



Robustness of spaceborne–ground radar reflectivity comparisons to conditioning coordinate and observational support: application to FY-3G PMR and an S-band radar network

Ping Li¹, Wei Wei², Yonghua Zhang^{2,3*}, Yinhe Zhang^{2*}, and Shuo Liu¹

¹Liaoning Provincial Meteorological Observatory, Shenyang 110016, China

²Guangdong Meteorological Service Center, Guangzhou, China

³State Key Laboratory of Severe Weather Meteorological Science and Technology, Beijing 100081, China

Correspondence to: Yonghua Zhang (zhangyh@gd121.cn); Yinhe Zhang (zhangyinhe@gz121.com)

Abstract. Spaceborne–ground radar reflectivity comparisons can be sensitive to how reflectivity level is represented, how comparable observations are selected, and how uneven observational contributions are aggregated. Using warm-season observations from 2024 and 2025, we volume-matched Fengyun-3G (FY-3G) Precipitation Measurement Radar (PMR) observations with those from five operational S-band radars in Liaoning Province, China, yielding more than 1.6 million matched reflectivity pairs from 146 independent FY-3G overpasses. The overall PMR–S-band reflectivity difference varied with radar composition, aggregation scheme, and warm season. Using exactly the same matched pairs, the reflectivity-dependent slope changed from $-0.532 \text{ dB dBZ}^{-1}$ under conventional ground-radar conditioning to $-0.101 \text{ dB dBZ}^{-1}$ under the paired-mean coordinate, and the weaker negative dependence was reproduced across observational subsets and radar sites. Stratiform and convective precipitation were then compared within the same radar–overpass task and paired-mean reflectivity interval. The task-balanced convective-minus-stratiform contrast was 4.06 dB and decreased to 2.59 dB when both reflectivities were restricted to at least 15 dBZ, while remaining predominantly positive. These results show that conditioning-coordinate choice, common observational support, and aggregation level materially influence the magnitude and interpretation of spaceborne–ground radar reflectivity comparisons.

1. Introduction

Spaceborne precipitation radars provide three-dimensional observations of precipitating systems (Hou et al., 2014; Skofronick-Jackson et al., 2017; Battaglia et al., 2020), while ground-based radar networks provide dense regional sampling in space and time. Quantitative comparisons between the two systems are widely used for validation and consistency assessment (Schwaller and Morris, 2011; Warren et al., 2018; Keem et al., 2019). Early TRMM-era studies demonstrated the use of spaceborne precipitation-radar observations for ground-radar calibration and reflectivity intercomparison (Anagnostou et al., 2001; Wang and Wolff, 2009), and later work extended such comparisons to operational multi-radar networks and GPM observations (Gabella et al., 2017; Zhong et al., 2017). Since the Tropical Rainfall Measuring Mission (TRMM) and Global Precipitation



Measurement (GPM) eras, temporal collocation and volume matching have been routinely applied to improve observational comparability (Bolen and Chandrasekar, 2003; Schwaller and Morris, 2011; Biswas and Chandrasekar, 2018). However, volume matching addresses only part of the comparability problem. Even after geometric and temporal matching, residual measurement and sampling differences can remain (Warren et al., 2018; Keem et al., 2019). Beyond matching, the inferred relationships may also depend on how reflectivity level is conditioned, how comparable samples are defined, and how uneven radar and overpass contributions are aggregated.

Fengyun-3G (FY-3G) was launched in 2023, and its Precipitation Measurement Radar (PMR) established China's capability for spaceborne active precipitation-radar observations (Zhang et al., 2023). Recent studies have evaluated PMR using Chinese operational weather radars, the U.S. Next Generation Weather Radar (NEXRAD) network, and the GPM Dual-frequency Precipitation Radar (DPR) from regional, multi-radar, network, and cross-satellite perspectives. These studies generally report good correspondence between PMR and ground-radar reflectivity, together with differences related to reflectivity level, precipitation type, bright-band position, radar frequency, and observing conditions (He et al., 2025; Chen et al., 2025; Hu et al., 2025; Shang et al., 2026; Zhang et al., 2026). Comparisons with GPM DPR have also examined dual-frequency consistency and detection sensitivity (Liu et al., 2024, 2025). The magnitude and main observational dependencies of PMR–ground-radar differences have therefore been increasingly documented. Less attention has been given to the robustness of the inferred relationships themselves. In particular, it remains unclear how much they depend on the coordinate used to represent reflectivity level, the observational support of the compared samples, and the way uneven radar and overpass contributions are aggregated.

Reflectivity-dependent differences are often examined as a function of ground-radar reflectivity. In this formulation, ground-radar reflectivity is used both to construct the difference and to represent reflectivity level. The shared use of the same measurement can contribute to an apparent difference–magnitude relationship through measurement-error coupling. Measurement-method comparison studies have shown that plotting the difference between two measurements against either component measurement can induce such dependence, whereas the paired mean provides a symmetric coordinate for representing measurement level (Bland and Altman, 1995). The paired mean is not an error-free reference, but it does not privilege either measurement in defining the conditioning coordinate. By holding the matched-pair membership fixed and changing only the conditioning coordinate, the conventional ground-radar formulation can therefore be compared directly with the paired-mean formulation to test whether the inferred reflectivity dependence is sensitive to coordinate choice. This distinction is important when difference–reflectivity relationships are interpreted as evidence of calibration nonlinearity or used to derive intensity-dependent corrections.

A related comparability problem arises in stratiform–convective comparisons. The two regimes differ in reflectivity structure and bright-band characteristics (Awaka et al., 2016; Le et al., 2016) and occupy different reflectivity ranges in the present matched dataset. Their pooled comparison can therefore reflect not only regime-related differences but also unequal radar and overpass contributions. To reduce this confounding, regime contrasts should be evaluated under common observational support defined by radar, overpass, and reflectivity level. In addition, because matched pairs are unevenly distributed among tasks and



radar sites, the resulting summaries can depend on the level at which observations are aggregated. The present study therefore examines regime contrasts under common support and tests their sensitivity to sample-, task-, and radar-level weighting.

65 We use FY-3G PMR observations from the warm seasons of 2024 and 2025 together with five operational S-band radars in Liaoning Province, China. The overall PMR–S-band reflectivity correspondence is first described as a baseline. We then compare reflectivity-dependent relationships obtained from exactly the same matched pairs using ground-radar reflectivity and the paired mean as conditioning coordinates. Stratiform–convective differences are examined under common radar, overpass, and reflectivity support. Finally, the sensitivity of the main comparison metrics to observational aggregation and warm-season
70 composition is assessed. Neither PMR nor the ground radars are treated as an absolute reflectivity reference. The focus is on which comparison features remain stable under these analysis choices.

2. Materials and methods

2.1. FY-3G PMR and ground-based S-band radar data

The analysis covered the warm seasons (May–October) of 2024 and 2025. Five operational S-band weather radars in Liaoning
75 Province—Shenyang, Dalian, Dandong, Yingkou, and Huludao—were used for the quantitative comparison. The Chaoyang C-band radar was included only for regional network context and was not used in the quantitative analysis. Diagnostic information on ground-radar clutter, anomalous propagation, low copolar correlation, and possible beam blockage was retained for quality assessment. The ground-based radars were used as an independent comparison system rather than as an absolute reflectivity reference. The radar locations and the spatial and radar-specific observational support of the CORE sample are shown
80 in Fig. 1.

Spaceborne observations were taken from the FY-3G PMR Level-2 products. Instrument configuration and precipitation-product characteristics are described by Zhang et al. (2023). The primary variable used in the reflectivity comparison was the official frequency-corrected S-band-equivalent reflectivity (FRE; `zFactorFrequencyCorrectionS`; NSMC, 2023; Wu et al., 2023), which was used throughout the primary PMR–S-band reflectivity analysis. Precipitation regime was identified from the
85 `typePrecip` field in the KuR-CSF product group, with codes 1 and 2 representing stratiform and convective precipitation, respectively (NSMC, 2023). This classification was used to construct the regime-specific samples and was not treated as an independent ground-based precipitation classification.

