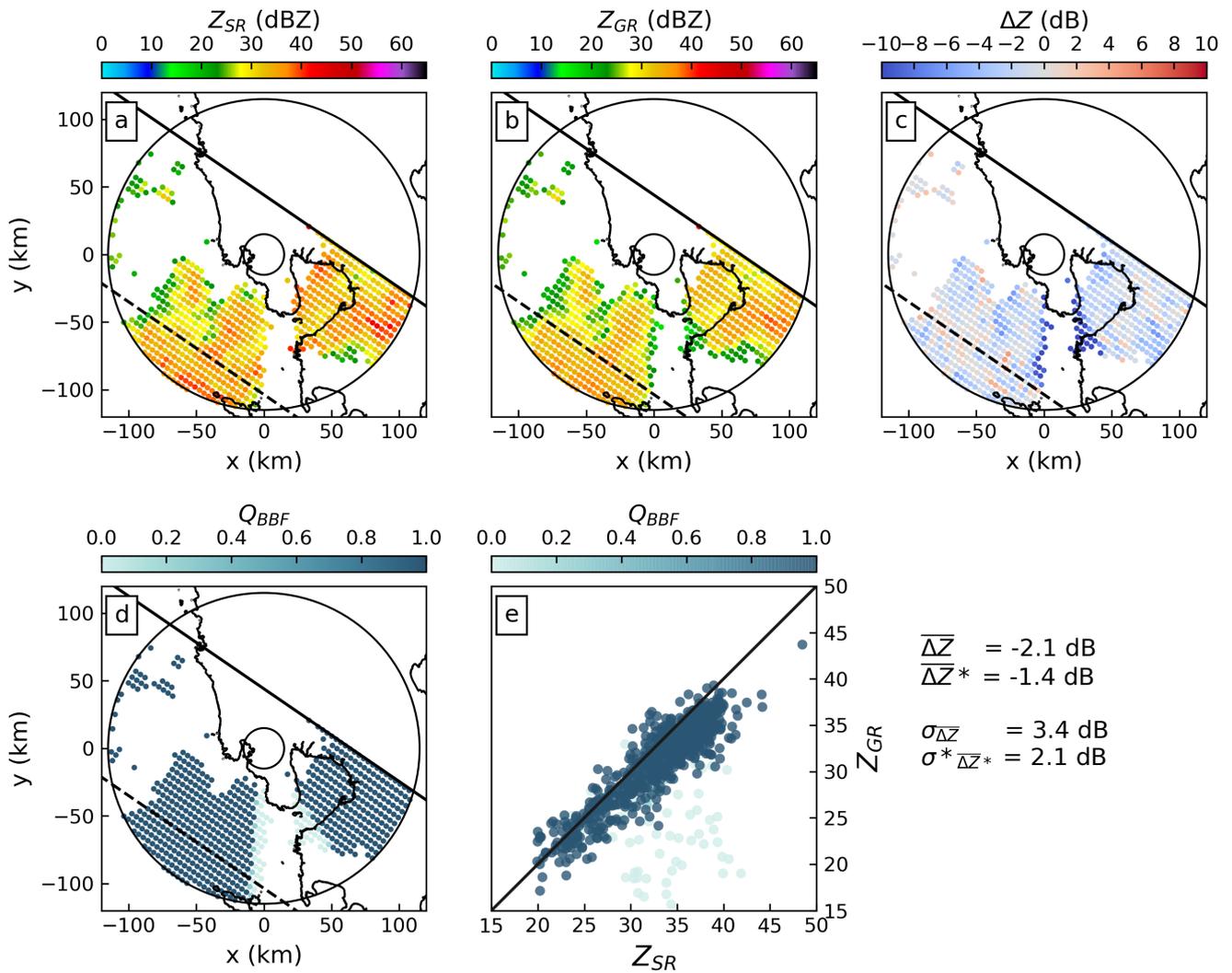
Figure 6. Same as in Figure 5 but for 1.5° elevation angle
