

# Deep Learning for Tropical Rainfall: Enhancing Quantitative Estimation and Extreme Event Detection Using Polarimetric Radar

N'guessan Ghislain Kouadio<sup>1</sup>, Augustin Kadjo Koffi<sup>2</sup>, Ibrahim Bamba<sup>2</sup>, Eric-Pascal Zahiri<sup>2</sup>, Modeste Huberson Kacou<sup>2</sup>, Adama Sahouarizie Ouattara<sup>2</sup>, N'guessan Apolline Yapi<sup>3</sup>

<sup>1</sup>Unité de Recherche et d'Expertise Numérique, Université Virtuelle de Côte d'Ivoire, Abidjan, Côte d'Ivoire

<sup>2</sup>Laboratoire des Sciences de la Matière, de l'Environnement et de l'Energie Solaire, Université Félix Houphouët-Boigny, Abidjan, Côte d'Ivoire

<sup>3</sup>Laboratoire de Physique Fondamentale et Appliquée, Université Nanguy Abergoua, Abidjan, Côte d'Ivoire

Email: ghislain.kouadio@uvci.edu.ci

**How to cite this paper:** Kouadio, N.G., Koffi, A.K., Bamba, I., Zahiri, E.-P., Kacou, M.H., Ouattara, A.S. and Yapi, N.A. (2026) Deep Learning for Tropical Rainfall: Enhancing Quantitative Estimation and Extreme Event Detection Using Polarimetric Radar. *Open Journal of Modern Hydrology*, **16**, 128-157.

<https://doi.org/10.4236/ojmh.2026.162009>

**Received:** November 20, 2025

**Accepted:** March 7, 2026

**Published:** March 10, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Accurate quantitative precipitation estimation (QPE) in tropical regions remains a major challenge due to the high spatio-temporal variability of rainfall and the limitations of traditional parametric radar-rainfall relationships. This study develops a Multi-Layer Perceptron (MLP) Deep Learning model to estimate rainfall rates from polarimetric radar observations collected during the AMMA (African Monsoon Multidisciplinary Analysis) campaign in northern Benin. The model uses three primary radar variables: horizontal reflectivity ( $Z_h$ ) differential reflectivity ( $Z_{dr}$ ), and specific differential phase ( $K_{dp}$ ) complemented by derived features to enhance data representation. Performance evaluation against conventional parametric algorithms, including optimized  $Z_h - R$  relationships, multiparametric regressions, and T-Matrix-based microphysical simulations, demonstrates the clear superiority of the MLP. The proposed model achieves the lowest RMSE ( $7.04 \text{ mm}\cdot\text{h}^{-1}$ ), NRMSE (0.53), and MAE ( $2.85 \text{ mm}\cdot\text{h}^{-1}$ ), with a near-zero normalized bias (0.88%) and the highest coefficient of determination ( $R^2 = 0.72$ ). Residual analysis confirms the absence of systematic bias, while the ROC curve ( $\text{AUC} = 0.96$ ) highlights excellent skill in detecting heavy rainfall events ( $\geq 20 \text{ mm}\cdot\text{h}^{-1}$ ). The Diebold-Mariano test further validates the statistical significance of the MLP's improvements over reference models. These results confirm that the integration of deep learning with polarimetric radar variables offers a robust and accurate approach for QPE in tropical environments, with strong potential for operational hydrometeorological applications and early warning systems.

---

## Keywords

Quantitative Precipitation Estimation, Polarimetric Radar, Deep Learning, Extreme Rainfall, West Africa

---

## 1. Introduction

Accurate rainfall estimation using weather radar represents a major challenge due to the vital role of rain in human society and life in general, namely in water resource management, agriculture, and the development of early warning systems for managing hydrological risks such as floods. The main difficulty lies in the high spatiotemporal variability of precipitation, which complicates its precise estimation using traditional methods derived from radar and in situ measurements.

Traditionally, radar rainfall estimates have been derived from a parametric relationship between horizontal reflectivity  $Z_h$  and precipitation rate  $R$  ( $Z_h - R$  relationships) since the work of [1]. However, the main difficulty in precipitation estimation remains the high variability of  $Z_h - R$  relationships. Indeed, numerous studies [2]-[9] on the development of these  $Z_h - R$  relationships have highlighted this difficulty, which limits their performance. In the case of West Africa (tropical zone), the variability of  $Z_h - R$  relationships has been widely documented [9]-[13]. Such an approach has met with limited success due to the high observed variability of  $Z_h - R$  relationships. Furthermore, polarimetric radars have established themselves as an ambitious alternative thanks to their ability to introduce, on the one hand, additional observables, and on the other, to extract multi-dimensional information on the physical properties of targets [14]-[16]. Conversely, a non-polarimetric radar provides information only on backscattered power, allowing access solely to reflectivity. However, converting polarimetric radar variables into accurate rainfall rate estimates remains an open research area, requiring robust and adaptive approaches.