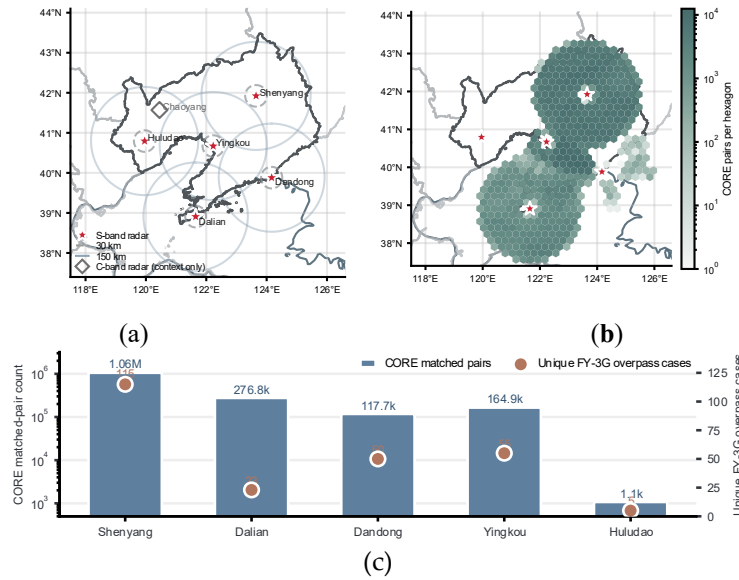


Figure 1: Study area, radar network, and observational support of the CORE dataset over Liaoning Province. (a) Locations of the five S-band radars used in the quantitative analysis and their 30- and 150-km range rings; the Chaoyang C-band radar is shown for regional network context only. (b) Spatial distribution of CORE matched reflectivity pairs, with color indicating the number of pairs per hexagonal grid cell. (c) Numbers of CORE matched reflectivity pairs and independent FY-3G overpass cases for each radar. Administrative boundary data were derived from the ChinaAdminDivisonSHP dataset, which is based on the administrative-region query of the AMap Web Service API (GaryBikini, 2024).

2.2 Quality control, volume matching, and sample construction

Spaceborne and ground-based radar observations were collocated using a volume-matching approach following Bolen and Chandrasekar (2003), Schwaller and Morris (2011), and Biswas and Chandrasekar (2018). This procedure reduced major differences in sampling time, horizontal position, and sampling height between the two observing systems before reflectivity comparison. Figure 2 summarizes the workflow.

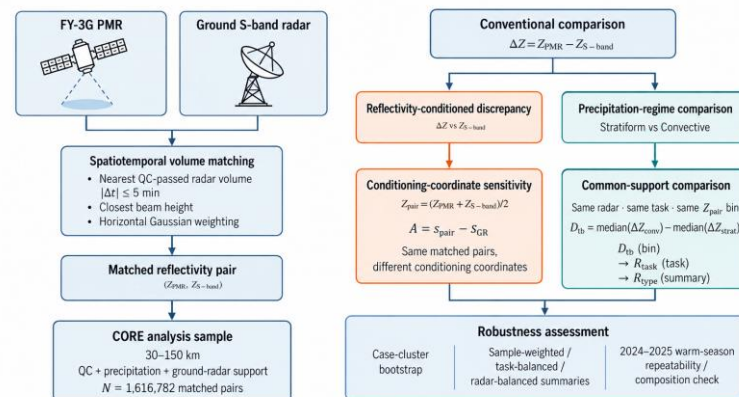


Figure 2: Framework for FY-3G PMR and ground-based S-band radar reflectivity intercomparison. The left part shows spatiotemporal volume matching and construction of the CORE analysis sample. The right part summarizes the conditioning-coordinate, common-support precipitation-regime, and robustness analyses.

One FY-3G overpass paired with one ground-based radar was defined as a radar–overpass task. For each task, the ground-
 105 radar volume nearest in time to the satellite overpass was selected, with $|Δt_{\text{case}}| \leq 5$ min. Tasks outside this window were
 excluded. When one FY-3G overpass intersected more than one radar domain, the corresponding tasks were assigned to the
 same overpass case. PMR sampling cells with valid geolocation, S-band-equivalent reflectivity between -10 and 80 dBZ, and
 sampling heights between 0 and 20 km were retained for volume matching. Ground-radar gates with finite reflectivity between
 -20 and 75 dBZ were treated as valid. For each PMR sampling cell, the ground-radar elevation sweep with beam height closest
 110 to the PMR sampling height was selected, and matches with a vertical separation greater than 1.5 km were discarded.

Within the selected sweep, valid ground-radar gates within 5 km of the PMR sampling position were combined using horizontal
 Gaussian distance weighting,

$$w_i = \exp\left(-\frac{d_i^2}{2\sigma^2}\right), \quad \sigma = 2.5 \text{ km} \quad (1)$$

where d_i is the horizontal distance between ground-radar gate i and the PMR sampling location. Reflectivity was averaged in
 linear reflectivity-factor space and converted back to dBZ:

$$Z_{\text{S-band}} = 10\log_{10}\left[\frac{\sum_i w_i 10^{Z_i/10}}{\sum_i w_i}\right] \quad (2)$$

115 where Z_i is the reflectivity of ground-radar gate i in dBZ.

Each matched value therefore represents a horizontally weighted estimate from the single elevation sweep closest to the PMR
 sampling altitude. After removal of pairs with non-finite weighted ground-radar reflectivity, $3,479,480$ valid matched pairs
 formed the ALL sample. The CORE and STABLE samples were derived from ALL using the criteria in Table 1.

Table 1: Definitions and screening criteria for the ALL, CORE, and STABLE samples.

Sample	Role	Screening criteria	Matched pairs (N)
ALL	Valid paired-reflectivity sample	Finite paired PMR and weighted ground-radar reflectivity from the five S-band radars.	3,479,480
CORE	Primary analysis sample	ALL restricted to 30 – 150 km ground-radar range; require <code>flagPrecip = 1</code> , <code>qualitySLV = 0</code> , at least 3 valid ground-radar gates, and valid-gate fraction $\geq 20\%$.	1,616,782
STABLE	Sensitivity-analysis sample	CORE with shallow precipitation and samples within the PMR bright band additionally excluded.	1,575,047



- 120 Shallow precipitation was identified from flagShallowRain. When a bright band was identified and both heightBB and widthBB were valid, a PMR sample with sampling height h_{PMR} satisfying

$$\text{heightBB} - \text{widthBB}/2 \leq h_{\text{PMR}} \leq \text{heightBB} + \text{widthBB}/2$$

was classified as lying within the bright band. The matching and screening procedures reduce major differences in sampling geometry and product quality, but do not eliminate them completely.

2.3 Intercomparison diagnostics and statistical analysis

- 125 For each matched pair,

$$\Delta Z = Z_{\text{PMR}} - Z_{\text{S-band}} \quad (3)$$

where Z_{PMR} is FY-3G PMR S-band-equivalent reflectivity and $Z_{\text{S-band}}$ is the volume-matched ground-radar reflectivity. Reflectivity is expressed in dBZ and ΔZ in dB. Positive ΔZ indicates larger PMR reflectivity for the matched pair.

Seven predefined samples were used to separate baseline statistics from sensitivity to reflectivity threshold, precipitation regime, shallow precipitation, and bright-band screening (Table 2).

130

Table 2: Predefined samples used in the intercomparison and sensitivity analyses.

Sample	Definition	Purpose
B0	CORE	Baseline
B1	CORE; $Z_{\text{PMR}} \geq 15$ dBZ and $Z_{\text{S-band}} \geq 15$ dBZ	Reflectivity-threshold sensitivity
B2	CORE stratiform	Stratiform
B3	CORE stratiform; $Z_{\text{PMR}} \geq 15$ dBZ and $Z_{\text{S-band}} \geq 15$ dBZ	Stratiform sample with reflectivity threshold
B4	STABLE stratiform; $Z_{\text{PMR}} \geq 15$ dBZ and $Z_{\text{S-band}} \geq 15$ dBZ	Additional shallow-rain and bright-band screening
B5	CORE convective	Precipitation-regime sensitivity
B6	CORE convective; $Z_{\text{PMR}} \geq 15$ dBZ and $Z_{\text{S-band}} \geq 15$ dBZ	Convective sample with reflectivity threshold

2.3.1 Baseline comparison metrics

Baseline reflectivity correspondence was summarized using the mean and median of ΔZ , interquartile range (IQR), mean absolute difference (MAE), root-mean-square difference (RMSE), and Pearson correlation coefficient r . These statistics were calculated for CORE and the predefined samples in Table 2, and differences among the samples were summarized descriptively.

