Response Sheet for Referee 01 Comments (egusphere-2025-6551)

We thank Referee #1 for their careful review and constructive suggestions, which have helped improve the manuscript. In the following sections, we respond to each comment and describe the corresponding revisions made to the manuscript.

The reviewer's comments are presented in bold red text, our responses are provided in black text, and the revised manuscript text is shown in blue. All line numbers and references correspond to the originally submitted manuscript.

Minor Comment 01

The introduction would benefit from incorporating several important recent studies that have applied innovative deep learning approaches and explicitly accounted for spatial autocorrelation in precipitation estimation frameworks. Including these studies would better position the current work within the existing literature and highlight methodological differences and contributions. Some relevant examples include (there are also several papers missing related to GNNs methods and their recent applications):

https://doi.org/10.3390/rs15174160

https://doi.org/10.1016/j.rse.2023.113723

https://doi.org/10.1016/j.atmosres.2022.106159

Thank you for this valuable suggestion. We agree that the Introduction would benefit from a stronger discussion of recent deep learning and spatial-dependence-aware precipitation estimation studies. In response, we have revised the Introduction to include recent studies that applied advanced deep learning, spatiotemporal fusion, and spatially aware machine learning approaches for precipitation estimation and merging.

Specifically, we now discuss Fang et al. (2023), who developed an attention-based ConvLSTM framework for multi-source precipitation fusion; Gavahi et al. (2023), who proposed a deep learning-based precipitation data fusion network combining 3D-CNN and ConvLSTM components to capture spatiotemporal dependencies; and Zandi et al. (2022), who compared stacking machine learning models with a locally weighted linear model for high-resolution

monthly precipitation estimation in complex terrain. We have also expanded the discussion of GNN-related studies to better position the proposed GraphIDW framework within recent efforts to explicitly represent spatial autocorrelation and spatial dependencies in rainfall estimation.

These additions clarify the methodological context of the present study and highlight its contribution, namely the integration of graph-based spatial learning with residual IDW correction for satellite-gauge precipitation merging over a tropical region. The following revised section will be added to the Introduction between Lines 57 and 69. The reference list will also be updated accordingly.

Recent studies have further demonstrated the potential of advanced DL and spatially explicit modelling frameworks for precipitation estimation, fusion, and downscaling. Fang et al. (2023) developed an attention-based ConvLSTM framework for multi-source precipitation spatiotemporal fusion, demonstrating the ability of attention mechanisms and recurrent convolutional structures to capture rainfall variability across space and time. Gavahi et al. (2023) proposed a deep learning-based precipitation data fusion network that combines three-dimensional convolutional neural networks and ConvLSTM components to represent spatiotemporal dependencies in multi-source precipitation products. Zandi et al. (2022) compared stacking ML approaches with a locally weighted linear model for high-resolution monthly precipitation estimation over complex terrain, emphasizing the importance of spatial non-stationarity in precipitation modelling. Collectively, these studies indicate that precipitation estimation can benefit from methods that explicitly represent nonlinear relationships, spatial heterogeneity, and spatiotemporal dependence.

While conventional ML models have shown strong predictive capability, many still implicitly assume that training samples are independent and identically distributed, an assumption that contradicts the spatial and temporal autocorrelation intrinsic to precipitation fields (Li & Heap, 2014; Hengl et al., 2018). Consequently, such models often fail to adequately represent the spatial dependence structures that control regional rainfall variability. Most ML-based PDM frameworks therefore incorporate additional covariates alongside precipitation estimates, including meteorological variables, topographic attributes, and spatial location information. However, when these auxiliary predictors are unavailable, conventional ML approaches frequently exhibit limited transferability and poor performance at ungauged locations. As a result, although traditional ML

methods can provide localized improvements, their effectiveness for large-scale precipitation merging remains constrained, highlighting the need for hybrid or advanced spatially explicit modelling strategies that directly account for spatial autocorrelation (Shen et al., 2018; Wu et al., 2020). Notably, most previous studies on SPP merging have overlooked the spatial autocorrelation that inherently exists among gauge observations during the merging process (Lei et al., 2022).

Recent advances in hydrological deep learning have increasingly focused on methods capable of representing spatial connectivity and dependency structures. Graph Neural Networks (GNNs) provide a more direct framework for representing such spatial relationships by modelling stations, grid cells, catchments, or monitoring wells as nodes and their spatial or hydrological connections as edges.

Several recent studies have emphasized the value of GNNs for hydrological and environmental applications where spatial autocorrelation and connectivity are important. Xiang and Demir (2022) developed a GNN-based fully distributed rainfall-runoff model that explicitly represents river network topology and inter-cell connectivity, thereby improving the representation of spatially heterogeneous rainfall-runoff processes. Bai and Tahmasebi (2023) applied a GNN framework for groundwater level forecasting by representing monitoring wells as graph nodes and learning inter-well spatial dependencies through graph convolution and a self-adaptive adjacency matrix. More recently, Taghizadeh et al. (2025) developed a physics-informed GNN framework for flood forecasting that incorporates mass conservation constraints while capturing complex spatiotemporal flood dynamics.

Despite these advances, the use of GNN-based frameworks for satellite-gauge precipitation merging remains limited, particularly in tropical regions where rainfall fields are highly heterogeneous and auxiliary predictors may be limited or unavailable. Furthermore, although conventional ML approaches can improve satellite precipitation estimates, their ability to represent spatial dependency among gauge observations remains restricted when spatial relationships are not explicitly embedded into the model structure. This provides the motivation for developing a graph-based PDM framework that can directly represent spatial relationships among rainfall observations while remaining applicable under limited auxiliary-data conditions.

Minor Comment 02

Although the study primarily uses IMERG V6, the Data Availability section mentions the use of IMERG V7. This inconsistency should be clarified.

Thank you for pointing out this inconsistency. We confirm that all analyses in this study were conducted using **IMERG V6**. The mention of **IMERG V7** in the Data Availability section was an error. We have revised the Data Availability section to consistently refer to **IMERG V6** throughout the manuscript.

Section **2.2.2** will be updated from Lines **107–108** as follows:

Original text:

The latest release, IMERG Version 06, extends coverage back to June 2000 through retrospective processing, thereby facilitating the assessment of long-term precipitation patterns (Huffman et al., 2019).

Revised text:

IMERG Version 06B extends coverage back to June 2000 through retrospective processing, thereby facilitating the assessment of long-term precipitation patterns (Huffman et al., 2019).