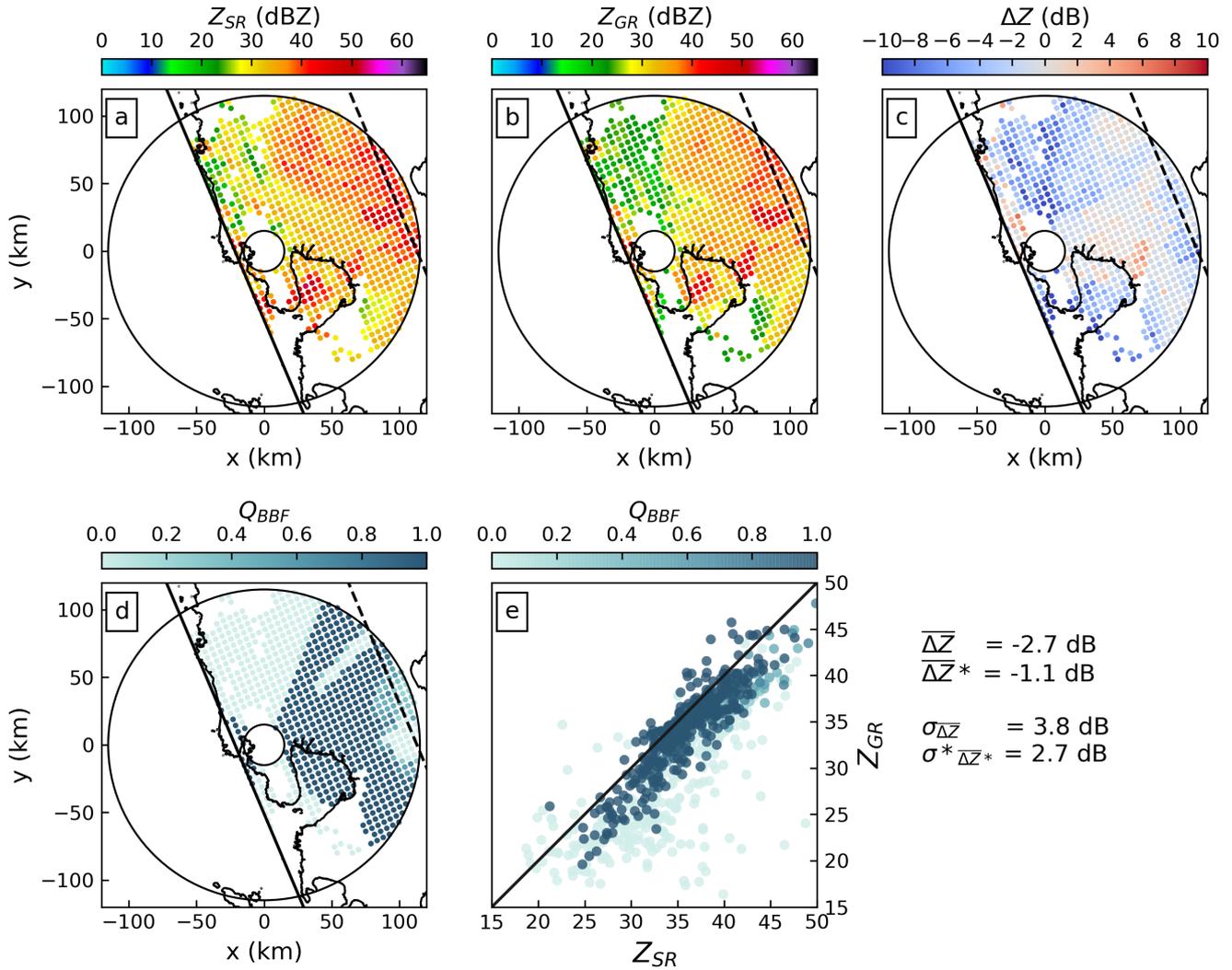
Figure 7. As in Figure 5 but for the overpass 01 October 2015