- 135 **2.3.2 Conditioning-coordinate sensitivity**

Reflectivity dependence was first examined using $Z_{\text{S-band}}$ as the conditioning variable. Median ΔZ and IQR were calculated in 2-dBZ bins. Slope estimation was restricted to $15 \leq Z_{\text{S-band}} < 45$ dBZ. A Theil–Sen estimator was fitted to the binwise median



ΔZ values (Sen, 1968), giving the conventional slope s_{GR} in units of dB dBZ⁻¹. A symmetric conditioning coordinate was then defined as

$$Z_{pair} = \frac{Z_{PMR} + Z_{S-band}}{2}, \quad (4)$$

following the measurement-comparison rationale of Bland and Altman (1995). The corresponding slope was denoted s_{pair} , and conditioning-coordinate sensitivity was defined as

$$A = s_{pair} - s_{GR} \quad (5)$$

Thus, A quantifies the change in the fitted reflectivity dependence between the two conditioning coordinates. The two slopes were estimated from exactly the same matched pairs. Common fitting support required both $15 \leq Z_{S-band} < 45$ dBZ and $15 \leq Z_{pair} < 45$ dBZ. The retained pairs were then binned separately according to the two coordinates. Pair membership was identical between the fits, although an observation could fall into different 2-dBZ bins. When $s_{GR} < 0$, positive A indicates a weaker negative dependence under Z_{pair} conditioning.

The exact-common-pair analysis was performed at network, radar-overpass-task, and radar levels. Network-level estimates required at least eight valid bins, at least 500 matched pairs per bin, and a reflectivity span of at least 16 dBZ. Task-level estimates required at least six valid bins, at least 30 pairs per bin, and a span of at least 12 dBZ. Radar-level estimates required at least eight valid bins, at least 200 pairs per bin, and a span of at least 16 dBZ. The analysis was applied to B0–B6. B3 and its below-bright-band subset were also examined. At the independent-overpass level, A_{case} was defined as the median task-level A across eligible radar tasks belonging to the same FY-3G overpass.

2.3.3 Precipitation-regime contrast under common observational support

Stratiform and convective samples differ in reflectivity distribution and in their contributions from individual radars and overpasses. Regime contrasts were therefore evaluated within common support defined jointly by radar-overpass task and reflectivity level. The analysis used 2-dBZ Z_{pair} bins over $15 \leq Z_{pair} < 45$ dBZ. A task-bin cell was retained only when both stratiform and convective precipitation contained at least 30 matched pairs within the same radar-overpass task and the same Z_{pair} bin. The two regimes were therefore compared for the same radar, the same FY-3G overpass, and a similar symmetric reflectivity level. For each retained task t and bin b ,

$$D_{tb} = \text{median}(\Delta Z_{conv,t,b}) - \text{median}(\Delta Z_{strat,t,b}), \quad (6)$$

Positive D_{tb} indicates a larger PMR–S-band reflectivity difference for convective precipitation within the same common-support cell. Two common-support samples were used. P0 was based on CORE. P1 additionally required both Z_{PMR} and Z_{S-band} to be at least 15 dBZ. A radar-overpass task was retained only when at least four comparable Z_{pair} bins were available over a span of at least 8 dBZ. For each eligible task,



$$R_{\text{task}} = \text{median}_b(D_{tb}), \quad (7)$$

and the task-balanced network estimate was

$$R_{\text{type}} = \text{median}_t(R_{\text{task}}), \quad (8)$$

Variability among independent FY-3G overpasses was summarized using R_{case} , defined as the median R_{task} among eligible radar tasks from the same overpass. This procedure reduces major differences in radar, overpass, and reflectivity support between the two regimes. Residual contrasts can still include effects of three-dimensional precipitation structure and unresolved measurement or sampling differences (Grecu et al., 2016; Warren et al., 2018; Keem et al., 2019).

2.3.4 Robustness and uncertainty assessment

Robustness was examined in terms of observational aggregation and sampling uncertainty. Three aggregation schemes were used: sample-weighted, task-balanced, and radar-balanced. Sample-weighted estimates retain the unequal numbers of matched pairs contributed by individual tasks, task-balanced estimates give equal weight to eligible radar–overpass tasks, and radar-balanced estimates give equal weight to radars with sufficient support. The three schemes were applied to the overall reflectivity difference ΔZ , the conditioning-coordinate metric A , and the precipitation-regime contrast, with median-based summaries calculated at the corresponding pair, task, and radar levels. Warm-season composition was examined separately for the May–October 2024 and May–October 2025 CORE samples, focusing on overall ΔZ , radar contribution, overpass composition, and conditioning-coordinate response.

Sampling uncertainty was estimated using a cluster bootstrap with independent FY-3G overpass cases as the resampling unit (Field and Welsh, 2007). All matched observations and radar–overpass tasks belonging to the same overpass were resampled together. Unless otherwise stated, 1000 bootstrap replicates were used to obtain percentile-based 95 % confidence intervals.

3. Results

3.1. Baseline reflectivity correspondence and sample dependence

The CORE sample comprised 1,616,782 FY-3G PMR–S-band matched reflectivity pairs from 146 independent FY-3G overpasses and 248 radar–overpass tasks. PMR and S-band reflectivities were positively correlated overall, with some variation across reflectivity levels (Fig. 3a). The network-wide mean and median ΔZ were 3.57 and 3.25 dB, respectively, and the interquartile range was −1.67 to 8.33 dB (Fig. 3b).

The ΔZ distributions differed among the five S-band radars (Fig. 3c). Shenyang, Dandong, and Yingkou had positive central differences, Dalian was close to zero, and Huludao had a negative difference. The distribution spread also varied among radars.

These differences, together with the unequal radar contributions shown in Fig. 1, indicate that the pooled network statistics are influenced by radar composition.

At the network scale, the mean ΔZ had a 95 % overpass-cluster bootstrap confidence interval of 2.84–4.38 dB (Fig. 3d). MAE and RMSE were 7.09 and 9.36 dB, with 95 % confidence intervals of 6.58–7.63 and 8.66–10.09 dB, respectively. The Pearson correlation coefficient was 0.489, with a 95 % confidence interval of 0.405–0.566.

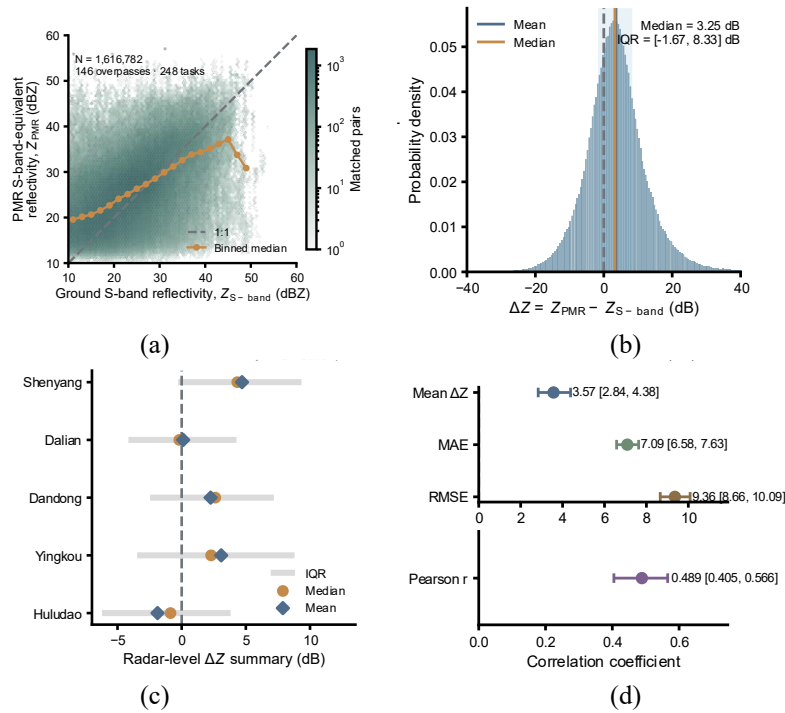


Figure 3: Overall reflectivity comparison for the CORE sample between FY-3G PMR and the five-radar S-band network. (a) Joint distribution of FY-3G PMR S-band-equivalent reflectivity (Z_{PMR}) and volume-matched ground-based S-band reflectivity (Z_{S-band}); color shading denotes matched-pair density, the dashed line denotes 1:1 correspondence, and connected symbols show median PMR reflectivity in 2-dBZ Z_{S-band} bins. (b) Probability-density distribution of ΔZ ; vertical lines denote zero, the mean, and the median, and the annotation gives the interquartile range. (c) Radar-level distributions of ΔZ ; gray segments denote the interquartile range, circles the median, and diamonds the mean. (d) Network-scale mean ΔZ , MAE, RMSE, and Pearson r , with 95% confidence intervals obtained by FY-3G-overpass case-cluster bootstrap.