Lines **637–638** will also be updated as follows:

IMERG V06B can be accessed from the NASA Global Precipitation Measurement (GPM) website: https://gpm.nasa.gov/data.

Minor Comment 03

It is recommended to include the Kling–Gupta Efficiency (KGE) metric in the evaluation. KGE has become a widely accepted performance metric in hydrological studies because it simultaneously accounts for correlation, bias, and variability, providing a more balanced assessment of model performance.

Thank you for this valuable suggestion. We agree that the Kling–Gupta Efficiency (KGE) is a widely accepted and informative metric for hydrological model evaluation, as it simultaneously accounts for correlation, bias, and variability and therefore provides a more balanced assessment of model performance.

Accordingly, KGE has been included in the revised evaluation framework. The KGE values were calculated for each machine-learning-based product/model and incorporated into the comparative performance assessment together with RMSE, MAE, PCC, and relative bias. In particular, the box plots showing the variation of KGE across the evaluated ML products have been presented and discussed under the responses to **Major Comments 1 and 2**, where the overall model performance improvements and spatial/temporal variability are evaluated.

Minor Comment 04

A more comprehensive statistical analysis of the gauge observations is needed. For example, it would be helpful to present seasonal variability of precipitation, mean precipitation distribution across stations, and the elevation-precipitation correlation.

Thank you for this valuable suggestion. We agree that a more comprehensive statistical characterization of gauge observations is important for describing the rainfall regime of the study area and for supporting the interpretation of model performance. In response, we have revised the manuscript to include additional information on the seasonal and spatial variability of observed rainfall.

First, the seasonal precipitation variability of the Wet Zone has been incorporated into Section 2.1: Study area. The revised section now describes Sri Lanka's four major rainfall seasons, namely the First Inter-Monsoon, Southwest Monsoon, Second Inter-Monsoon, and Northeast Monsoon, and explains the seasonal rainfall distribution across the Wet Zone. The revision also notes that, during 2001–2015, the Southwest Monsoon contributes the largest proportion of annual rainfall in the Wet Zone, followed by the Second Inter-Monsoon, First Inter-Monsoon, and Northeast Monsoon.

Second, to represent the spatial distribution of mean precipitation across the study region, we have added a spatial map of observed annual average rainfall variability in Section 4.5: Spatial Evaluation. This map is generated using IDW interpolation of station-wise observed annual average rainfall values and provides a clearer visualization of regional rainfall gradients across the Wet Zone.

Regarding the elevation-precipitation relationship, we have clarified that rainfall does not increase monotonically with altitude across the Wet Zone. Instead, rainfall generally increases from the coastal lowlands toward the foothills and central highlands, reaches relatively high values around

approximately 1000 m elevation, and then decreases at higher elevations. This pattern is especially evident under the influence of southwest monsoon rainfall along the western slopes of the central highlands. Therefore, the revised manuscript describes the elevation–precipitation relationship as a non-linear topographic influence rather than a simple positive linear correlation. These additions provide better context for the spatial heterogeneity of rainfall across the study region and support the need for spatially explicit precipitation estimation methods.

The revised Section 2.1: Study area is shown below:

The Wet Zone of Sri Lanka (Fig. 1), covering an extent of approximately 12,665 km², is selected as the study region. Sri Lanka experiences a tropical monsoonal climate and is commonly classified into three principal climatic regions: the Wet Zone, Intermediate Zone, and Dry Zone, based on long-term mean annual rainfall patterns and associated biophysical characteristics (Punyawardena, 2020). The southwestern region of the country, including the central highlands, constitutes the Wet Zone, with elevations ranging from sea level to approximately 2,530 m above mean sea level. This zone receives high mean annual rainfall, varying between approximately 1,750 mm and 5,500 mm, and does not exhibit a clearly defined dry season.

Sri Lanka's rainfall regime is generally bimodal, although the seasonal distribution is not symmetrical. The four major rainfall seasons are the First Inter-Monsoon (FIM; March-April), Southwest Monsoon (SWM; May-September), Second Inter-Monsoon (SIM; October-November), and Northeast Monsoon (NEM; December-February). A representative monthly rainfall cycle for the Wet Zone, synthesized from long-term monthly normals at representative Wet Zone stations, indicates comparatively low rainfall during January-February, followed by a strong rainfall build-up during March-April. The primary rainfall maximum occurs during the Southwest Monsoon from May to September, while a second pronounced rainfall peak occurs during the Second Inter-Monsoon in October-November. Based on a zonal comparison for 2001–2015, approximately 46% of the annual Wet Zone rainfall occurs during the SWM, followed by 25% during the SIM, 15% during the FIM, and 14% during the NEM (Bandara et al., 2022).

However, according to the studies, rainfall does not increase monotonically with elevation across the study area (Nilanthi, 2016). Along the western slopes of the Central Highlands, southwest monsoon rainfall generally increases from the coastal lowlands toward the foothills and continues to increase up to approximately 1000 m elevation. Above this level, rainfall tends to decrease with

further increases in altitude. A similar pattern is observed for annual average rainfall, which increases up to around 1000 m before declining at higher elevations. This suggests that the elevation–precipitation relationship in the Wet Zone is non-linear and is strongly influenced by orographic effects, monsoonal circulation, and local topographic exposure.

The Wet Zone is therefore characterized by pronounced spatial and temporal rainfall variability, driven by the combined influence of monsoonal circulation, convective rainfall processes, and complex topographic gradients. This heterogeneity poses significant challenges for hydrological analysis, water resources management, and satellite-based precipitation estimation. In particular, the combined influence of high rainfall intensity, seasonal variability, and complex terrain can lead to substantial discrepancies in satellite precipitation products. These characteristics make the Wet Zone a suitable testbed for evaluating spatially explicit machine learning models for precipitation estimation and merging.

Minor Comment 05

In the manuscript, it is stated in the table that only monthly CHIRPS data were used. However, the methodology section and Figure 3 indicate that daily CHIRPS data were also used for downscaling. This discrepancy should be clarified.

We appreciate the reviewer's careful observation regarding this inconsistency. The data table previously listed only monthly CHIRPS, which did not fully describe the datasets used in the initial satellite precipitation product evaluation. We have revised the table to include both daily and monthly CHIRPS data and clarified their specific roles in the methodology.

