

Author's Response to Referee Comments

We thank the topic editor and both referees for their careful evaluation of our manuscript and dataset. All comments have been addressed point by point below. Each item is structured as follows:

- (1) Referee comment (*grey box*)
- (2) Author's response (*blue*)
- (3) Changes in manuscript (*green*)

Unless otherwise stated, all line and page numbers refer to the **marked-up (track-changes) version** of the revised manuscript appended to this document.

Referee #1 (RC1)

General comments

The paper describes a 10 - minute merged radar-gauge precipitation data set for South Korea with spatial resolution of 0.5 km and temporal coverage from 2016 to 2024. The manuscript is well written; the methods are comprehensible and the results regarding the data set quality are convincing. The data itself are very useful, well described and accessible. I have only some minor comments for improvement (see below).

Author's response:

We sincerely thank the referee for the positive and constructive assessment of our manuscript and dataset. We have carefully addressed all comments. In particular, we removed the unsupported statement from the abstract, revised Figure 2 to ensure consistent notation and a clearer data flow, simplified Figure 5, and clarified the rationale for the file organization and for dividing the dataset between two repositories. Detailed responses are provided below.

Comment 1.1

Line 24: "The dataset is evaluated through comparison with conventional geostatistical interpolation methods." I could not find this comparison in the text. This would be an interesting and necessary exercise if this would be a research paper. Since this is a data set paper this comparison is not absolute necessary but this sentence should be removed.

Author's response:

Thank you for pointing this out. We agree that the manuscript does not present a systematic comparison with conventional geostatistical interpolation methods, and we have therefore removed this statement.

Changes in manuscript:

The corresponding sentence has been removed from the abstract (Lines 24–25).

Comment 1.2

Figure 2 can be improved:

In “radar rainfall field” and “gauge rainfall field” the symbols “R” and “G” should be introduced.

Sometimes variables are functions of (t) sometimes not. Be consistent.

Sometimes variables are in brackets some times not. Be consistent.

Once a dot is used for multiplication, once not. Be consistent.

The arrow from the left lower box should go out at the end and input the right box on top.

Author's response:

Thank you for these helpful and detailed suggestions. We have carefully revised Figure 2 to improve the consistency and clarity of both the notation and the workflow, addressing each of the points raised.

Changes in manuscript:

- (1) The symbols $R(\mathbf{s}, t)$ and $G_i(t)$ are now explicitly introduced to denote the radar-derived rainfall field and the rainfall observed at the i -th gauge, respectively. The radar rainfall at the location of the i -th gauge is denoted $R(\mathbf{s}_i, t)$ (Lines 134–135, p. 5).
- (2) All time-dependent variables are now consistently written as functions of t (Figure 2, p. 7).
- (3) The use of parentheses in the mathematical expressions has been standardized throughout Figure 2.
- (4) Multiplication is now denoted consistently in all equations in Figure 2.
- (5) The arrow from the lower-left box has been repositioned so that it leaves from the end of the box and enters the “Residual computation” box from the top (Figure 2).

Comment 1.3

Line 194: There are more than 470.000 single files. Experience has shown that handling (e.g. copying) thousands of single files is much more time consuming than handling bigger files. Is this really the best way to store these files for each 10-minute interval separately?

Author's response:

We thank the referee for raising this important practical concern. The 10-minute file structure was adopted to remain consistent with the near-real-time production design of the RRBf framework, in which each precipitation field is generated and released as soon as it becomes available (approximately one minute of processing time per 10-minute analysis, as noted in Sect. 1). The Korea Meteorological Administration (KMA) follows a similar practice operationally, distributing its composite radar reflectivity product (HSR) as individual files at 5-minute intervals. Because one purpose of this dataset is to support future operational implementation, we intentionally preserved this per-interval file structure so that it directly reflects the output format of the underlying algorithm.

At the same time, we agree that distributing several hundred thousand individual files would be impractical for end users. For this reason, the dataset is not distributed as loose 10-minute files. Instead, the individual files are aggregated into monthly compressed archives (*.tar.gz) before being uploaded to Zenodo, which also limits the number of files per record. Users therefore download and extract the data by month rather than handling individual 10-minute files.

Changes in manuscript:

Sect. 4.3 has been revised to explain this design rationale (Lines 228–233). The monthly archive structure is described in Sect. 6 (Data availability).

Comment 1.4

Figure 5: What is the additional information when using line-dots in addition to the bar plot. One of the two is sufficient.

Author's response:

We thank the referee for this observation. We agree that displaying both the bars and the line-dot markers is redundant, as both convey the same annual values. The bar plot alone is sufficient to represent the interannual variability described in the text.

Changes in manuscript:

Figure 5 has been revised to show the bar plot only (p. 14, Line 310).

Comment 1.5

Data availability: The data should not be in two separate publications. Why are they not stored in a single archive?

Author's response:

We thank the referee for raising this point. The dataset was divided into two Zenodo repositories because of practical constraints of the Zenodo platform, which limits each record (DOI) to at most 100 files and a total storage volume of 50 GB. The full dataset (2016–2024) comprises approximately 473,000 individual 10-minute precipitation fields at 0.5 km resolution, and its total volume and file count exceed the limits for a single Zenodo record. The dataset was therefore split into two parts (2016–2020 and 2021–2024) solely to comply with these platform constraints, not for scientific or organizational reasons.