3.2. Conditioning-coordinate sensitivity of reflectivity dependence

The conditioning-coordinate comparison was first examined using B3, defined as the CORE stratiform sample with both Z_{PMR} and Z_{S-band} restricted to at least 15 dBZ. This sample provides a relatively homogeneous precipitation regime while excluding weak-reflectivity pairs. Application of the exact-common-pair criteria retained 969,567 matched pairs, corresponding to 99.82 % of the parent B3 sample. Under conventional Z_{S-band} conditioning, the slope was $s_{GR} = -0.532$ dB dBZ⁻¹. Using the same matched pairs but conditioning on the symmetric Z_{pair} coordinate, the slope weakened to $s_{pair} = -0.101$ dB dBZ⁻¹ (Fig. 4a,b). The resulting conditioning-coordinate metric was $A = s_{pair} - s_{GR} = 0.430$ dB dBZ⁻¹, with a 95 % FY-3G-



overpass cluster-bootstrap confidence interval of 0.328–0.702 dB dBZ⁻¹ (Fig. 4c,d). A similar response was found for the B3 subset below the PMR bright band. In this subset, s_{GR} was -0.507 dB dBZ⁻¹, s_{pair} was 0.005 dB dBZ⁻¹, and A was 0.513 dB dBZ⁻¹ (95 % CI: 0.384–0.755 dB dBZ⁻¹).

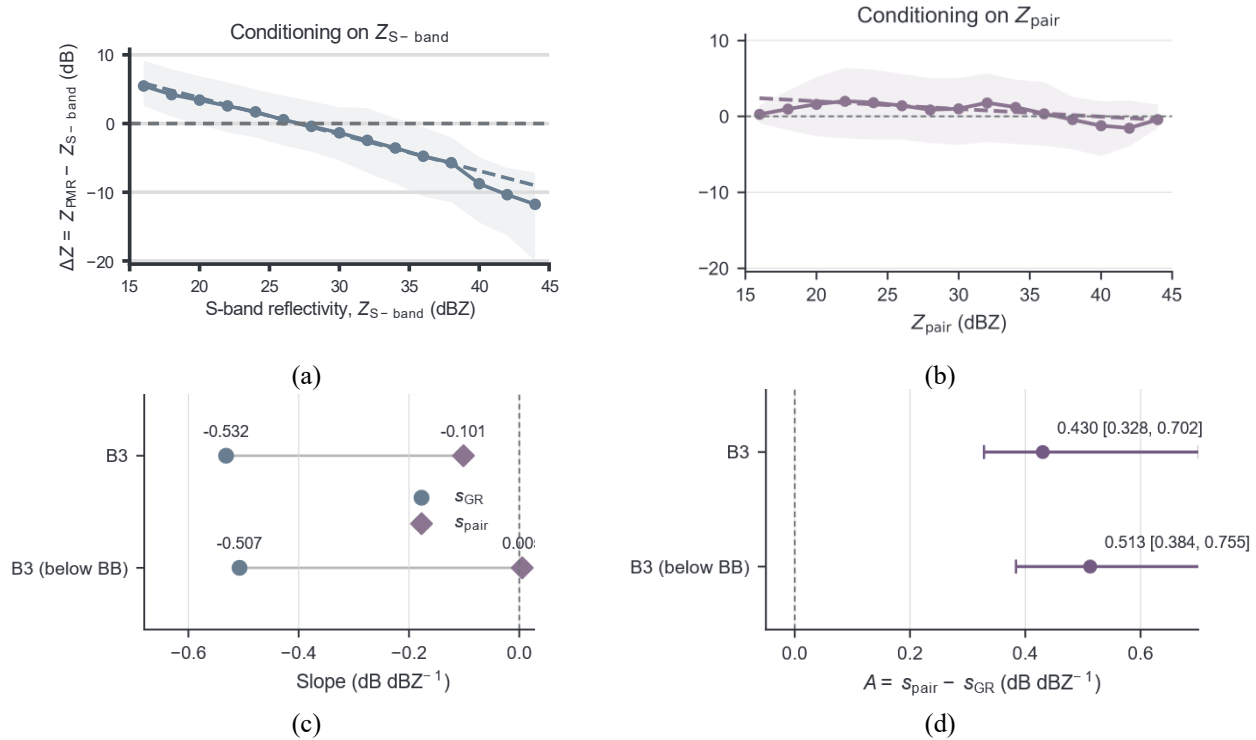


Figure 4: Conditioning-coordinate sensitivity of the reflectivity-dependent PMR–S-band difference. Panels (a) and (b) use exactly the same individual B3 matched reflectivity pairs satisfying both $15 \leq Z_{S\text{-band}} < 45$ dBZ and $15 \leq Z_{pair} < 45$ dBZ. The common observations are then binned separately according to the respective conditioning coordinate. (a) ΔZ conditioned on $Z_{S\text{-band}}$, with the corresponding Theil–Sen fit. (b) ΔZ conditioned on the symmetric diagnostic coordinate Z_{pair} , with the corresponding Theil–Sen fit. Shading in (a,b) denotes the interquartile range. (c) Paired s_{GR} and s_{pair} estimates for B3 and the below-bright-band subset of B3. (d) $A = s_{pair} - s_{GR}$ with FY-3G-overpass case-cluster bootstrap 95% confidence intervals. Z_{pair} is used only as a symmetric diagnostic coordinate.

Positive A values were obtained for all seven predefined samples under exact common-pair support (Fig. 5a), with network-scale estimates ranging from 0.430 to 0.580 dB dBZ⁻¹. Differences between the two conditioning-coordinate fits were larger for B0, B2, and B5, which included more low-reflectivity observations, and smaller for B3 and B4.

At the task level, most B3 estimates lay above the equal-slope diagonal in the s_{GR} – s_{pair} plane (Fig. 5b). A was positive in 100 of 104 eligible radar–overpass tasks (96.2 %). The task-level A distributions for B3 and the below-bright-band subset were both centered above zero (Fig. 5c), with positive values in 48 of 51 below-bright-band tasks (94.1 %). At the independent-overpass level, $A_{case} > 0$ for 73 of 76 B3 overpasses (96.1 %) and 38 of 40 below-bright-band overpasses (95.0 %). The corresponding median A_{case} values were 0.438 and 0.513 dB dBZ⁻¹.

Radar-level A also remained positive wherever the support criteria were met (Fig. 5d). For B3, A was 0.476, 0.344, 0.795, and 0.716 dB dBZ⁻¹ at Shenyang, Dalian, Dandong, and Yingkou, respectively. Across B0–B6, Dandong and Yingkou generally



230 had larger A than Shenyang and Dalian. Support was more limited for the convective B5 and B6 samples, particularly at Dalian, where only four tasks met the radar-level criteria. Huludao did not meet the radar-level support requirement for any predefined sample.