Daily CHIRPS data were used only during the initial satellite precipitation product evaluation to compare the performance of CHIRPS, CHIRP, and IMERG at the daily temporal scale. Monthly CHIRPS data were used for the monthly-scale evaluation and subsequently selected for Eq. (1), because CHIRPS showed better agreement with gauge observations at the monthly scale. In Eq. (1), monthly CHIRPS provides the high-resolution 0.05° monthly rainfall constraint, while IMERG daily data provide the intra-monthly temporal fractions. Therefore, daily CHIRPS data were not directly used in the downscaling formulation; they were used only for the preliminary daily-scale product evaluation.

To avoid implying that the developed 0.05° daily product is purely IMERG-derived, we have also revised the notation from $\boldsymbol{IMERG_{daily,0.05^\circ}}$ to **IM_CH** throughout the manuscript. This revised notation reflects the hybrid nature of the product, where IMERG contributes the daily temporal distribution and CHIRPS contributes the monthly rainfall magnitude at 0.05° resolution.

The revised Table 1 is shown below.

Product	Spatial Resolution	Temporal Resolution
IMERG	0.1° × 0.1°	Daily, Monthly
CHIRP	0.05° × 0.05°	Daily, Monthly
CHIRPS	0.05° × 0.05°	Daily, Monthly
SRTM DEM	30 m	-

The text in Lines 155–160 of Section 3.1 will be revised as follows:

To combine these advantages, monthly CHIRPS rainfall at 0.05° resolution is used as the high-resolution monthly rainfall constraint, while daily IMERG rainfall is used to provide the intra-monthly temporal distribution. Monthly CHIRPS is selected for this step based on the preliminary SPP evaluation, which shows that CHIRPS produces better agreement with gauge observations at the monthly scale than the other evaluated products. In addition, CHIRPS provides rainfall estimates at a finer 0.05° spatial resolution and incorporates gauge-adjusted information, making it more suitable for representing monthly rainfall magnitude across the study region. In contrast, IMERG shows better performance at the daily scale and is therefore used to define the daily rainfall fraction within each month. Specifically, for each 0.1° IMERG grid cell, the ratio between daily IMERG rainfall and the corresponding monthly IMERG rainfall is calculated to represent the fractional contribution of each day to the monthly total. This IMERG-derived daily fraction is then applied to the corresponding CHIRPS monthly rainfall total at 0.05° resolution, as shown in Eq. (1). Therefore, the generated IM_CH product preserves the daily temporal variability of IMERG while constraining the monthly rainfall magnitude to CHIRPS.

$$IM\_CH_{daily,0.05^\circ} = CHIRPS_{monthly,0.05^\circ} \times \frac{IMERG_{daily,0.1^\circ}}{IMERG_{monthly,0.1^\circ}} \quad (1)$$

where $IM_CH_{daily,0.05^\circ}$represents the generated hybrid daily rainfall estimate at 0.05° resolution, $CHIRPS_{monthly,0.05^\circ}$is the CHIRPS monthly rainfall total at 0.05° resolution, $IMERG_{daily,0.1^\circ}$is the daily IMERG rainfall at 0.1° resolution, and $IMERG_{monthly,0.1^\circ}$is the corresponding monthly IMERG rainfall total at 0.1° resolution. Since the daily estimates are constrained to sum to the CHIRPS monthly total, IM_CH should be interpreted as a CHIRPS-constrained, IMERG-temporally-disaggregated rainfall product rather than a purely IMERG-derived 0.05° product.

Minor Comment 06

Additionally, the manuscript should include a comparison of IMERG before and after downscaling. Downscaling should ideally lead to at least some improvement in accuracy; otherwise, simple interpolation techniques such as bilinear interpolation or nearest neighbor resampling might produce similar results.

Thank you for this valuable comment. We agree that the manuscript should explicitly evaluate whether the proposed downscaling procedure improves IMERG relative to the original product. In response, we have added a comparison between the generated **IM_CH** product, the original IMERG product, and CHIRPS using gauge observations as the reference.

The added results show that **IM_CH** provides improved performance compared with the original IMERG product. In particular, **IM_CH** shows the highest median Pearson correlation coefficient among the evaluated products, indicating improved agreement with the temporal variability observed at the gauge stations. The RMSE distribution also shows that **IM_CH** has a slightly lower median RMSE than IMERG and CHIRPS, suggesting that the proposed CHIRPS-constrained, IMERG-temporally-disaggregated formulation provides added value beyond the original IMERG product. These results have been added to the revised manuscript to clarify the accuracy improvement achieved through the downscaling procedure.

We also acknowledge the reviewer's point that simple spatial resampling methods, such as bilinear interpolation or nearest-neighbor resampling, may produce similar results if the improvement is due only to increasing spatial resolution. Therefore, we have clarified that the proposed **IM_CH**

formulation is not a simple spatial resampling of IMERG. Instead, it combines the daily temporal variability of IMERG with the monthly rainfall magnitude of CHIRPS at 0.05° resolution. This distinction has been added to the methodology.

Following part will be added to the Section 3.1.

The performance of the developed IM_CH product was compared with the original IMERG and CHIRPS products using gauge observations as the reference. As shown in Fig. 4, IM_CH produces the highest median PCC among the evaluated products, indicating improved representation of daily rainfall variability compared with the original IMERG and CHIRPS products. The RMSE distribution also indicates that IM_CH has the lowest median RMSE, although the improvement relative to IMERG is modest. These results suggest that the proposed downscaling formulation improves the agreement between satellite-based rainfall estimates and gauge observations by combining the stronger daily-scale temporal representation of IMERG with the more reliable monthly rainfall magnitude of CHIRPS.