dominant contribution from the northern part, but also clear effects from the eastern and southern sectors. The scatter plot of Z_{GR} over Z_{SR} (e) demonstrates how the consideration of partial beam blockage increases the consistency between GR and SR observations and allows for a more reliable estimation of the GR calibration bias: Ignoring partial beam blockage (simple mean) yields a bias of -2.7 dB, while the quality-weighted average bias is -1.1 dB. Taking into account quality decreases the standard deviation from 3.8 dB to 2.7 dB.
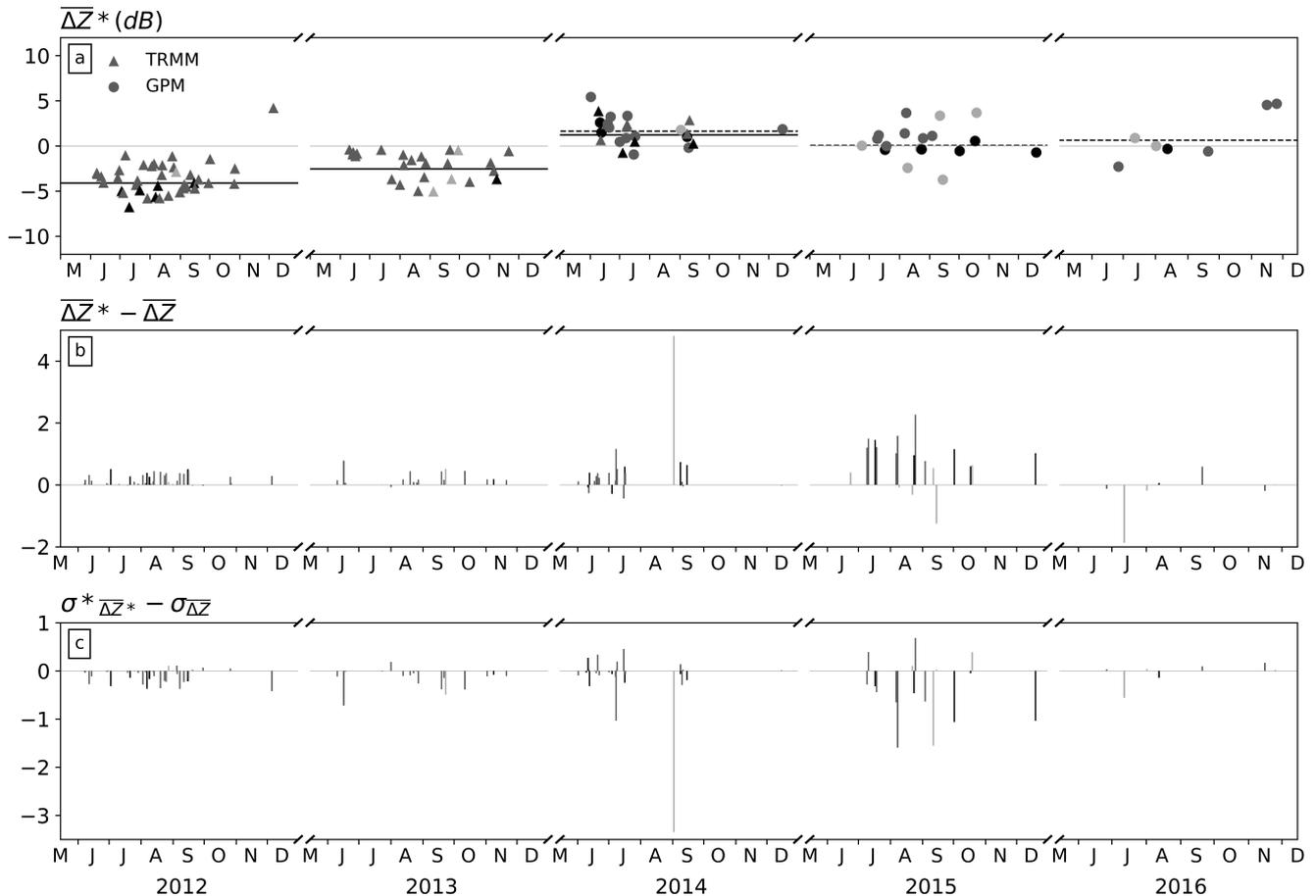
Figure 8. (a) Time series of the weighted mean bias (ΔZ^*) from 2012 to 2016. Triangle markers represent comparison with TRMM overpasses while circle markers are comparisons with GPM overpasses. Symbols are colored according to the number of volume-matched sample pairs on a logarithmic scale: light gray = 10-99, medium gray = 100-999, and black = 1000+. Solid and dashed horizontal lines represent the weighted average of all individual matched points within the year for TRMM and GPM, respectively. (b) The difference between the weighted mean biases (ΔZ^*) and the simple mean biases (ΔZ). (c) The standard deviation of the weighted mean bias values minus the standard deviation of the simple mean bias values.

