

## **Responses to Reviewer 1 Comments**

Reviewer's comments are in black. Responses are in blue.

The authors are thankful to the Reviewer's constructive input and an appropriate acknowledgement for this has been added in the revised version.

-----

The manuscript presents a two-step machine learning framework for estimating precipitation from X-band radar observations in Cyprus. In the first step, a neural network is used to correct ground-based reflectivities using DPR data, and in the second step, the corrected reflectivities are used to estimate precipitation rates. While the general problem of inferring precipitation rates from ground-based X-band radar measurements is relevant for any region where they are deployed, the current study exhibits fundamental shortcomings in both methodology and results. Most importantly, the reported performance metrics raise serious concerns about the validity of the proposed approach and its implementation. The coefficients of determination ( $r^2$ ) reported for one of the PFO radar from 0.078 to -67.66. Negative values of  $r^2$  indicate that the model performs substantially worse than a trivial baseline indicating either implementation or data errors. This strongly suggests that the model, in its current form, does not capture a meaningful relationship between radar observations and precipitation. In addition, the precipitation fields shown in Figures 6 and 10 appear to be dominated by artifacts rather than coherent precipitation structures. Taken together, these results do not provide evidence that the proposed method produces reliable precipitation estimates. Given these issues, I cannot recommend publication of the manuscript in its current form.

### **General Response:**

We thank the reviewer for the thorough evaluation of our manuscript and for the constructive comments provided. We acknowledge the concerns raised regarding the validation of the proposed methodology and the interpretation of the results. In response, we carefully revised the manuscript and implemented a number of modifications aimed at improving the clarity, robustness, and completeness of the study.

In particular, we expanded the validation of the proposed framework by introducing additional diagnostic analyses, including joint-distribution heatmaps between radar and GPM Ku-band reflectivities, as well as boxplot analyses examining the relationship between corrected reflectivity and gauge rainfall rates. We also strengthened the discussion of the assumptions and limitations associated with the use of GPM DPR observations and clarified several methodological aspects throughout the manuscript.

We believe that these revisions strengthen the validation of the proposed framework and improve the presentation and interpretation of the results. Detailed responses to each comment are provided below.

### **General Comment 1**

#### **Comment:**

The manuscript does not provide convincing evidence that the proposed method yields realistic or skillful precipitation estimates. Before considering publication, the authors need to demonstrate that their method produces meaningful results. This could include comparison against a simple, working baseline or validation against independent estimates from spaceborne products (e.g., DPR L2 precipitation).

#### **Response:**

To strengthen the validation of the proposed methodology, several additional analyses and figures have been included in the revised manuscript. In particular, we expanded the evaluation of the reflectivity correction stage by introducing additional diagnostic analyses and visualisations. The revised manuscript now includes comparisons between raw and corrected radar reflectivities and GPM Ku-band observations for both radars, demonstrating a substantial increase in agreement after correction. Furthermore, boxplot analyses were added to investigate the relationship between corrected reflectivity and gauge rainfall rates. The boxplots show that higher corrected reflectivity intervals are generally associated with higher median rainfall rates for both radars, indicating that the corrected reflectivity fields preserve a physically meaningful relationship with rainfall intensity. These additions provide further evidence that the proposed framework captures meaningful relationships between radar observations and precipitation.

The following text was added to Section 5:

*"Figures 12 and 13 illustrate the performance of the reflectivity correction stage for the LCA and PFO radars, respectively. Panel (a) presents the relationship between the original radar reflectivities and the corresponding GPM Ku-band observations, while panels (b) and (c) show the corresponding relationships after the application of the neural-network-based correction for the training and validation datasets. For both radars, the corrected reflectivities exhibit substantially stronger agreement with the GPM observations than the original measurements."*

*"Figures 14 and 15 present boxplots of gauge rainfall rates grouped by corrected reflectivity intervals for the LCA and PFO radars, respectively. Higher corrected reflectivity intervals are generally associated with higher median rainfall rates, indicating that the corrected reflectivity fields preserve a physically meaningful relationship with rainfall intensity."*

Regarding the suggestion to validate the rainfall retrievals using DPR Level-2 rainfall products, we note that GPM DPR observations are already used during the reflectivity correction stage of the proposed framework. Consequently, DPR-derived rainfall products would not constitute a fully independent validation dataset. For this reason, rain-gauge measurements were retained as the primary independent reference for evaluating the rainfall estimation stage.

## **General Comment 2**

### **Comment:**

Please include plots of the training data statistics such as joint distributions of the principal input variables and corresponding target variables. Please also explain how the model is expected to generalize to precipitation rates outside the range of recorded precipitation rates.