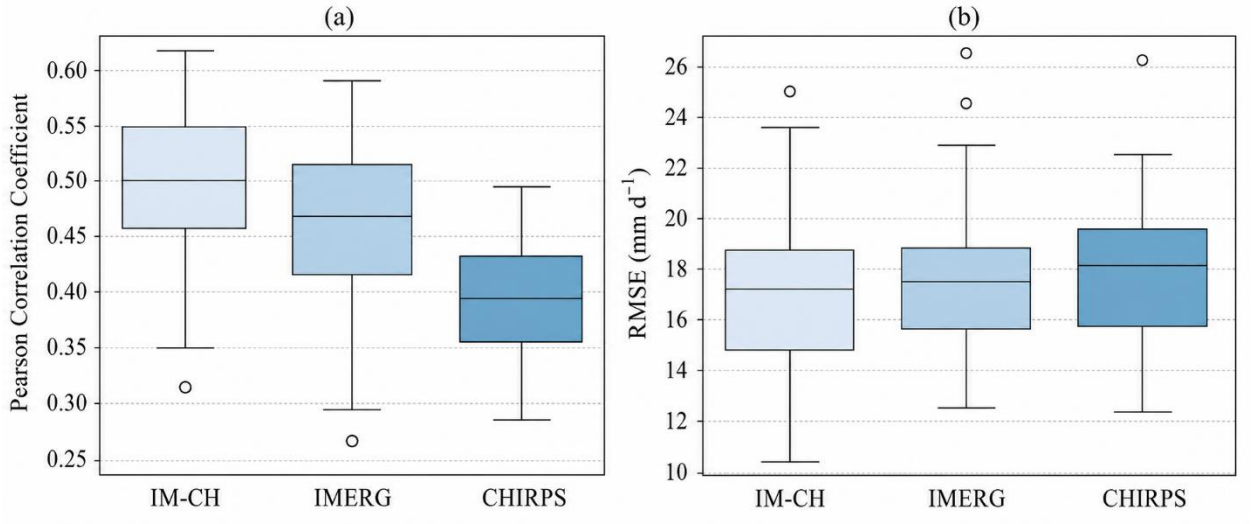

Figure 4. Box plots of spatial variability of (a) Pearson correlation coefficient and (b) root mean square error between IM-CH, IMERG and CHIRPS and observed rainfall

Minor Comment 07

Latitude and longitude were used as input features in the model. However, these variables are static spatial attributes, while satellite observations are dynamic temporal features. It can be concluded that the location is already encoded in the graph structures through adjacency and edge weight matrices.

We appreciate the reviewer's careful observation regarding the use of latitude and longitude as input features. We agree that latitude and longitude are static spatial attributes, whereas satellite precipitation estimates are temporally dynamic variables. We also agree that part of the spatial information is already represented through the graph structure, particularly through the adjacency and edge-weight matrices.

In the proposed GraphIDW framework, latitude and longitude are included to provide each node with its absolute geographic position, while the graph structure represents the relative spatial connectivity among neighbouring nodes. Therefore, these two components encode different aspects of spatial information. The coordinate features describe where each node is located in geographic space, whereas the adjacency and edge-weight matrices describe how nodes are connected and how strongly neighbouring nodes influence each other during message passing.

However, we acknowledge that some degree of redundancy may exist because spatial proximity is also partially encoded through the graph structure.

Following text will be added from **Line 208, Section 3.2.2 GNN-IDW model (GraphIDW) structure,**

Latitude and longitude are retained as auxiliary node attributes to provide an absolute geographic reference, whereas the adjacency matrix and edge weights encode relative neighbourhood connectivity. These signals are partly redundant, and their incremental contribution is not isolated through ablation in the present study. However, prior geospatial GNN work suggests that coordinate-based positional information can complement graph-based local neighbourhood structure in modelling spatially heterogeneous processes (Klemmer et al., 2023).

Minor Comment 08

Since the grid structure appears to be regular, the edge weights between nodes are likely identical. In such cases, a binary adjacency matrix may be sufficient. The manuscript should clarify whether weighted edges provide additional benefits in this context.

We appreciate this helpful comment. We agree that when graph nodes are arranged on a perfectly regular grid, the distances between neighbouring nodes may be identical or nearly identical, and in such cases a binary adjacency matrix can be sufficient. However, in this study, weighted edges are used to represent distance-dependent spatial interaction among neighbouring nodes. This allows closer nodes to exert greater influence during graph-based message passing, which is consistent with the spatial autocorrelation behaviour of precipitation fields.

Line 202 in Section 3.2.2 will be updated as follows:

Edge weights are assigned as the inverse of the inter-node distance to represent distance-dependent spatial interaction among neighbouring nodes. This weighting scheme allows closer nodes to exert stronger influence during graph-based message passing, consistent with the spatial autocorrelation of precipitation.

Minor Comment 09

Please clarify which software packages or libraries were used to implement the GraphIDW model and the other machine learning methods. Providing implementation details improves reproducibility.

Thank you for highlighting the importance of implementation details for reproducibility. We agree that including software and library details improves reproducibility. In response, we have revised the methodology section to specify the Python libraries used to implement the proposed GraphIDW model and the benchmark machine learning models.

The GraphIDW model was implemented using **PyTorch** and **PyTorch Geometric**, including *torch, torch.nn, torch.nn.functional, Data*, and *GCNConv*. These details will be added to **Section 3.2.2: GNN-IDW model (GraphIDW) structure**.

The benchmark models were implemented using **scikit-learn** and **XGBoost**. Specifically, Random Forest was implemented using *RandomForestRegressor*, Support Vector Regression was

implemented using *SVR* with *StandardScaler* and *Pipeline*, ANN was implemented using *MLPRegressor*, and XGBoost was implemented using *XGBRegressor*. These details will be added to the corresponding subsections for each machine learning algorithm under **Section 3.3: Traditional ML algorithms**.

Data preprocessing and statistical analyses were performed using ***pandas***, ***NumPy***, and ***SciPy***, while figures were generated using ***Matplotlib***.

Major Comments 01 and 02

The post-processing residual correction was applied only to the GraphIDW approach, while the other machine learning methods were evaluated without this correction. This introduces an inconsistency in the comparison. It is recommended to apply the correction method to all machine learning models, which would allow for more direct (apple-to-apple) comparison between GraphIDW and the traditional ML approaches.

Inverse Distance Weighting (IDW) is a simple yet effective spatial interpolation method and is commonly used as a benchmark method. However, its spatial patterns are often strongly influenced by the uniform weighting scheme and bull's-eye effect. The manuscript should clearly position IDW as a baseline and discuss its limitations relative to more advanced methods.

Thank you for these important comments. We agree that applying the residual correction only to GraphIDW introduced an inconsistency in comparison with the traditional ML models. In response, we have revised the analysis by applying the same post-processing residual correction procedure to all benchmark ML models. For each model, residuals are calculated at the training stations as the difference between observed rainfall and model-predicted rainfall. These residuals are then spatially interpolated using IDW and added back to the original model predictions to generate residual-corrected estimates. This allows a more direct and consistent comparison between GraphIDW and the traditional ML approaches.

We have revised the relevant results figures and performance tables to include both the residual-corrected ML models and the proposed GraphIDW framework. This revision enables an "apple-to-apple" comparison by evaluating all models under the same corrective post-processing strategy.