Moreover, numerous studies have demonstrated the significant contribution of polarimetric radars to improving precipitation estimation. Indeed, [17] highlighted that the integration of these variables allows for better microphysical characterization of hydrometeors, thereby improving quantitative precipitation estimation. Thus, [18] introduced dual-parameter precipitation estimates using differential reflectivity ( $Z_{dr}$ ) and specific differential phase ( $K_{dp}$ ) to address the precipitation estimation problem by achieving a better characterization of the raindrop size distribution. Some authors [19]-[21] have also used polarimetric radar variables ( $Z_h$ ,  $Z_{dr}$ ,  $K_{dp}$ ) in a multi-parametric approach to estimate precipitation. Works such as those by [22] [23] have confirmed that relationships combining  $Z_h$ ,  $Z_{dr}$  and  $K_{dp}$  generally offer more robust estimates, particularly in complex meteorological situations such as intense convective showers. Although multi-parametric radars have improved precipitation estimates, they remain sensitive to the variability of the Drop Size Distribution (DSD) and microphysical

processes, just like traditional empirical  $Z_h - R$  laws. Furthermore, for very intense rainfall, parametric precipitation estimation models are generally less suitable and less robust [24]-[27].

In parallel, machine learning-based methods, notably deep neural networks (deep learning), have recently distinguished themselves as particularly effective tools for precipitation estimation from radar data [28]-[32]. Among these authors, some, such as [28] have demonstrated that Artificial Neural Networks (ANNs), particularly Multi-Layer Perceptron (MLP) architectures, outperform traditional empirical models due to their ability to capture complex non-linear relationships between radar variables and ground rainfall. Furthermore, the introduction of deep learning techniques has led to notable improvements in model generalization, robustness to noise, and adaptability to different meteorological contexts [33]. Artificial neural networks are used to address highly complex and non-linear phenomena. In the literature, artificial neural networks have shown great success in statistical analysis, data processing, forecasting, and the estimation of environmental parameters. The advantage of using neural networks in radar rainfall estimation lies in their ability to account for the non-linearity of the relationship between rainfall and radar observables. Indeed, without assuming parametric relationships between variables, this precipitation estimation technique using artificial neural networks extracts representations from input data and directly associates polarimetric radar variables with ground rain gauge measurements. Through training based on rain gauge measurements (the target), the MLP models the link between the received data and the actual rainfall at ground level.