### **Response:**

We thank the reviewer for this suggestion. To better document the training data characteristics, we have added a new figure per radar (Fig. 12 & 13) presenting joint distributions between the principal input variable (raw radar reflectivity) and the corresponding target variable (GPM Ku-band reflectivity) used during the reflectivity correction stage. In each figure, panel (a) shows the relationship between raw radar reflectivity and GPM Ku-band reflectivity, while panels (b) and (c) show the corresponding relationships between the corrected radar reflectivities and GPM Ku-band reflectivities for the training and validation datasets, respectively. The new figures demonstrate the

substantial increase in agreement achieved after the reflectivity correction stage for both radars.

The following text was added to Section 5:

*“Figures 12 and 13 illustrate the performance of the reflectivity correction stage for the LCA and PFO radars, respectively. Panel (a) presents the relationship between the original radar reflectivities and the corresponding GPM Ku-band observations, while panels (b) and (c) show the corresponding relationships after the application of the neural-network-based correction for the training and validation datasets. For both radars, the corrected reflectivities exhibit substantially stronger agreement with the GPM observations than the original measurements.”*

Regarding the model behaviour outside the range of observed precipitation rates, the following clarification was added:

*“As with most supervised machine-learning approaches, the models are expected to perform most reliably within the range of conditions represented in the training dataset. Predictions outside the observed training domain should therefore be interpreted with caution, as extrapolation performance cannot be guaranteed.”*

### **General Comment 3**

#### **Comment:**

The approach of using DPR reflectivities to correct the X-band reflectivities is questionable. In addition to operating at different frequencies and being 400 km far away, the DPR has a blind zone of up to 2 km over land and reflectivities below that are extrapolated. Furthermore, the resolution of the DPR is extremely low compared to the ground based radar, which may lead to the loss of information when using the DPR-based reflectivities to correct the ground-based ones. If you decide to continue to use the two-step approach based on the DPR reflectivities, these limitations need to at least be discussed.

#### **Response:**

Additional discussion has been added to the revised manuscript acknowledging the limitations associated with the use of GPM DPR observations for reflectivity correction. In particular, we now discuss differences in operating frequency, spatial resolution, sampling geometry, and the potential implications of the DPR blind zone over land. We also clarify that the objective of the first-stage model is not to directly replace ground-based radar observations with DPR measurements, but rather to exploit the DPR observations as an independent reference dataset for reducing systematic discrepancies in the ground-based radar reflectivity fields. The limitations and potential sources of uncertainty associated with this approach are now explicitly acknowledged in the manuscript. These considerations are now discussed in the revised manuscript to provide a balanced assessment of the strengths and limitations of the proposed framework.

The following text was added to the Discussion section:

*“The use of GPM DPR observations for reflectivity correction introduces several limitations that should be considered when interpreting the results. The DPR operates at a different frequency than the ground-based X-band radars and provides observations at a substantially coarser spatial resolution. In addition, the DPR measurements are affected by a near-surface blind zone over land, and reflectivities within this region are estimated through extrapolation procedures. These factors may introduce uncertainties into the reference dataset used during model development. Nevertheless, the DPR provides an independent and spatially consistent source of observations that can be exploited to reduce systematic*

*discrepancies in ground-based radar reflectivity measurements. The proposed approach should therefore be viewed as a data-driven bias-correction framework rather than a direct replacement of ground-based radar observations with satellite measurements.”*

### **Specific Comment 1**

**Comment:**

Sect. 3.2: Please provide the absolute number of samples.

**Response:**

We thank the reviewer for this suggestion. The absolute number of samples used in the analysis has now been explicitly reported in Section 3.2.

The following text was added:

*“After quality-control procedures and temporal matching, the final datasets comprised 25,741 samples for the LCA radar and 21,915 samples for the PFO radar.”*

### **Specific Comment 2**

**Comment:**

Sections 3.2.1 and 3.2.2 combine into one.

**Response:**

Sections 3.2.1 and 3.2.2 have been merged in the revised manuscript to improve readability.

### **Specific Comment 3**

**Comment:**

Section 3.2: A robust method should not require the use of different configurations for the two different radars. Please motivate or try to use the same approach for both radars.

**Response:**

Additional justification has been added to Section 3.2 regarding the use of radar-specific model configurations.