We also agree that IDW should be clearly positioned as a baseline method. In the revised manuscript, IDW is included as a benchmark interpolation method and is evaluated using the same validation framework and performance metrics as the ML-based approaches. We have also expanded the discussion of IDW limitations, including its sensitivity to gauge distribution, its reliance on distance-based weighting, its inability to incorporate nonlinear relationships or auxiliary predictors, and its tendency to generate localized bull's-eye patterns around gauge stations. These revisions clarify the role of IDW as a simple but important baseline against which the added value of the advanced methods can be interpreted.

The revised Figs. 5,6 and 7 are shown below.

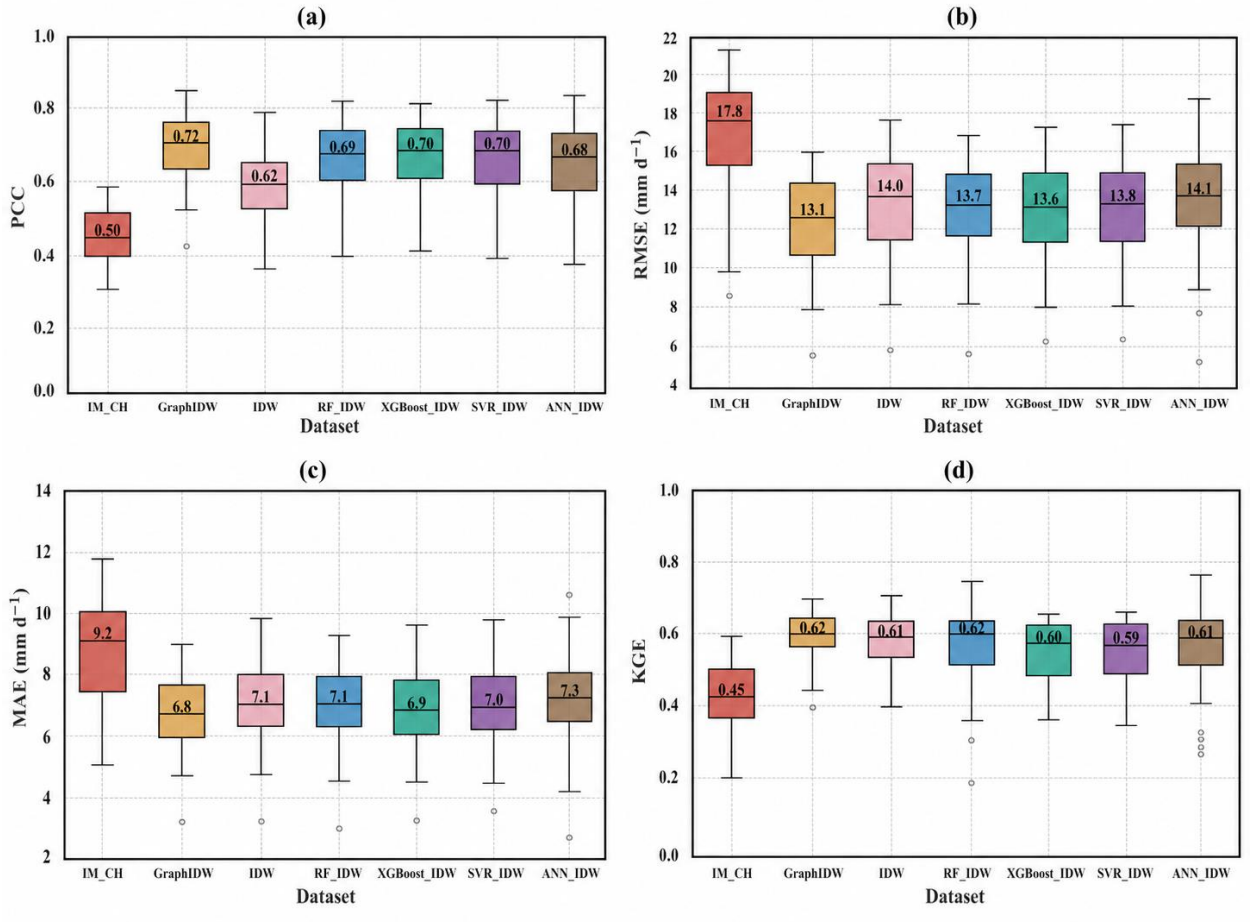

Figure 5. Boxplots of four quantitative metrics (PCC (a), RMSE (b), MAE (c), and (d) KGE) for seven products, including original (IM_CH) product