3.2 Calibration bias over time

Finally, we applied both the simple and the quality-weighted mean bias estimation to each of the TRMM and GPM overpasses from 2012 to 2016 that met the criteria specified in Section 2.3 Table 2. As pointed out in Section 2.3, the matching procedure itself is carried out per GR sweep, ı.e. separately for each antenna elevation angle.

5   As a result, we obtain a time series of bias estimates for GR calibration, as shown in Figure 8. In this figure, the calibration bias for each overpass is computed from the full GR volume, ı.e. including matched samples from all available antenna ele-

vations. In the upper panel (a), each marker represents the quality-weighted mean bias for a specific SR overpass (circles for GPM, triangles for TRMM). The center panel (b) highlights the differences between the quality-weighted and the simple mean approach, ı.e. it quantifies the effect of taking into account GR data quality (ı.e. partial beam blockage). The bottom panel (c) shows the differences between the quality-weighted standard deviation and the simple standard deviation of differences, illustrating how taking into account GR quality affects the precision of the bias estimates.

The time series provide several important insights:

(1) **Short term variability**: There is a strong variability of the estimated calibration bias between overpasses (Figure 8a). That variability appears to have a strong random component, and it is clearly not a desirable property: typically, we would not expect changes in calibration bias to occur at the observed frequency and amplitude. Yet, it is beyond the scope of this study to disentangle the sources of this variability. We have to assume that the fluctuations are a cumulative result of various effects that might include e.g. hardware instability, fluctuations in the refractive index of the atmosphere, dynamic changes in the atmosphere in the time interval between SR overpass and GR sweep, non-meteorological echoes, residual errors in the geometric intersection of the volume samples, and other random error components, as noted in other studies (Anagnostou et al., 2001; Kim et al., 2014; Schwaller and Morris, 2011; Wang and Wolff, 2009; Warren et al., 2017).

(2) **Change of bias over time**: Despite the variability of bias estimates between the individual overpass events, the time series still provides us with a clear signal: The bias estimates appear to fluctuate around an average value that appears to be quite persistent over the duration of the corresponding wet seasons of the different years, so over intervals of several months. Considering the average calibration bias over the different wet seasons (horizontal lines in Figure 8a), we can clearly observe changes in calibration bias over time. The bias was most pronounced in 2012 and 2013, with average bias estimates around -4.1 dB for 2012 and -2.5 dB for 2013. For 2014, the absolute calibration bias was much smaller, at a level of 1.4 dB while for 2015 and 2016, the situation improved further, with an average bias of 0.03 dB in 2015 and 0.6 dB in 2016. It is important to note that these values were computed as the average bias across all matched volumes and not as the average of biases across overpasses, as the overpasses have different sample sizes. We have to assume that a fundamental issue with regard to calibration maintenance was addressed between 2013 and 2014, although we could not confirm that with the radar operator.

(3) **Effect of quality weighting on bias estimation**: Figure 8b and c together illustrate the benefit of taking into account GR data quality (i.e. beam blockage) when we estimate GR calibration bias. It is important to understand that only the center and the bottom panel together prove that benefit: It does not come as a surprise that the quality-weighted mean bias is consistently higher than the simple mean bias (Figure 8b) since we give less weight to those GR bins affected by partial beam blockage. But only Figure 8c shows that the quality weighted bias estimates are also consistently more precise: in the vast majority of overpasses the quality weighted standard deviation is substantially smaller than the simple standard deviation. That result is also consistent with the case study result shown above. However, when the matched sample size is already small, further filtering of samples increases the standard deviation.