Changes in manuscript:

This rationale has been added to Sect. 6 (Data availability; Lines 429–430).

Referee #2 (RC2)

General comment

This paper describes a tremendous amount of work. It is very well written – and therefore I have only few specific comments. Most of them can be regarded as suggestions for future work.

Author's response:

We thank the referee for the encouraging assessment of our work and for the valuable suggestions, which we have addressed point by point below.

Comment 2.1

When it comes to the interpretation of remote sensing data there is always a debate about the meaning of the term “resolution”. I understand it as “physical resolution”. You mention (line 72 ...) a “spatial resolution of 500 m × 500 m” – which I would understand as the gridding of the data.

By comparing radar estimates with high-density station fields (~1.5 km distance), we found that radar estimates are smoother than the station-based estimates, which suggest that the “physical” resolution of the radar estimates is less than indicated.

I think it would be sufficient to mention that the radar data are available on a 500 m × 500 m grid – which does not necessarily represent the physical resolution of the data.

Author's response:

We thank the referee for this clarification. We agree that the term “spatial resolution” as used in the manuscript refers to the native grid spacing of the HSR product rather than the physical resolution of the

underlying radar measurements, which degrades with distance from each radar owing to beam broadening.

Changes in manuscript:

A clarifying sentence has been added to Sect. 2.1 to make this distinction explicit (Lines 73–75).

Comment 2.2

We noticed quite some fundamental changes when comparing gauge data with estimates from single and dual polarized radar data. For the validation period in 2022 this is not important anymore (except Incheon), but in further work you could check there have been larger differences in the time before. It would be interesting to know, when this update happened within 2014(2016) - 2019.

Author's response:

We thank the referee for this helpful suggestion. We have added a new table (Table B1, Appendix B) that lists, for each radar site in the KMA network, the year operation started, the year of the dual-polarization upgrade, the antenna altitude, and the frequency band. As shown in the table, most radars were upgraded sequentially from single- to dual-polarization systems between 2014 and 2019; among the ten operational sites, Baengnyeongdo was the first to be upgraded (2014) and Gangneung the last (2019). The Incheon International Airport radar was upgraded separately to a dual-polarization system in August 2022.

Changes in manuscript:

- (1) New Table B1 has been added in Appendix B (pp. 24–25).
- (2) A reference to Table B1 has been added to Sect. 2.1, together with a brief description of the upgrade timeline and a note that potential impacts of these hardware and configuration changes on the long-term homogeneity of reflectivity cannot be fully excluded (Lines 89–93).

Comment 2.3

Figure 1: Most radar stations are close to the coast, but some are further inland. Since there also altitude-dependent errors it would be interesting to know the altitudes of the station (e.g., in a table, showing also the replacement time (comment 2), or by mentioning the altitude range.

Author's response:

We thank the referee for this suggestion. The antenna altitudes are now listed in the new Table B1 (see also our response to Comment 2.2). The radar network spans a wide range of elevations, from low-lying coastal sites (e.g., Gangneung, 99 m; Gosan, 103 m) to high-altitude inland mountain sites (e.g., Myeonbongsan, 1136 m; Gwangdeoksan, 1064 m), reflecting the mix of coastal and inland sites visible in Fig. 1. We agree that altitude-dependent beam propagation and blockage effects are an important consideration, and the table allows readers to relate the characteristics of individual stations to the spatial error patterns discussed in Sect. 5.2.3.

Changes in manuscript:

Antenna altitudes of all radar sites are listed in Table B1 (Appendix B), which is referenced in Sect. 2.1 (Lines 89–93).

Comment 2.4

Most gauge stations (AWS) just resolve 0.5 mm increments, but you apply (e.g., line 142) a “minimum rainfall threshold of 0.1 mm h⁻¹”. I am not sure if this is fully consistent (but I would also not know how to make it better, without losing the light rain events).

Author's response:

We thank the referee for this careful observation, which prompted us to clarify the basis for this threshold. The value of 0.1 mm h⁻¹ was applied uniformly to both the 10-minute and the 60-minute analyses for consistency. It corresponds exactly to the smallest non-zero rainfall rate that the ASOS gauges (0.1 mm resolution) can resolve when accumulated over a 60-minute window, and is therefore physically meaningful at that timescale.

For 10-minute accumulations, the smallest non-zero rainfall rate that any gauge in the network can report is higher owing to gauge resolution (0.6 mm h⁻¹ for ASOS and 3.0 mm h⁻¹ for AWS, once 10-minute accumulations are converted to hourly rates). Consequently, the 0.1 mm h⁻¹ threshold effectively acts as a simple non-zero criterion at the 10-minute timescale, while retaining a direct physical meaning at the 60-minute timescale used in Table 2.

Changes in manuscript:

Sect. 3.2 has been revised to explain this distinction explicitly (Lines 154–158).