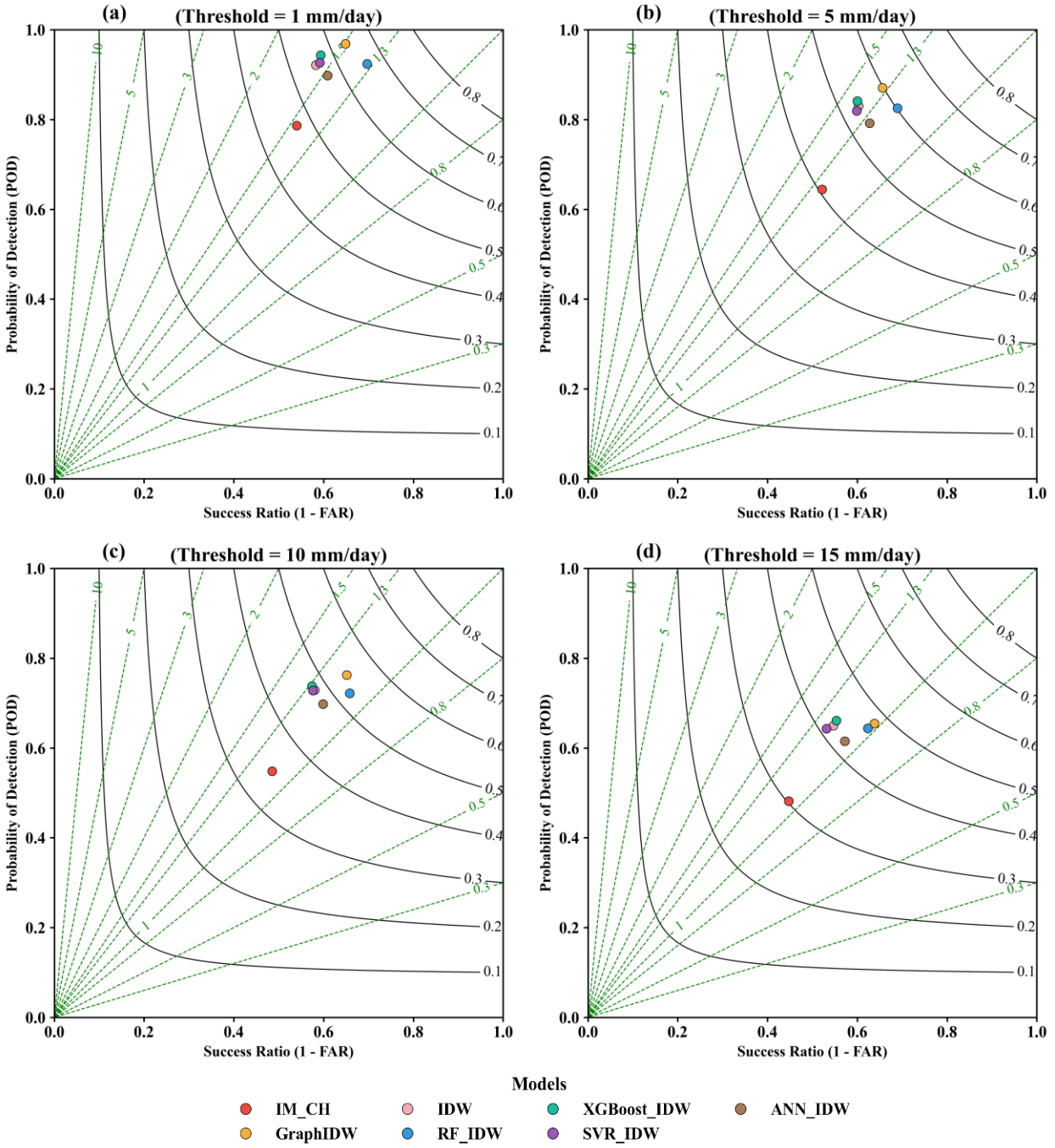

Figure 6. Roebber's performance diagram for the original (IM_CH) and the merged precipitation products. The green dashed lines represent the bias score (BS), while the black contour lines indicate the critical success index (CSI). The four diagrams present the detection performance for the precipitation thresholds of (a) 1 mm d^{-1}, (b) 5 mm d^{-1}, (c) 10 mm d^{-1}, and (d) 15 mm d^{-1}, respectively

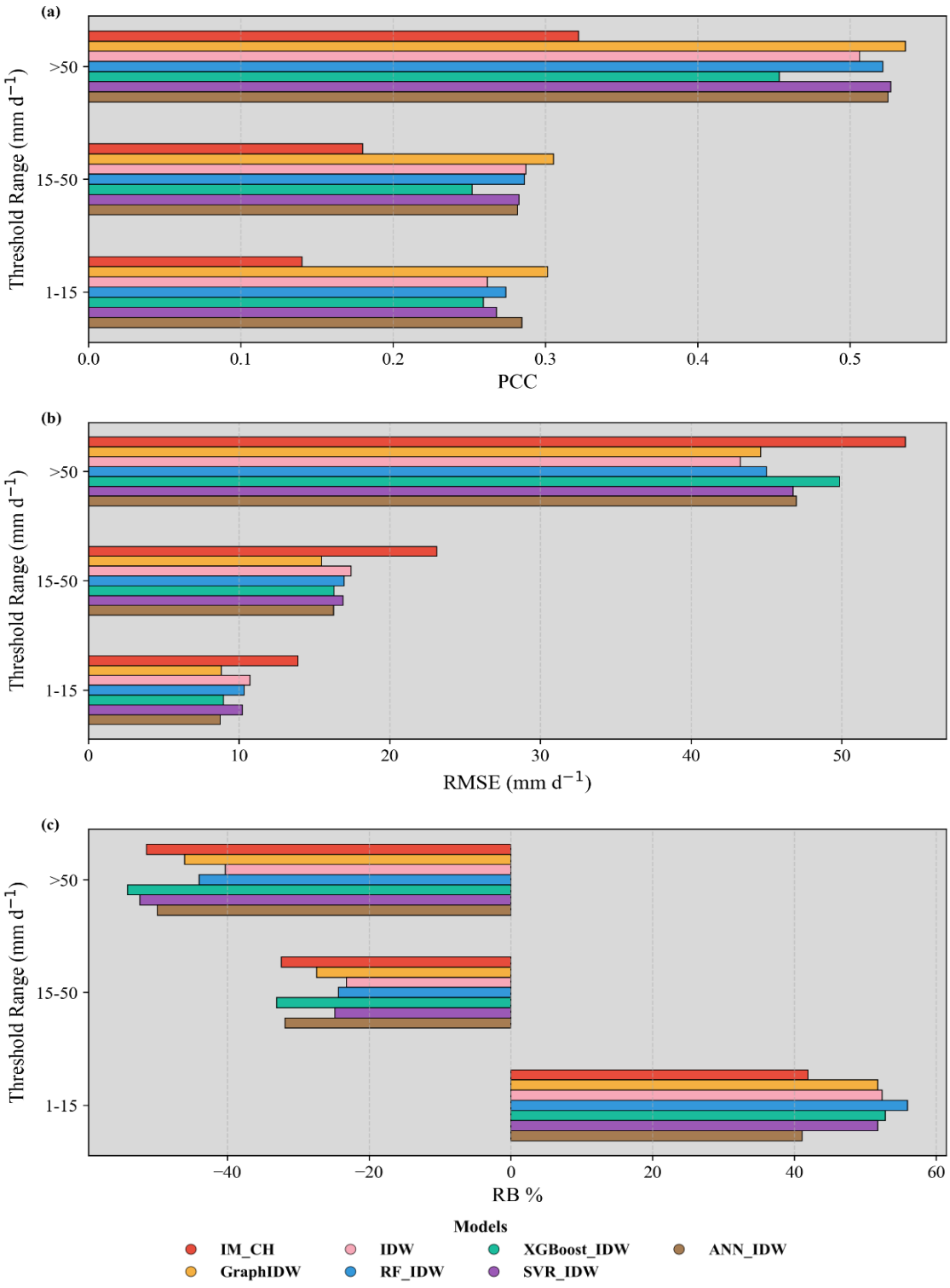

Figure 7. Comparison of (a) PCC, (b) RMSE, and (c) RB% of seven precipitation products against ground observations under different rain intensities. (Heavy rain: >50 mm d^{-1}; Moderate: 15-50 mm d^{-1}; Light rain: 1-15 mm d^{-1}).