(4) **GPM and TRMM radars are consistent**: In 2014, both TRMM and GPM overpasses are available. That period of overlap shows that the GR calibration bias estimates that are based on both TRMM and GPM observations can be considered as homogeneous. Using TRMM data, the average calibration bias for all 2014 overpasses amounts to 1.6 ± 1.3 dB, while using

the GPM overpasses yields a bias of 1.8 ± 1.5 dB. That makes us confident that the substantial year-to-year changes of our bias estimates are based on changes in GR calibration, not based on differences between TRMM and GPM sensors.

4   Conclusions

In 2011, Schwaller and Morris presented a new technique to match spaceborne radar (SR) and ground-based radar (GR)
5   reflectivity observations, with the aim to determine the GR calibration bias. Our study extends that technique by a coherent approach to take into account the quality of the ground radar observations: Each GR bin was assigned a quality index between 0 and 1, which was again used to assign a quality value to each matched volume of SR and GR observations. For any sample of matched volumes (e.g. all matched volumes of one overpass, or a combination of multiple overpasses), the calibration bias can then be computed as a quality-weighted average of the differences between GR and SR reflectivity in all samples.

10   We exemplified that approach by inferring a GR data quality index based on the beam blockage fraction, and we demonstrated the added value for both TRMM and GPM overpasses over the 120 km range of the Subic S-band radar in the Philippines for a five year period.

Although the variability of the calibration bias estimates between overpasses is high, we could show that taking into account partial beam blockage leads to more consistent and more precise estimates of GR calibration bias. Analyzing five years of
15   archived data from the Subic S-band radar (2012-2016), we could also demonstrate that the calibration standard of the Subic radar substantially improved over the years, from bias levels around -4.1 dB to bias levels around 0.6 dB.

Case studies for specific overpass events also showed that the necessity to account for partial beam blockage might even increase for higher antenna elevations. That applies when sectors with total beam blockage (in which no valid matched volumes are retrieved at all) turn into sectors with partial beam blockage at higher elevation angles.

20   Considering the scatter between SR and GR reflectivity in the matched volumes of one overpass (see case studies), as well as the variability of bias estimates between satellite overpasses (see time series), it is obvious that we do not yet account for various sources of systematic and random sources of observational errors. Also the simulation of beam blockage itself might still be prone to various errors (e.g. from geometric inaccuracies, variability of atmospheric refractivity, interpolation effects, errors in the DEM, or limited DEM resolution). Nevertheless, the idea of the quality-weighted estimation of calibration bias
25   presents a consistent framework that allows for the integration of any quality variables that are considered important in a specific environment or setting. For C-band radars, e.g. considering path-integrated attenuation would be vital. The framework could also be extended by explicitly assigning a quality index to SR observations, too. In the context of this study, that was implicitly implemented by filtering the SR data e.g. based on bright band membership.

Despite the fact that there is still ample room for improvement, our tool that combines SR-GR volume matching and
30   quality-weighted bias estimation is readily available for application or further scrutiny. In fact, our analysis is the first of its kind that is entirely based on open source software, and thus fully transparent, reproducible and adjustable (see also Heistermann et al. (2015)). Therefore this study, for the first time, demonstrates the utilization of wradlib functions that have just recently been implemented to support the volume matching procedure and the simulation of partial beam block-

age. We also make available the complete workflow that put together the different steps of the processing chain, together with the underlying ground and space-borne radar data. Both code and results can be accessed at the following repository `https://github.com/wradlib/radargpm-beamblockage` upon the publication of this manuscript.

Through these open-source resources, our methodology provides both research institutions and weather services with a
5 valuable tool that can be applied to monitor the radar calibration, and - perhaps more importantly - to quantify the calibration bias for long time series of archived radar observations, basically beginning with the availability of TRMM radar observations in December 1997.