Comment 2.5

In Line 196 and Table 1 you use “Local Standard Time (LST)”, but in Line 74 you have already introduced the “Korea Standard Time (KST)” – which seems to be the same. I would suggest to use just one name throughout.

Author's response:

We thank the referee for catching this inconsistency. We now use “Korea Standard Time (KST, UTC+9)” throughout the manuscript, consistent with the definition given in Sect. 2.1 (Line 76).

Changes in manuscript:

All instances of “Local Standard Time (LST, UTC+9)” have been replaced with “Korea Standard Time (KST, UTC+9)”, including in Table 1 (Line 207) and Sect. 4.3 (Line 211).

Comment 2.6

Line 214: “(summer 2022” – there is just a missing “)”.

Author's response:

Thank you. This has been corrected.

Changes in manuscript:

The section heading now reads “... method (summer 2022)” (Line 235).

Comment 2.7

Line 222: “60 min accumulated rainfall” are these conventional hours, or can this be any 60 min interval? In the former case, you could consider just using “hours”?

Author's response:

We thank the referee for pointing out this ambiguity. These are not conventional clock hours. Each 60-minute accumulation is the sum of six consecutive, non-overlapping 10-minute values from the underlying 10-minute time series, and the 60-minute windows are not necessarily aligned with the top of the clock hour (e.g., a window may consist of the 10-minute values from 00:30 to 01:20). We have

therefore retained the term “60 min accumulated rainfall” rather than adopting “hours”, which could incorrectly imply alignment with standard clock hours.

Changes in manuscript:

A clarifying note has been added at the first use of the term in Sect. 5.1.1 (Lines 243–244).

Comment 2.8

Figure 6: The distribution in the wet year 2020 shows still considerable structure. Could this be related to typhoons?

Author's response:

We thank the referee for this insightful comment. We agree that the pronounced spatial structure in the 2020 annual precipitation field may be partly related to typhoon-induced rainfall. The year 2020 was characterized by an exceptionally prolonged summer rainy season, followed by several typhoons that affected the Korean Peninsula in August and September. In particular, Typhoons Maysak and Haishen made landfall near Busan and Ulsan, respectively, and subsequently tracked through Gyeongsang and Gangwon Provinces, producing substantial rainfall over southern and eastern Korea, including the mountainous regions of Gangwon Province (KMA National Typhoon Center, 2021). These typhoon-related events likely contributed to some of the pronounced spatial features in the 2020 annual accumulation field.

However, the spatial distribution in 2020 cannot be attributed solely to typhoons. The annual accumulation integrates precipitation over the entire year, and the exceptionally wet conditions in 2020 were also strongly influenced by the prolonged summer rainy season and several other heavy-rainfall events.

Changes in manuscript:

Text has been added stating that the spatial structure in 2020 likely reflects a combination of seasonal precipitation variability, orographic and coastal effects, and typhoon-related rainfall (Lines 316–321).

Comment 2.9

Line 357: You mention the bias on Jeju Island. Isn't there the highest mountain in entire South Korea? I might be worth mentioning this.

Author's response:

We thank the referee for this helpful suggestion and agree that it is worth mentioning. The bias likely reflects beam blockage by Hallasan (1,947 m), the highest mountain in South Korea, located near the center of Jeju Island. The blockage requires the HSR algorithm to substitute higher-elevation scans, which can lead to underestimation by the Z–R estimates. The RRBF estimates are less affected because gauge observations are directly incorporated into the bias correction.

Changes in manuscript:

A brief note explaining this has been added (Lines 393–395, p. 18).

Comment 2.10

Table 2 shows results for different threshold values, but I could not find a discussion in the text. It would be nice to summarize these results.

Author's response:

We thank the referee for pointing out this omission. As the rainfall threshold increases, the negative bias of the Z–R estimates becomes progressively larger (NBE from -13.13% to -30.73%) and their correlation with gauge observations decreases (COR from 0.7703 to 0.6852), reflecting greater underestimation by radar-only estimates during higher-intensity rainfall. In contrast, the RRBF estimates maintain a bias close to zero across all thresholds (NBE within $\pm 3.3\%$) and show only a modest decrease in correlation (COR from 0.9251 to 0.8699). This demonstrates that the benefit of the gauge-based bias correction persists even at higher rainfall intensities.

Changes in manuscript:

A summary of these threshold-dependent results has been added to the text (starting at Line 353).

Comment 2.11

Line 388: I assume that the partition into two data sets is due to the high data volume?

Author's response:

Yes, the referee's assumption is correct. The dataset was split into two Zenodo repositories because of platform-imposed limits on the number of files and the total storage volume per record (DOI); see also our response to Comment 1.5.

Changes in manuscript:

This rationale has been added to Sect. 6 (Data availability; Lines 429–430).

Comment 2.12

Do you plan to update the dataset in the future? If yes then it would be good to indicate this.

Author's response:

We thank the referee for raising this point. At present, we have no specific plans to extend the dataset beyond the 2016–2024 period covered in this study.

Changes in manuscript:

No changes were made.

Closing remark

All the best for your future work!

Author's response:

We sincerely thank the referee for the thoughtful and constructive comments, as well as for the kind wishes. We greatly appreciate the time and care taken in reviewing our manuscript.