The following text was added:

*“The two radars are located at substantially different elevations above mean sea level (100 m for LCA and 392 m for PFO), resulting in differences in radar beam geometry and sampling volumes. Consequently, the relationship between radar-observed reflectivity aloft and rainfall measured at the surface is not expected to be identical for the two sites. Precipitation particles sampled by the higher-elevation PFO radar are, on average, observed farther from the surface and may undergo additional microphysical evolution before reaching the ground. Furthermore, differences in beam height, viewing geometry, and local topographic influences can affect the representativeness of radar observations with respect to surface rainfall measurements. For this reason, radar-specific model configurations were explored and selected based on validation performance for each radar individually.”*

### **Specific Comment 4**

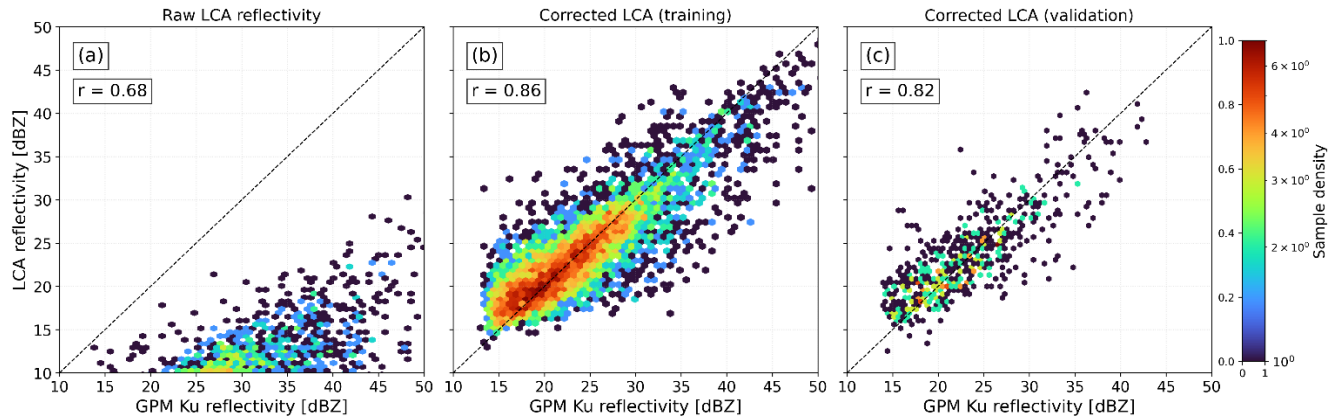
**Comment:**

Figure 3 + 7: Please use heat maps for all scatter plots as they are more informative.

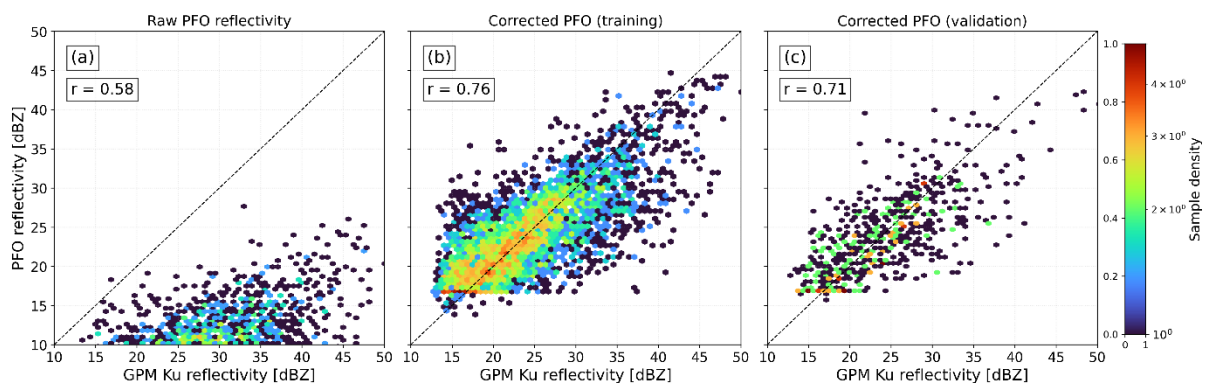
### Response:

The original scatter plots have been replaced by density-based heatmap (hexbin) representations. In addition, the newly introduced joint-distribution figures (Figs. 12 & 13) is presented using heatmaps for both radars.