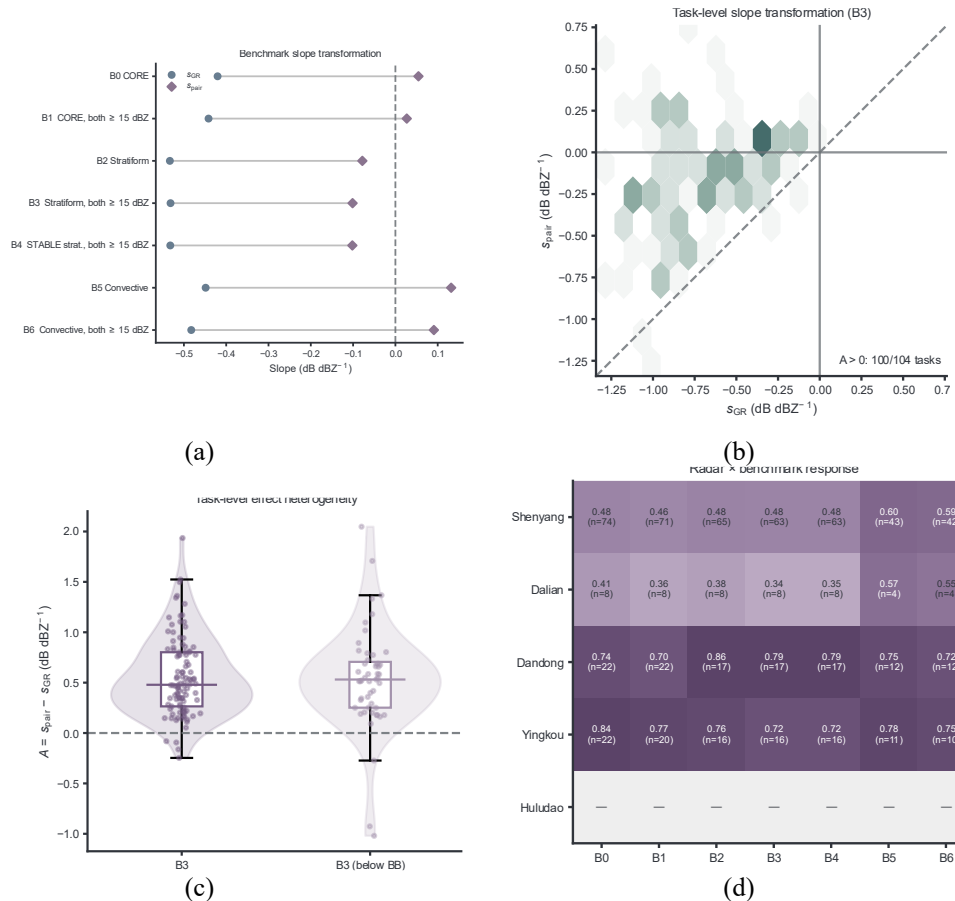


Figure 5: Cross-sample, task-level, and radar-level variability of conditioning-coordinate sensitivity under exact common-pair support. (a) Network-scale paired s_{GR} and s_{pair} estimates across B0–B6; their difference represents A. (b) Task-level joint distribution of s_{GR} and s_{pair} for B3; the diagonal denotes equal slopes and color denotes task count. (c) Task-level distributions of A for B3 and its below-bright-band subset. (d) Radar × sample-set matrix of radar-level A; annotations give A and the number of supporting tasks, while gray cells indicate insufficient support. All paired slope comparisons use the same individual matched pairs for the two conditioning coordinates.

3.3. Precipitation-regime contrast under common observational support

240 To improve comparability between stratiform and convective precipitation, the two regimes were compared only within common-support cells defined by the same radar, the same FY-3G overpass, and the same 2-dBZ Z_{pair} bin. This restriction reduces differences in radar contribution, overpass composition, and reflectivity range between the two regimes. Under these criteria, P0 retained 84 radar–overpass tasks from 65 independent FY-3G overpasses, while P1 retained 75 tasks from 58 overpasses. Support varied with Z_{pair} (Fig. 6a). For P0, the largest number of supporting tasks occurred between about 18 and 26 dBZ.



245 Requiring both reflectivities to be at least 15 dBZ in P1 mainly reduced support at the low-reflectivity end. Above about 26 dBZ, P0 and P1 had similar task support. In both samples, support decreased rapidly toward higher reflectivity, with only a few tasks remaining above 40 dBZ.

Within the P0 common-support cells, both stratiform and convective ΔZ generally decreased with increasing Z_{pair} , while convective values were higher over most of the supported range (Fig. 6b). Considerable spread remained among task-bin cells.
 250 The within-cell convective-minus-stratiform contrast D_{cb} was positive in 14 of 15 Z_{pair} bins for P0 and 11 of 15 bins for P1 (Fig. 6c). For P0, median D_{cb} was generally about 4–6 dB over 22–40 dBZ, with larger values near 28–30 dBZ. For P1, D_{cb} was close to zero or slightly negative in the lowest-reflectivity bins and positive at moderate and higher reflectivity.

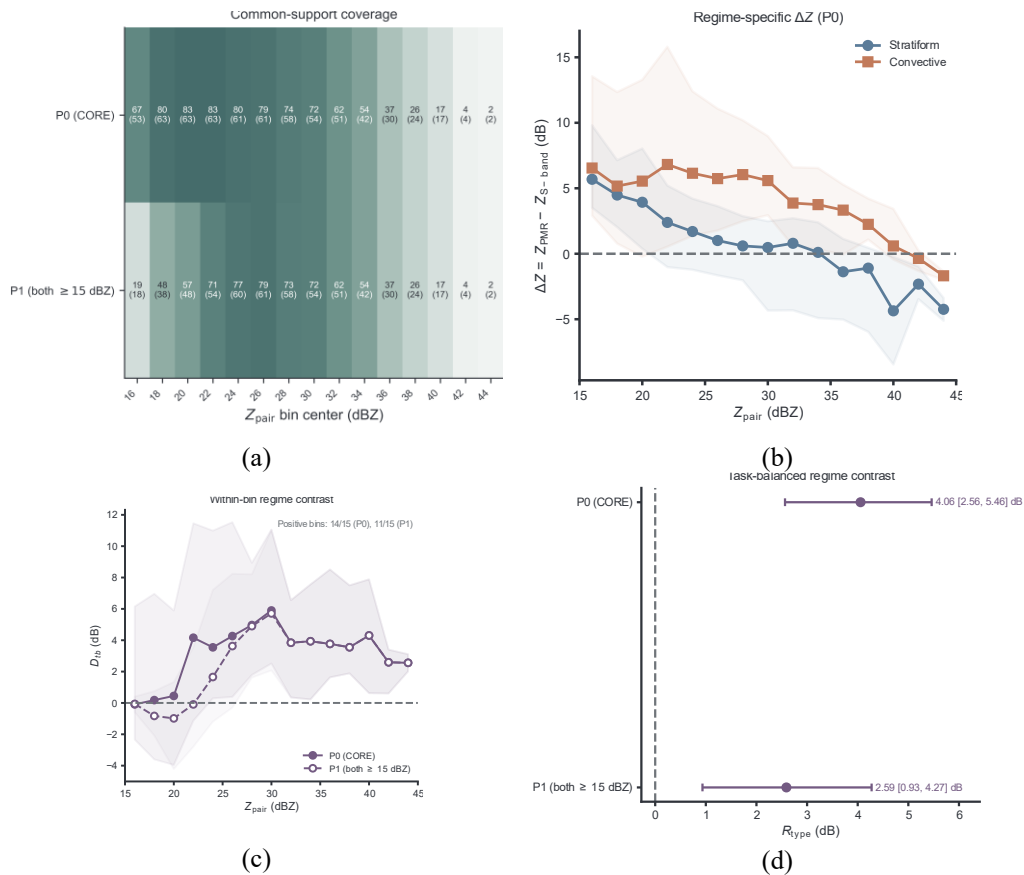


Figure 6: Precipitation-regime-related reflectivity differences under common observational support. Common support is defined within the same radar, the same FY-3G overpass task, and the same 2-dBZ Z_{pair} bin. (a) Numbers of supporting tasks across Z_{pair} bins for P0 and P1; numbers in parentheses denote independent FY-3G overpass cases. (b) Median stratiform and convective ΔZ within P0 common-support cells; shading denotes the interquartile range across supported task-bin cells. (c) Median within-cell convective-minus-stratiform difference D_{cb} as a function of Z_{pair} for P0 and P1; shading denotes the interquartile range. (d) Task-balanced network estimates R_{type} with FY-3G-overpass case-cluster bootstrap 95% confidence intervals. P0 uses CORE; P1 additionally requires both reflectivities to be ≥ 15 dBZ.



260 After aggregation across eligible tasks, the task-balanced regime contrast R_{type} was 4.06 dB for P0, with a 95 % confidence interval of 2.56–5.46 dB (Fig. 6d). For P1, R_{type} was 2.59 dB, with a 95 % confidence interval of 0.93–4.27 dB.