Major Comments 03 and 04

The manuscript mentions the Single Mass Curve method, but it is not sufficiently explained. Please provide a brief description of the method and explain how it contributes to assessing the reliability of the products.

It is strongly recommended to include maps showing the spatial distribution of mean precipitation (or representative high-intensity events) across the study region. Such visual comparisons are important because realistic spatial precipitation patterns are a critical indicator of model performance.

Thank you for these valuable comments. We agree that the explanation of the **Single Mass Curve (SMC)** method in the original manuscript was not sufficiently clear, and that a clearer spatial visualization of rainfall patterns would improve the presentation.

Regarding the **Single Mass Curve method**, the original intention was to use the station-wise SMC-derived values to illustrate rainfall variability across the study region. However, upon further consideration, we agree that this is not the most suitable way to represent regional rainfall variation or to assess the realism of spatial precipitation patterns. Therefore, in the revised manuscript, the previous SMC-based figure (Figure 11) has been removed and replaced with a map showing the spatial distribution of observed annual average rainfall, generated by applying IDW interpolation to the station-wise observed annual average rainfall values. This revised figure provides a more direct and interpretable representation of rainfall variability and is better suited for visual comparison of precipitation patterns.

In response to the second comment, we agree that maps showing the spatial distribution of precipitation are important for assessing whether the rainfall products reproduce realistic spatial patterns. Accordingly, the revised manuscript now includes a spatial map of **annual average rainfall variability** over the study region. This figure more clearly illustrates the spatial rainfall regime and provides a stronger basis for interpreting the spatial realism of the products. Section **4.5 Spatial Evaluation** will be updated accordingly.

The revised Figs. 11 and 12 are shown here.

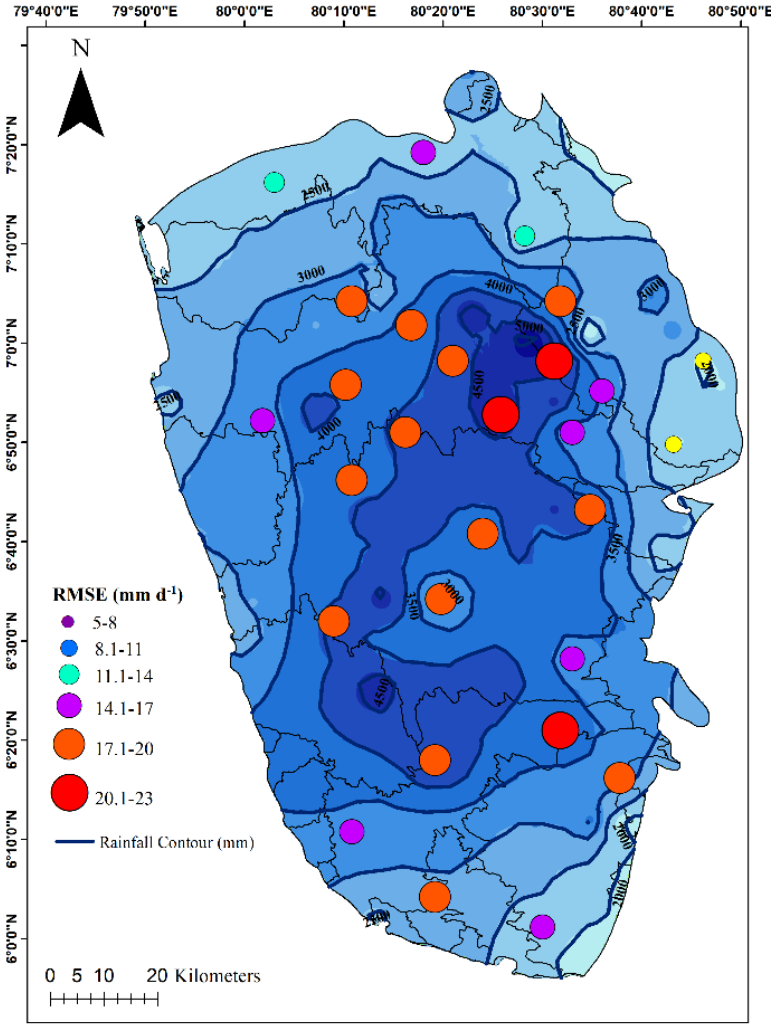

Figure 11. Spatial distribution of RMSE for the original IM_CH product overlaid on the IDW-interpolated observed annual average rainfall across the Wet Zone of Sri Lanka.

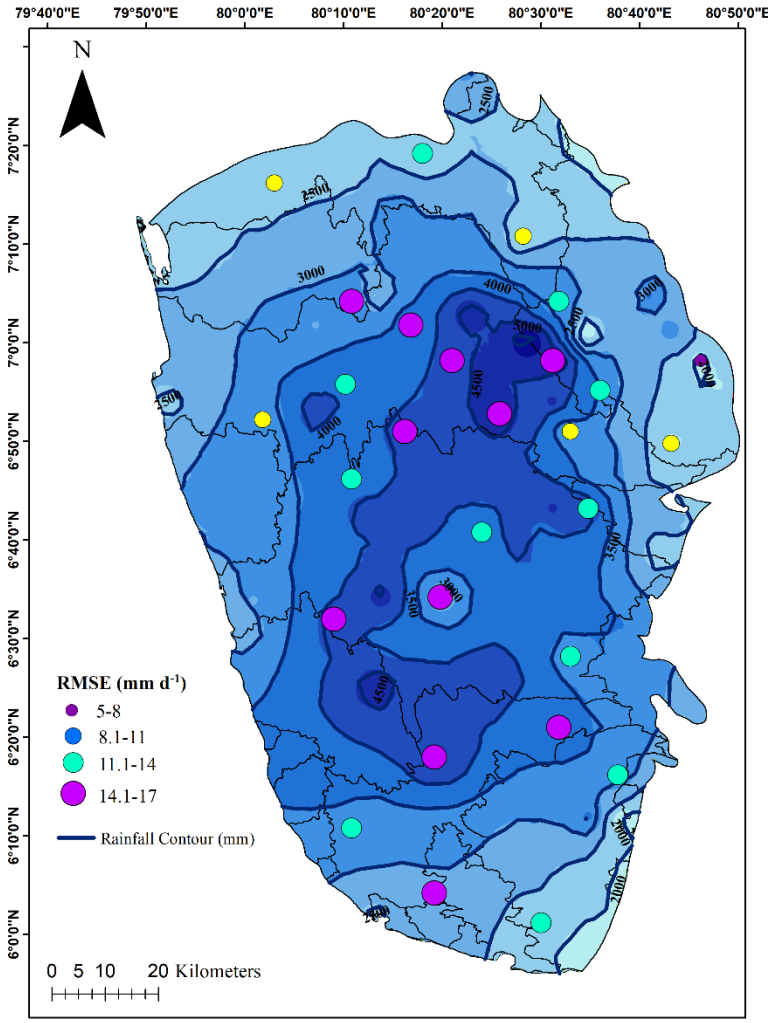

Figure 12. Spatial distribution of RMSE for GraphIDW-predicted rainfall overlaid on the IDW-interpolated observed annual average rainfall across the Wet Zone of Sri Lanka.