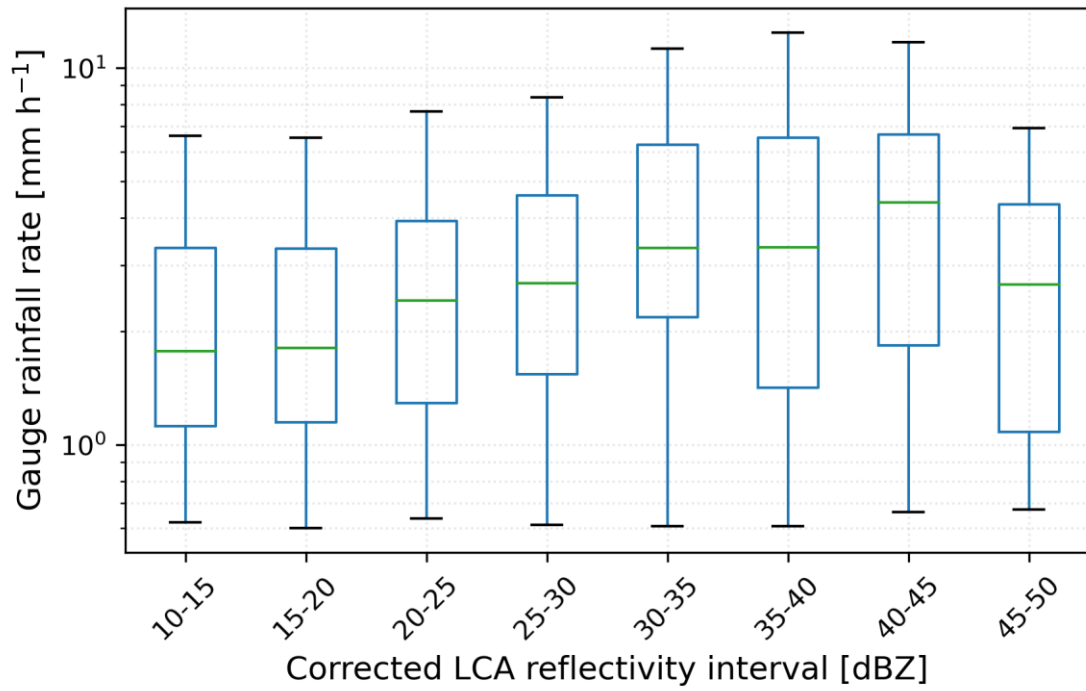
### Newly added figures:



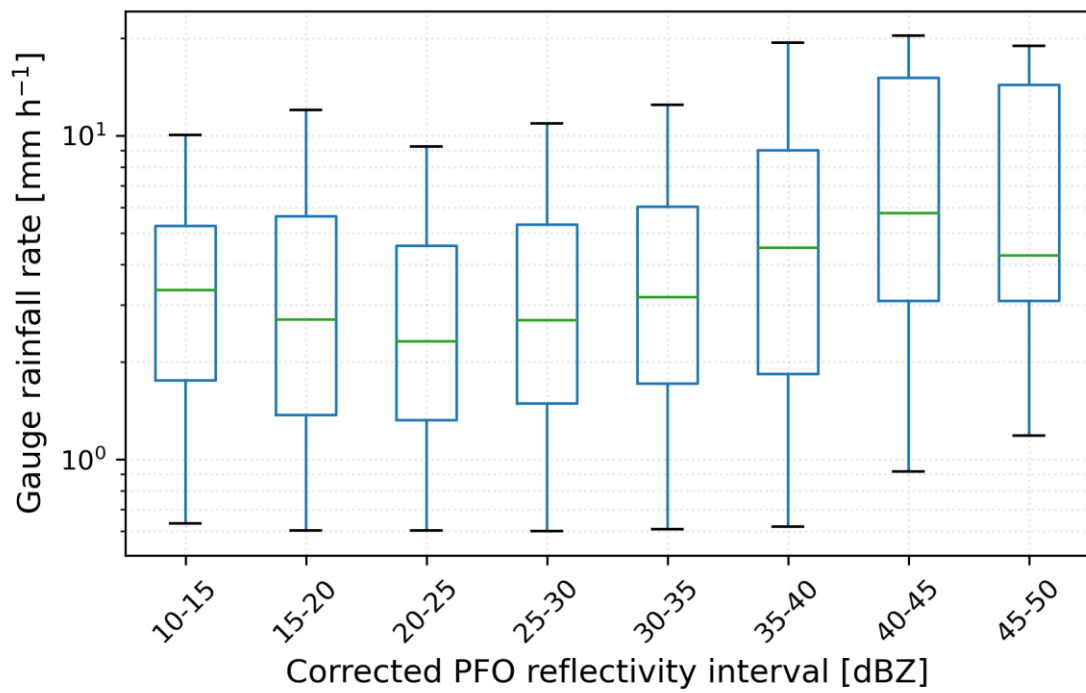
**Figure 12.** Joint distributions between GPM Ku-band reflectivity and LCA radar reflectivity: (a) raw reflectivity, (b) corrected reflectivity for the training dataset, and (c) corrected reflectivity for the validation dataset. The dashed line indicates the 1:1 relationship.



**Figure 13.** Joint distributions between GPM Ku-band reflectivity and PFO radar reflectivity: (a) raw reflectivity, (b) corrected reflectivity for the training dataset, and (c) corrected reflectivity for the validation dataset. The dashed line indicates the 1:1 relationship.



**Figure 14.** Boxplots of gauge rainfall rates grouped by corrected LCA reflectivity intervals.



**Figure 15.** Boxplots of gauge rainfall rates grouped by corrected PFO reflectivity intervals.