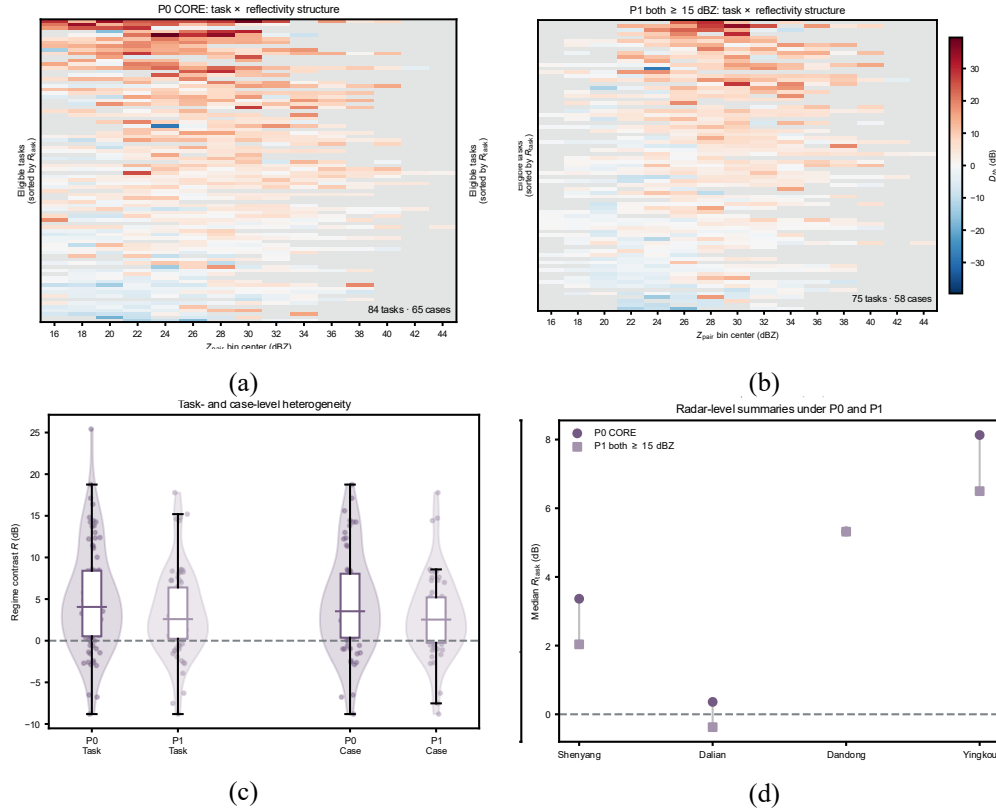


Figure 7: Task-, case-, and radar-level heterogeneity of precipitation-regime-related reflectivity differences. (a,b) Task × Z_{pair} matrices of D_{tb} for P0 and P1 using the same zero-centered diverging color scale; gray cells indicate insufficient common support. Tasks are ordered separately by R_{task} within each sample set, so rows in (a) and (b) do not represent paired task identities. (c) Distributions of R_{task} and R_{case} for P0 and P1. (d) Radar-level median R_{task} for P0 and P1 across radars with sufficient support.

Regime contrasts varied among individual radar–overpass tasks and reflectivity intervals (Fig. 7a,b). Both positive and negative D_{tb} values occurred, and their magnitude and reflectivity dependence differed among tasks. Positive R_{task} values occurred in 68 of 84 P0 tasks (81.0 %) and 59 of 75 P1 tasks (78.7 %). At the independent-overpass level, $R_{case} > 0$ for 51 of 65 P0 cases (78.5 %) and 43 of 58 P1 cases (74.1 %) (Fig. 7c). The task- and case-level distributions crossed zero, but positive values were more frequent.

Radar-level summaries also varied among sites (Fig. 7d). All four radars with sufficient support had positive median R_{task} for P0. Under P1, Shenyang, Dandong, and Yingkou remained positive, whereas Dalian was close to zero and slightly negative. The largest radar-level contrasts occurred at Yingkou and Dandong.

3.4. Robustness to aggregation and warm-season composition

Summary estimates varied with the observational unit used for aggregation (Fig. 8a,b). The median-based ΔZ summaries were 3.25, 3.77, and 2.74 dB under sample-weighted, task-balanced, and radar-balanced aggregation, respectively. The conditioning-coordinate metric A remained positive under all three schemes, with values of 0.475 dB dBZ⁻¹ (95 % CI: 0.372–0.720), 0.673 dB dBZ⁻¹ (0.586–0.851), and 0.610 dB dBZ⁻¹ (0.524–0.866), respectively. The corresponding precipitation-regime contrasts R_{type} were 3.24, 4.06, and 4.34 dB.

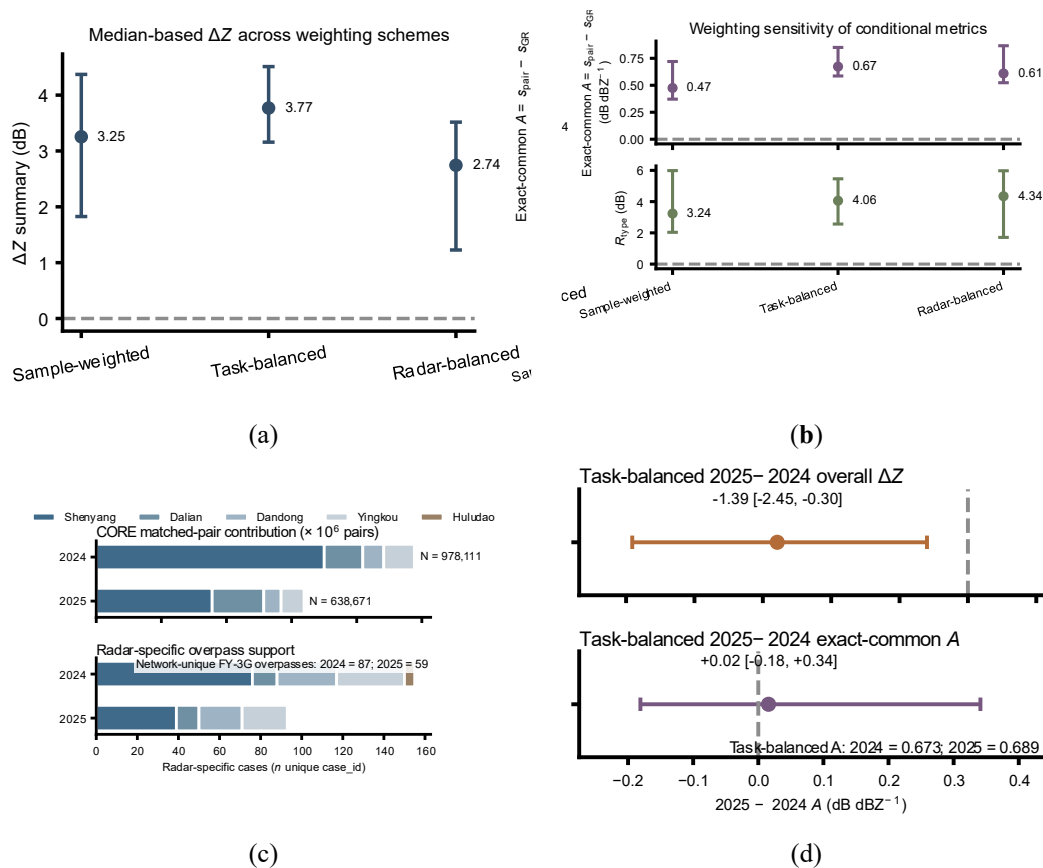


Figure 8: Sensitivity of the overall reflectivity difference, conditioning-coordinate metric, and precipitation-regime difference to statistical weighting and warm-season composition. (a) Median-based ΔZ summaries under sample-weighted, task-balanced, and radar-balanced aggregation; error bars denote 95% confidence intervals. (b) Exact-common conditioning-coordinate metric A and precipitation-regime metric R_{type} under the same weighting schemes, with 95% confidence intervals. (c) Radar contributions to CORE matched-pair counts and radar-specific FY-3G overpass support in 2024 and 2025. (d) Signed task-balanced 2025-minus-2024 differences in overall ΔZ and A , with 95% confidence intervals.

Sample size and radar contribution also differed between the two warm seasons (Fig. 8c). CORE contained 978,111 matched pairs from 87 network-unique FY-3G overpasses in 2024 and 638,671 pairs from 59 overpasses in 2025. The relative contributions of individual radars and their overpass support also changed between years.