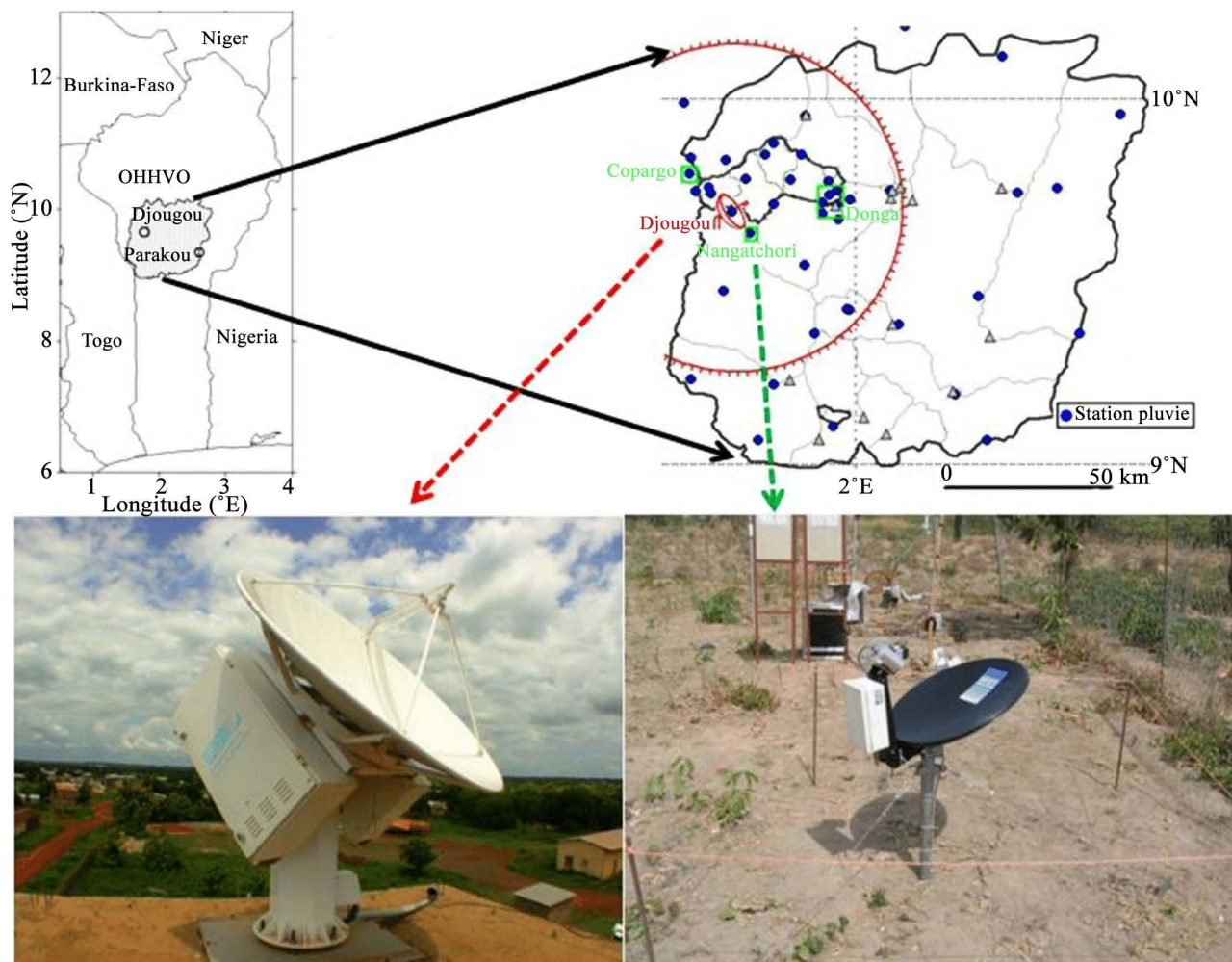
This study aims to develop a Deep Learning-based Multilayer Perceptron (MLP) model for quantitative precipitation estimation (QPE) using polarimetric radar variables, specifically horizontal reflectivity ( $Z_h$ ), differential reflectivity ( $Z_{dr}$ ) and specific differential phase ( $K_{dp}$ ) collected during the AMMA (African Monsoon Multidisciplinary Analysis) field campaign in northern Benin. The objective is to assess whether a data-driven neural architecture can outperform conventional parametric radar-rainfall estimation algorithms commonly used in tropical environments.

The remainder of the manuscript is organized as follows. Section 2 describes the experimental framework, including the radar instrumentation, rainfall measurements, and the preprocessing procedures applied to the dataset. It also details the methodologies employed for developing the artificial neural network and the approaches used for performance evaluation. Section 3 presents the results of the MLP model, along with a comparative analysis against traditional parametric algorithms, including standard  $Z_h - R$  relationships and multiparametric or microphysics-based estimators. This section also discusses the model's predictive accuracy and the diagnostic analyses performed to assess error structure and potential biases. Finally, section 4 provides the concluding remarks, summarizing the key findings of the study and outlining the implications for operational hydrometeorological applications as well as directions for future research.

## 2. Materials and Methods

### 2.1. Experimental Setup

The data used for this study were collected in 2006 and 2007, during the intensive phase of the African Monsoon Multidisciplinary Analysis (AMMA) campaign at the OHHVO site (Upper Ouémé Valley Hydrometeorological Observatory) of AMMA-CATCH (Coupling the Tropical Atmosphere and the Hydrological Cycle), specifically in northern Benin in the tropical zone (**Figure 1**). One of the objectives of the AMMA campaign was to analyze mesoscale convective systems and precipitation in northern Benin. This mesoscale site covered an area of approximately 10,000 km<sup>2</sup> equipped with a network of fifty-four (54) tipping-bucket rain gauges [20] [34]. The instrumentation of this site included, among others, an X-band polarimetric radar, a vertical Micro Rain Radar (MRR), as well as several optical disdrometers [12]. The X-band polarimetric radar (Xport radar), which



**Figure 1.** AMMA-CATCH experimental area in Benin showing the measuring instruments installed. Top (on the left) we show a view of Benin (West Africa) and (on the right) the AMMA-CATCH site. Rain gauges are indicated by black dots. Bottom: (on the left), the Xport radar, positioned in Djougou; (on the right), the Micro Rain Radar (MRR) in Nangatchori (10 km far from radar Xport).

was located in the city of Djougou (1.66°E, 9.69°N), is a dual-polarization polarimetric radar. It offers a significant advantage over single-polarization systems, as it allows for multi-parameter measurements using orthogonal polarizations.