Code and data availability. Code and sample data can be accessed at `https://github.com/wradlib/radargpm-beamblockage`

Competing interests. The authors declare that they have no conflict of interest.

10 *Acknowledgements.* The radar data for this analysis were provided by the Philippine Atmospheric, Geophysical and Astronomical Services Administration (PAGASA, http://pagasa.dost.gov.ph). The study was also funded by the German government through the German Academic Exchange Service (https://www.daad.de/en/).

References

Abon, C. C., Kneis, D., Crisologo, I., Bronstert, A., David, C. P. C., and Heistermann, M.: Evaluating the potential of radar-based rainfall estimates for streamflow and flood simulations in the Philippines, Geomatics, Natural Hazards and Risk, 7, 1390–1405, https://doi.org/10.1080/19475705.2015.1058862, http://www.tandfonline.com/doi/full/10.1080/19475705.2015.1058862, 2016.

5 Anagnostou, E. N., Morales, C. A., and Dinku, T.: The use of TRMM precipitation radar observations in determining ground radar calibration biases, Journal of Atmospheric and Oceanic Technology, 18, 616–628, http://journals.ametsoc.org/doi/abs/10.1175/1520-0426(2001)018%3C0616%3ATUOTPR%3E2.0.CO%3B2, 2001.

Bech, J., Codina, B., Lorente, J., and Bebbington, D.: The sensitivity of single polarization weather radar beam blockage correction to variability in the vertical refractivity gradient, Journal of Atmospheric and Oceanic Technology, 20, 845–855, http://journals.ametsoc.org/doi/abs/10.1175/1520-0426(2003)020%3C0845:TSOSPW%3E2.0.CO;2, 2003.

Bolen, S. M. and Chandrasekar, V.: Methodology for aligning and comparing spaceborne radar and ground-based radar observations, Journal of Atmospheric and Oceanic Technology, 20, 647–659, http://journals.ametsoc.org/doi/abs/10.1175/1520-0426(2003)20%3C647%3AMFAACS%3E2.0.CO%3B2, 2003.

Bringi, V. N., Huang, G.-J., Munchak, S. J., Kummerow, C. D., Marks, D. A., and Wolff, D. B.: Comparison of Drop Size Distribution

15 Parameter (D_0) and Rain Rate from S-Band Dual-Polarized Ground Radar, TRMM Precipitation Radar (PR), and Combined PR–TMI: Two Events from Kwajalein Atoll, Journal of Atmospheric and Oceanic Technology, 29, 1603–1616, https://doi.org/10.1175/JTECH-D-11-00153.1, http://journals.ametsoc.org/doi/abs/10.1175/JTECH-D-11-00153.1, 2012.

Cao, Q., Zhang, G., and Xue, M.: A Variational Approach for Retrieving Raindrop Size Distribution from Polarimetric Radar Measurements in the Presence of Attenuation, Journal of Applied Meteorology and Climatology, 52, 169–185, https://doi.org/10.1175/JAMC-D-12-

20 0101.1, http://journals.ametsoc.org/doi/abs/10.1175/JAMC-D-12-0101.1, 2013.

Chandrasekar, V., Bolen, S., and Gorgucci, E.: Microphysical cross validation of spaceborne radar and ground polarimetric radar, IEEE Transactions on Geoscience and Remote Sensing, 41, 2153–2165, https://doi.org/10.1109/TGRS.2003.817186, http://ieeexplore.ieee.org/document/1237374/, 2003.

Chen, H. and Chandrasekar, V.: Validation of NASA's Global Precipitation Measurement mission with a high-resolution ground radar

25 network, in: URSI Asia-Pacific Radio Science Conference (URSI AP-RASC), pp. 836–839, IEEE, http://ieeexplore.ieee.org/abstract/document/7601343/, 2016.

Gabella, M., Joss, J., Perona, G., and Michaelides, S.: Range adjustment for ground-based radar, derived with the spaceborne TRMM precipitation radar, IEEE Transactions on Geoscience and Remote Sensing, 44, 126–133, https://doi.org/10.1109/TGRS.2005.858436, http://ieeexplore.ieee.org/document/1564402/, 2006.