Major Comment 05

The manuscript states that the proposed approach follows the methodology of Baez-Villanueva et al. (2020) and Zhang et al. (2021). However, based on the description provided in Section 3.3, it appears that the implementation corresponds only to the method proposed by Baez-Villanueva et al. (2020). The approach introduced by Zhang et al. (2021) differs slightly. Therefore, the statement that the study follows both approaches may need clarification. Incorporating the method proposed by Zhang et al. (2021) could improve model accuracy. It would therefore be helpful if the authors could clarify this point and explain which method they exactly implemented.

Thank you for this important clarification. We agree that the original wording could be interpreted as implying that the complete methodological frameworks of both Baez-Villanueva et al. (2020) and Zhang et al. (2021) were directly implemented in this study. This was not our intention. Rather, the present study adopts the **time-step-specific, day-to-day training strategy** used in these studies for developing models.

Specifically, Baez-Villanueva et al. (2020) used a time-step-based modelling strategy, where a separate regression model is derived for each time step to estimate precipitation at the desired temporal resolution. Similarly, Zhang et al. (2021) reported that a day-to-day training strategy was more effective and computationally efficient than periodic training strategies. Following this concept, the benchmark machine learning models in the present study were trained using a daily time-step strategy, where model training and prediction were performed at the daily scale rather than using monthly or annual aggregated training periods.

To avoid ambiguity, we have revised Section 3.2.3 to clarify that the present study follows the **daily time-step training strategy** inspired by Baez-Villanueva et al. (2020) and Zhang et al. (2021), rather than reproducing all methodological components of both studies. This clarification has been added to ensure that the implemented approach is accurately described.

Line 236 of Section 3.2.3 will be updated as follows:

Following the daily time-step training strategy adopted by Baez-Villanueva et al. (2020) and Zhang et al. (2021), an individual GraphIDW model is constructed at each daily time step to generate grid-level precipitation estimates.

Major Comment 06

It is recommended to reconsider the inclusion of Figure 13. The comparison of computational speed may not be fully informative without providing details about the computational hardware. For example, methods such as ANN and GNNs can be significantly accelerated when implemented on GPUs, whereas Random Forest (RF) models typically depend heavily on CPU-based multithreading. It is also unclear whether multithreaded training was used for the RF model and how many CPU cores were available. If the authors intend to keep this analysis, it is strongly recommended to report the hardware configuration used for training,

including CPU specifications, number of cores, GPU usage (if any), and relevant software settings.

Thank you for this valuable comment. We agree that the computational speed comparison should be interpreted carefully because runtime depends on the hardware configuration, GPU availability, CPU multithreading, and software settings. In response, we have revised the manuscript to report the hardware used for model training and to clarify that the runtime comparison is specific to the computational environment used in this study.

All model training and evaluation were conducted on an HP Pavilion laptop equipped with an Intel Core i7-1065G7 CPU with 4 Cores and 8 GB RAM. No GPU acceleration was used during the analyses; therefore, the reported computational times reflect CPU-based execution under this specific hardware configuration. We have also revised the caption and discussion of Figure 13 to clarify that the results are intended only as an indicative comparison of runtime under the stated setup, rather than a general benchmark of computational efficiency across algorithms.

Section 5. Discussion, Line 571 will be updated as follows:

Figure 13 illustrates the computational training time of the five ML models using a training sample of 60 rain gauges under the hardware configuration used in this study. All models are trained and evaluated on an HP Pavilion laptop equipped with an Intel Core i7-1065G7 CPU with 4 Cores and 8 GB RAM, without GPU acceleration.

References

Bai, T. and Tahmasebi, P.: Graph neural network for groundwater level forecasting, Journal of Hydrology, 616, 128792, https://doi.org/10.1016/j.jhydrol.2022.128792, 2023.

Bandara, U., Agarwal, A., Srinivasan, G., Shanmugasundaram, J., and Jayawardena, I. M. S.: Intercomparison of gridded precipitation datasets for prospective hydrological applications in Sri Lanka, Intl Journal of Climatology, 42, 3378–3396, https://doi.org/10.1002/joc.7421, 2022.

Fang, W., Qin, H., Liu, G., Yang, X., Xu, Z., Jia, B., and Zhang, Q.: A Method for Spatiotemporally Merging Multi-Source Precipitation Based on Deep Learning, Remote Sensing, 15, https://doi.org/10.3390/rs15174160, 2023.

Gavahi, K., Foroumandi, E., and Moradkhani, H.: A deep learning-based framework for multi-source precipitation fusion, Remote Sensing of Environment, 295, 113723, https://doi.org/10.1016/j.rse.2023.113723, 2023.

Klemmer, K., Safir, N., and Neill, D. B.: Positional Encoder Graph Neural Networks for Geographic Data, Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (AISTATS 2023), PMLR, 206, 1379–1389, 2023.

Nilanthi, K. W. G. R.: Rainfall variation on the windward and leeward side of the Central Highland of Sri Lanka, International Journal of Scientific and Research Publications, 6, 2016.

Taghizadeh, M., Zandsalimi, Z., Nabian, M. A., Shafiee-Jood, M., and Alemazkoor, N.: Interpretable physics-informed graph neural networks for flood forecasting, Computer-Aided Civil and Infrastructure Engineering, 40, 2629–2649, https://doi.org/10.1111/mice.13484, 2025.

Xiang, Z. and Demir, I.: Fully distributed rainfall-runoff modeling using spatial-temporal graph neural network, https://doi.org/10.31223/X57P74, 15 January 2022.

Zandi, O., Zahraie, B., Nasseri, M., and Behrangi, A.: Stacking machine learning models versus a locally weighted linear model to generate high-resolution monthly precipitation over a topographically complex area, Atmospheric Research, 272, 106159, https://doi.org/10.1016/j.atmosres.2022.106159, 2022.