Under task-balanced aggregation, the overall ΔZ was lower in 2025 than in 2024. The 2025-minus-2024 difference was -1.39 dB, with a 95 % confidence interval of -2.45 to -0.30 dB (Fig. 8d). By comparison, the task-balanced conditioning-coordinate metric was similar between years: $A=0.673$ dB dBZ $^{-1}$ in 2024 and 0.689 dB dBZ $^{-1}$ in 2025. The corresponding difference was 0.016 dB dBZ $^{-1}$ (95 % CI: -0.181 to 0.341 dB dBZ $^{-1}$).

4. Discussion

Previous FY-3G PMR evaluations have reported reflectivity-dependent differences between PMR and ground-based radar observations, with variations related to reflectivity level, precipitation type, radar network, and observing conditions (He et al., 2025; Chen et al., 2025; Hu et al., 2025; Shang et al., 2026; Zhang et al., 2026). In the present analysis, using exactly the same matched pairs, the negative slope obtained under conventional ground-radar conditioning became much weaker when the paired mean was used as the conditioning coordinate. Because ground-radar reflectivity enters both ΔZ and the conventional conditioning axis, measurement uncertainty in the ground-radar observation can contribute to the apparent dependence (Bland and Altman, 1995). The paired mean avoids assigning either radar as the conditioning reference, while calibration differences, beam sampling, attenuation, frequency-related effects, and residual collocation differences may still contribute to the remaining relationship (Warren et al., 2018; Keem et al., 2019; Wu et al., 2023; Chen et al., 2025). The positive conditioning-coordinate response was found across the predefined samples, most eligible radar-overpass tasks and independent overpasses, and all radars with sufficient support. It also remained positive under sample-, task-, and radar-balanced aggregation and was similar between the 2024 and 2025 warm seasons. By comparison, the overall ΔZ varied more with aggregation scheme and between years. A reflectivity-dependent slope obtained from a single conditioning coordinate therefore does not fully characterize the dependence between the two measurements.

Previous FY-3G studies have generally reported larger PMR-ground-radar discrepancies for convective than for stratiform precipitation (He et al., 2025; Chen et al., 2025). In the present dataset, the two regimes also differed in reflectivity range, radar contribution, and overpass composition. When the comparison was restricted to the same radar-overpass task and the same Z_{pair} interval, the convective-minus-stratiform contrast remained predominantly positive, although its magnitude decreased after both reflectivities were restricted to at least 15 dBZ. Differences in observational support therefore account for part of the pooled regime contrast. The remaining difference may be associated with three-dimensional precipitation structure, nonuniform beam filling, attenuation, frequency-dependent radar response, and vertical reflectivity gradients (Grecu et al., 2016; Liao and Meneghini, 2019; Meneghini et al., 2021; Wu et al., 2023). The contrast also varied among tasks, independent overpasses, and radar sites, with both positive and negative values occurring locally. Because precipitation type was defined from the PMR classification, the residual contrast represents the regime dependence of the PMR-ground-radar comparison. Taken together, the analysis shows that conditioning-coordinate choice, common observational support, and aggregation level can materially affect both the magnitude and interpretation of spaceborne-ground radar reflectivity comparisons.



5. Conclusions

Using FY-3G PMR observations from the 2024 and 2025 warm seasons and five operational S-band radars in Liaoning Province, this study examined the robustness of spaceborne–ground radar reflectivity comparisons to conditioning coordinate, observational support, and aggregation.

1. The overall PMR–S-band reflectivity difference varied with radar composition, aggregation scheme, and warm season. Network-scale ΔZ therefore depended on the observational composition of the matched dataset.
2. The inferred reflectivity dependence changed with the conditioning coordinate. Using exactly the same matched pairs, the negative ΔZ –reflectivity slope obtained under conventional ground-radar conditioning became much weaker under the paired-mean coordinate. This response was found across predefined samples, most eligible radar–overpass tasks and independent overpasses, and all radars with sufficient support.
3. The convective-minus-stratiform contrast remained predominantly positive after the two regimes were compared within the same radar–overpass task and Z_{pair} interval. Its magnitude decreased under the stricter reflectivity-support definition and varied among tasks, overpasses, and radars, indicating that observational support accounts for part of the pooled regime contrast.

Together, these results show that conditioning-coordinate choice, common observational support, and aggregation level affect the magnitude and interpretation of spaceborne–ground radar reflectivity comparisons.

Code and data availability

The analysis and figure-generation codes used in this study are available from the corresponding author upon reasonable request. The FY-3G Precipitation Measurement Radar (PMR) Level-2 products are publicly available from the official Fengyun Satellite Data Center of the National Satellite Meteorological Center at <https://satellite.nsmc.org.cn/DataPortal/cn/home/index.html>. The ground-based S-band radar observations used in this study are operational meteorological observations provided by the Liaoning Meteorological Service. Under the meteorological data-sharing regulations of the China Meteorological Administration, the authors are not authorized to redistribute these original observational data publicly. Access to the radar data may be requested from the relevant meteorological data provider and is subject to its data-access procedures and approval. The administrative boundary data used in Fig. 1 were derived from the ChinaAdminDivisonSHP dataset (GaryBikini, 2024).

Author contribution

Ping Li and Yonghua Zhang conceived the study and developed the methodology. Ping Li performed the data processing and analysis, developed the analysis code, curated the data, and prepared the figures. Ping Li, Wei Wei, and Yonghua Zhang



performed the validation, and Ping Li, Wei Wei, and Shuo Liu contributed to the investigation. Ping Li and Yonghua Zhang prepared the original manuscript. Yonghua Zhang and Yinhe Zhang supervised the study and coordinated the project. Wei Wei, Yonghua Zhang, and Yinhe Zhang contributed to funding acquisition. All authors contributed to the review and editing of the manuscript and approved the final version.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

The authors acknowledge the National Satellite Meteorological Center of the China Meteorological Administration for providing the FY-3G Precipitation Measurement Radar products, and the Liaoning meteorological service for providing the operational S-band weather radar observations used in this study.

Financial support

This research was funded by the Guangdong Basic and Applied Basic Research Foundation, grant number 2025A1515510029; the Liaoning Province Science and Technology Plan Joint Program (Natural Science Foundation–General Program), grant number 2025-MSLH-369; the State Key Laboratory of Severe Weather, grant number 2022LASW-B06; the Hebei Natural Science Foundation, grant number D2024112003; the Scientific and Technological Research Project of Guangdong Meteorological Bureau, grant number GRMC2023Z02; the Guangdong–Hong Kong–Macao Greater Bay Area Meteorological Science and Technology Collaborative Project, grant number GHMA2024Y05; the Guangzhou Municipal Science and Technology Project, grant number 2025B04J0018; and the Liaoning Meteorological Bureau Talent Program Project, grant number RC202501.

References

- Anagnostou, E. N., Morales, C. A., and Dinku, T.: The use of TRMM Precipitation Radar observations in determining ground radar calibration biases, *J. Atmos. Ocean. Technol.*, 18, 616–628, [https://doi.org/10.1175/1520-0426\(2001\)018<0616:TUOTPR>2.0.CO;2](https://doi.org/10.1175/1520-0426(2001)018<0616:TUOTPR>2.0.CO;2), 2001.
- Awaka, J., Le, M., Chandrasekar, V., Yoshida, N., Higashiawatoko, T., Kubota, T., and Iguchi, T.: Rain type classification algorithm module for GPM Dual-Frequency Precipitation Radar, *J. Atmos. Ocean. Technol.*, 33, 1887–1898, <https://doi.org/10.1175/JTECH-D-16-0016.1>, 2016.