30 Gabella, M., Morin, E., and Notarpietro, R.: Using TRMM spaceborne radar as a reference for compensating ground-based radar range degradation: Methodology verification based on rain gauges in Israel, Journal of Geophysical Research, 116, https://doi.org/10.1029/2010JD014496, http://doi.wiley.com/10.1029/2010JD014496, 2011.

Heistermann, M., Crisologo, I., Abon, C. C., Racoma, B. A., Jacobi, S., Servando, N. T., David, C. P. C., and Bronstert, A.: Brief communication "Using the new Philippine radar network to reconstruct the Habagat of August 2012 monsoon event around Metropolitan

35 Manila", Nat. Hazards Earth Syst. Sci., 13, 653–657, https://doi.org/10.5194/nhess-13-653-2013, http://www.nat-hazards-earth-syst-sci.net/13/653/2013/, 2013a.

Heistermann, M., Jacobi, S., and Pfaff, T.: Technical Note: An open source library for processing weather radar data (<i>wradlib</i>), Hydrology and Earth System Sciences, 17, 863–871, https://doi.org/10.5194/hess-17-863-2013, http://www.hydrol-earth-syst-sci.net/17/863/2013/, 2013b.

Heistermann, M., Collis, S., Dixon, M. J., Giangrande, S., Helmus, J. J., Kelley, B., Koistinen, J., Michelson, D. B., Peura, M., Pfaff, T.,
5 and Wolff, D. B.: The Emergence of Open-Source Software for the Weather Radar Community, Bull. Amer. Meteor. Soc., 96, 117–128, https://doi.org/10.1175/BAMS-D-13-00240.1, http://journals.ametsoc.org/doi/abs/10.1175/BAMS-D-13-00240.1, 2015.

Houze Jr, R. A., Brodzik, S., Schumacher, C., Yuter, S. E., and Williams, C. R.: Uncertainties in oceanic radar rain maps at Kwajalein and implications for satellite validation, Journal of Applied Meteorology, 43, 1114–1132, http://journals.ametsoc.org/doi/abs/10.1175/1520-0450(2004)043%3C1114:UIORRM%3E2.0.CO%3B2, 2004.

10 Islam, T., Rico-Ramirez, M. A., Han, D., Srivastava, P. K., and Ishak, A. M.: Performance evaluation of the TRMM precipitation estimation using ground-based radars from the GPM validation network, Journal of Atmospheric and Solar-Terrestrial Physics, 77, 194–208, https://doi.org/10.1016/j.jastp.2012.01.001, http://www.sciencedirect.com/science/article/pii/S1364682612000028, 2012.

Kawanishi, T., Kuroiwa, H., Kojima, M., Oikawa, K., Kozu, T., Kumagai, H., Okamoto, K., Okumura, M., Nakatsuka, H., and Nishikawa, K.: TRMM Precipitation Radar, Advances in Space Research, 25, 969–972, https://doi.org/10.1016/S0273-1177(99)00932-1, http://www.
15 sciencedirect.com/science/article/pii/S0273117799009321, 2000.

Kim, J.-H., Ou, M.-L., Park, J.-D., Morris, K. R., Schwaller, M. R., and Wolff, D. B.: Global Precipitation Measurement (GPM) Ground Validation (GV) Prototype in the Korean Peninsula, Journal of Atmospheric and Oceanic Technology, 31, 1902–1921, https://doi.org/10.1175/JTECH-D-13-00193.1, http://journals.ametsoc.org/doi/abs/10.1175/JTECH-D-13-00193.1, 2014.

Liao, L., Meneghini, R., and Iguchi, T.: Comparisons of rain rate and reflectivity factor derived from the TRMM precipitation radar and the
20 WSR-88D over the Melbourne, Florida, site, Journal of Atmospheric and Oceanic Technology, 18, 1959–1974, http://journals.ametsoc.org/doi/abs/10.1175/1520-0426(2001)018%3C1959:CORRAR%3E2.0.CO%3B2, 2001.