- 375 Battaglia, A., Kollias, P., Dhillon, R., Roy, R., Tanelli, S., Lamer, K., Grecu, M., Lebsock, M., Watters, D., Mroz, K.,
Heymsfield, G., Li, L., and Furukawa, K.: Spaceborne cloud and precipitation radars: Status, challenges, and ways forward,
Rev. Geophys., 58, e2019RG000686, <https://doi.org/10.1029/2019RG000686>, 2020.
- Biswas, S. K. and Chandrasekar, V.: Cross-validation of observations between the GPM Dual-Frequency Precipitation Radar
and ground based dual-polarization radars, Remote Sens., 10, 1773, <https://doi.org/10.3390/rs10111773>, 2018.
- 380 Bland, J. M. and Altman, D. G.: Comparing methods of measurement: why plotting difference against standard method is
misleading, Lancet, 346, 1085–1087, [https://doi.org/10.1016/S0140-6736\(95\)91748-9](https://doi.org/10.1016/S0140-6736(95)91748-9), 1995.
- Bolen, S. M. and Chandrasekar, V.: Methodology for aligning and comparing spaceborne radar and ground-based radar ob-
servations, J. Atmos. Ocean. Technol., 20, 647–659, [https://doi.org/10.1175/1520-0426\(2003\)20<647:MFAACS>2.0.CO;2](https://doi.org/10.1175/1520-0426(2003)20<647:MFAACS>2.0.CO;2),
2003.
- 385 Chen, P., Chen, L., Wu, Q., Wang, G., and Zhang, P.: Analysis of satellite-ground radar reflectivity consistency: First evalua-
tion results of FY-3G and GPM precipitation radar, Geophys. Res. Lett., 52, e2025GL115785,
<https://doi.org/10.1029/2025GL115785>, 2025.
- Field, C. A. and Welsh, A. H.: Bootstrapping clustered data, J. R. Stat. Soc. B, 69, 369–390, <https://doi.org/10.1111/j.1467-9868.2007.00593.x>, 2007.
- 390 Gabella, M., Speirs, P., Hamann, U., Germann, U., and Berne, A.: Measurement of precipitation in the Alps using dual-polar-
ization C-band ground-based radars, the GPM spaceborne Ku-band radar, and rain gauges, Remote Sens., 9, 1147,
<https://doi.org/10.3390/rs9111147>, 2017.
- GaryBikini: *ChinaAdminDivisonSHP*, version 24.02.06, Zenodo, <https://doi.org/10.5281/zenodo.10624971>, 2024.
- Grecu, M., Olson, W. S., Munchak, S. J., Ringerud, S., Liao, L., Haddad, Z., Kelley, B. L., and McLaughlin, S. F.: The GPM
395 combined algorithm, J. Atmos. Ocean. Technol., 33, 2225–2245, <https://doi.org/10.1175/JTECH-D-16-0019.1>, 2016.
- He, R., Li, H., Luo, J., Huang, H., and Zhu, Y.: Comparison of the reflectivities from Precipitation Measurement Radar onboard
the FY-3G satellite and ground-based S-band dual-polarization radars, Remote Sens., 17, 1117,
<https://doi.org/10.3390/rs17071117>, 2025.
- Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., and Iguchi,
400 T.: The Global Precipitation Measurement mission, Bull. Amer. Meteor. Soc., 95, 701–722, <https://doi.org/10.1175/BAMS-D-13-00164.1>, 2014.
- Hu, H., Pei, S., Wu, L., Shao, N., Zhang, C., Wen, H., and Liu, Y.: Design and application of two consistency verification
methods for weather radar networks in the South China region, Atmos. Meas. Tech., 18, 5507–5526,
<https://doi.org/10.5194/amt-18-5507-2025>, 2025.
- 405 Keem, M., Seo, B.-C., Krajewski, W. F., and Morris, K. R.: Inter-comparison of reflectivity measurements between GPM DPR
and NEXRAD radars, Atmos. Res., 226, 49–65, <https://doi.org/10.1016/j.atmosres.2019.04.010>, 2019.
- Le, M., Chandrasekar, V., and Biswas, S.: Evaluation and validation of GPM dual-frequency classification module after launch,
J. Atmos. Ocean. Technol., 33, 2699–2716, <https://doi.org/10.1175/JTECH-D-15-0253.1>, 2016.



- Liao, L. and Meneghini, R.: Physical evaluation of GPM DPR single- and dual-wavelength algorithms, *J. Atmos. Ocean. Technol.*, 36, 883–902, <https://doi.org/10.1175/JTECH-D-18-0210.1>, 2019.
- Liu, B., Li, H., Liu, L., Shang, J., Yin, H., and Kuo, K.-S.: Sea surface and snowflakes as natural targets connecting FY-3G and GPM-CO dual-frequency radars, *Geophys. Res. Lett.*, 51, e2024GL110878, <https://doi.org/10.1029/2024GL110878>, 2024.
- Liu, B., Li, H., Liu, L., Shang, J., Kuo, K.-S., Lu, C., Yuan, M., and Jiang, B.: On the detection sensitivities of dual-frequency radars onboard FY-3G and GPM-CO, *Atmos. Res.*, 316, 107935, <https://doi.org/10.1016/j.atmosres.2025.107935>, 2025.
- Meneghini, R., Kim, H., Liao, L., Kwiatkowski, J., and Iguchi, T.: Path attenuation estimates for the GPM Dual-Frequency Precipitation Radar (DPR), *J. Meteorol. Soc. Jpn.*, 99, 181–200, <https://doi.org/10.2151/jmsj.2021-010>, 2021.
- Schwaller, M. R. and Morris, K. R.: A ground validation network for the Global Precipitation Measurement mission, *J. Atmos. Ocean. Technol.*, 28, 301–319, <https://doi.org/10.1175/2010JTECHA1403.1>, 2011.
- Sen, P. K.: Estimates of the regression coefficient based on Kendall's tau, *J. Am. Stat. Assoc.*, 63, 1379–1389, <https://doi.org/10.1080/01621459.1968.10480934>, 1968.
- Shang, J., Zhang, P., Cao, L., Wu, Q., Wang, X., Zhang, X., Jiang, B., Yin, H., Yuan, M., Liu, D., Chen, Y., and Gu, S.: Comparison of the Precipitation Measurement Radar onboard the FY-3G meteorological satellite with ground-based radars in China, *Adv. Atmos. Sci.*, 43, 645–660, <https://doi.org/10.1007/s00376-025-5149-3>, 2026.
- Skofronick-Jackson, G., Petersen, W. A., Berg, W., Kidd, C., Stocker, E. F., Kirschbaum, D. B., Kakar, R., Braun, S. A., Huffman, G. J., Iguchi, T., Kirstetter, P.-E., Kummerow, C., Meneghini, R., Oki, R., Olson, W. S., Takayabu, Y. N., Furukawa, K., and Wilheit, T.: The Global Precipitation Measurement (GPM) mission for science and society, *Bull. Amer. Meteor. Soc.*, 98, 1679–1695, <https://doi.org/10.1175/BAMS-D-15-00306.1>, 2017.
- Wang, J. and Wolff, D. B.: Comparisons of reflectivities from the TRMM Precipitation Radar and ground-based radars, *J. Atmos. Ocean. Technol.*, 26, 857–875, <https://doi.org/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 TRMM and GPM, *J. Atmos. Ocean. Technol.*, 35, 323–346, <https://doi.org/10.1175/JTECH-D-17-0128.1>, 2018.
- Wu, Q.: FY-3G PMR level-2 product instructions document [in Chinese], National Satellite Meteorological Center, China Meteorological Administration, Beijing, China, <https://satellite.nsmc.org.cn/FY3G/html/IOTDATASET.html> (last access: 27 August 2026), 2023.
- Wu, Q., Yang, M., Chen, L., Yin, H., Shang, J., and Gu, S.: A frequency correction algorithm for spaceborne precipitation measurement radar and ground-based weather radar, *Acta Meteorol. Sin.*, 81, 353–360, <https://doi.org/10.11676/qxxb2023.20220093>, 2023.
- Zhang, C., Hu, H., Zhu, J., Zhou, Q., Wu, L., Li, J., and Zhu, X.: Collaborative assessment of reflectivity consistency between FY-3G Precipitation Measurement Radar and ground-based radars, *Adv. Atmos. Sci.*, 43, 1065–1078, <https://doi.org/10.1007/s00376-025-5120-3>, 2026.



Zhang, P., Gu, S., Chen, L., Shang, J., Lin, M., Zhu, A., Yin, H., Wu, Q., Shou, Y., Sun, F., Xu, H., Yang, G., Wang, H., Li, L., Zhang, H., Chen, S., and Lu, N.: FY-3G satellite instruments and precipitation products: First report of China's Fengyun rainfall mission in orbit, *J. Remote Sens.*, 3, 0097, <https://doi.org/10.34133/remotesensing.0097>, 2023.

445 Zhong, L., Yang, R., Wen, Y., Chen, L., Gou, Y., Li, R., Zhou, Q., and Hong, Y.: Cross-evaluation of reflectivity from the space-borne precipitation radar and multi-type ground-based weather radar network in China, *Atmos. Res.*, 196, 200–210, <https://doi.org/10.1016/j.atmosres.2017.06.016>, 2017.