Morris, K. R. and Schwaller, M.: Sensitivity of spaceborne and ground radar comparison results to data analysis methods and constraints, 2011.

Park, S., Jung, S.-H., and Lee, G.: Cross Validation of TRMM PR Reflectivity Profiles Using 3D Reflectivity Composite from the
25 Ground-Based Radar Network over the Korean Peninsula, J. Hydrometeorol., 16, 668–687, https://doi.org/10.1175/JHM-D-14-0092.1, wOS:000352735100013, 2015.

Schumacher, C. and Houze, Jr., R. A.: Comparison of Radar Data from the TRMM Satellite and Kwajalein Oceanic Validation Site., Journal of Applied Meteorology, 39, 2151–2164, https://doi.org/10.1175/1520-0450(2001)040<2151:CORDFT>2.0.CO;2, http://adsabs.harvard.edu/abs/2000JApMe..39.2151S, 2000.

30 Schumacher, C. and Houze Jr, R. A.: Stratiform rain in the tropics as seen by the TRMM precipitation radar, Journal of Climate, 16, 1739–1756, http://journals.ametsoc.org/doi/full/10.1175/1520-0442(2003)016%3C1739:SRITTA%3E2.0.CO%3B2, 2003.

Schwaller, M. R. and Morris, K. R.: A Ground Validation Network for the Global Precipitation Measurement Mission, J. Atmos. Oceanic Technol., 28, 301–319, https://doi.org/10.1175/2010JTECHA1403.1, http://journals.ametsoc.org/doi/citedby/10.1175/2010JTECHA1403.1, 2011.

35 Takahashi, N., Kuroiwa, H., and Kawanishi, T.: Four-year result of external calibration for Precipitation Radar (PR) of the Tropical Rainfall Measuring Mission (TRMM) satellite, IEEE Transactions on Geoscience and Remote Sensing, 41, 2398–2403, https://doi.org/10.1109/TGRS.2003.817180, 2003.

Wang, J. and Wolff, D. B.: Comparisons of Reflectivities from the TRMM Precipitation Radar and Ground-Based Radars, J. Atmos. Oceanic Technol., 26, 857–875, https://doi.org/10.1175/2008JTECHA1175.1, http://journals.ametsoc.org/doi/abs/10.1175/2008JTECHA1175.1, 2009.

Warren, R. A., Protat, A., Siems, S. T., Ramsay, H. A., Louf, V., Manton, M. J., and Kane, T. A.: Calibrating ground-based radars against
5   TRMM and GPM, J. Atmos. Oceanic Technol., https://doi.org/10.1175/JTECH-D-17-0128.1, http://journals.ametsoc.org/doi/abs/10.1175/JTECH-D-17-0128.1, 2017.

Wen, Y., Hong, Y., Zhang, G., Schuur, T. J., Gourley, J. J., Flamig, Z., Morris, K. R., and Cao, Q.: Cross Validation of Spaceborne Radar and Ground Polarimetric Radar Aided by Polarimetric Echo Classification of Hydrometeor Types, Journal of Applied Meteorology and Climatology, 50, 1389–1402, https://doi.org/10.1175/2011JAMC2622.1, http://journals.ametsoc.org/doi/abs/10.1175/2011JAMC2622.1,
10   2011.

Zhang, J., Qi, Y., Howard, K., Langston, C., and Kaney, B.: Radar quality index (RQI)—A combined measure of beam blockage and VPR effects in a national network, in: Proceedings, International Symposium on Weather Radar and Hydrology, http://www.nssl.noaa.gov/projects/q2/tutorial/images/mosaic/WRaH_Proceedings_Zhang-et-al_v3.pdf, 2011.

Zhong, L., Yang, R., Wen, Y., Chen, L., Gou, Y., Li, R., Zhou, Q., and Hong, Y.: Cross-evaluation of reflectivity from the space-
15   borne precipitation radar and multi-type ground-based weather radar network in China, Atmospheric Research, 196, 200–210, https://doi.org/10.1016/j.atmosres.2017.06.016, http://linkinghub.elsevier.com/retrieve/pii/S0169809516306913, 2017.