The experimental area was located in a Sudanian climate with an annual rainfall of approximately 1200 mm. In this region, the majority of precipitation is of convective origin and stems mainly from organized mesoscale convective systems (MCSs), supplemented by more local and sporadic convection [12] [34].

The choice of the 2006-2007 AMMA (African Monsoon Multidisciplinary Analysis) dataset is justified by its unique status as the most comprehensive and reliable high-resolution polarimetric reference ever collected for West African squall lines. Despite the time elapsed since the campaign, this dataset provides an exceptional density of synchronized radar and rain gauge measurements, which are essential for training complex models like the MLP. Given that the fundamental microphysical processes governing tropical convection in the Sahelian and Sudanian regions have not fundamentally changed, the 2006-2007 observations remain highly representative of current meteorological dynamics. We consider this work a critical baseline study, establishing a robust benchmark for machine learning applications in the region, while acknowledging that future validation with more recent datasets will be a natural extension of this research as such data become available.

## 2.2. Data Description

In this study, we focused on the polarimetric observables retrieved by the X-band radar and the rainfall rates measured by the rain gauge network and aggregated at 5-min time steps. The polarimetric observables considered and used in this study are: horizontal reflectivity ( $Z_h$ ), differential reflectivity ( $Z_{dr}$ ), and specific differential phase ( $K_{dp}$ ), measured at low elevation. These polarimetric radar variables were pre-processed [21]. Specifically, the data namely horizontal ( $Z_h$ ) and differential ( $Z_{dr}$ ) reflectivities were corrected for attenuation using the method proposed by [35], which relies on a linear relationship between total attenuation along a given radial and the total differential phase ( $\Phi_{dp}$ ), i.e., the integral of the specific differential phase ( $K_{dp}$ ). The correction coefficients used are 0.285 dB/deg and 0.051 dB/deg for horizontal and differential reflectivities, respectively.

## 2.3. Methodology

For this study, due to the nature of our data, we favored a fully connected deep neural network approach, capable of capturing the complex non-linear relationships linking radar features to in situ rain gauge observations. Indeed, our input data are mutually independent, without marked spatial or temporal correlation. Furthermore, our data were derived from physical measurements, and each observation was represented by a fixed feature vector ( $Z_h, Z_{dr}, K_{dp}, R$ ).

The input vector for the Multi-Layer Perceptron (MLP) model is composed of three primary polarimetric variables: horizontal reflectivity ( $Z_h$ ), differential re-

flectivity ( $Z_{dr}$ ), and specific differential phase ( $K_{dp}$ ). In the context of this study, these are considered “derived features” as they are not the raw signal backscattered by the particles, but are instead parameters calculated through signal processing to represent the physical and microphysical properties of the hydrometeors. Specifically,  $Z_h$  provides information on the intensity and concentration of the particles,  $Z_{dr}$  characterizes their mean shape (aspect ratio), and  $K_{dp}$  is sensitive to the liquid water content and less affected by attenuation. To ensure consistency across the training and validation phases, each feature vector  $[Z_h, K_{dp}, K_{dp}]$  undergoes a normalization process prior to being fed into the neural network, allowing the model to effectively map the non-linear relationship between these polarimetric descriptors and the target rainfall rate ( $R$ ).

### 2.3.1. Data Preprocessing and Analysis

Both model input and target data underwent preliminary processing to assess the suitability of using deep learning for this precipitation estimation task. Subsequently, the training and testing data were normalized at the feature level. For this task, we used Z-score normalization [36], using the mean and standard deviation of the features as normalization parameters.

Z-score normalization, also known as standardization, is indeed an essential step before training a neural network (such as an MLP), as it places all variables on the same scale. The normalization parameters, namely the mean  $\mu$  and the standard deviation  $\sigma$ , were calculated solely using the training data. Thus, for a variable  $x$  with mean  $\mu$  and standard deviation  $\sigma$ , the normalized variable  $z$  is expressed as follows:

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

Subsequently, to convert the normalized variable  $z$  back to the initial variable  $x$ , the following is used:

For the training of the MLP model, 80% of the data were used for training and 20% for testing. Furthermore, 30% of the training data were selected for model validation.

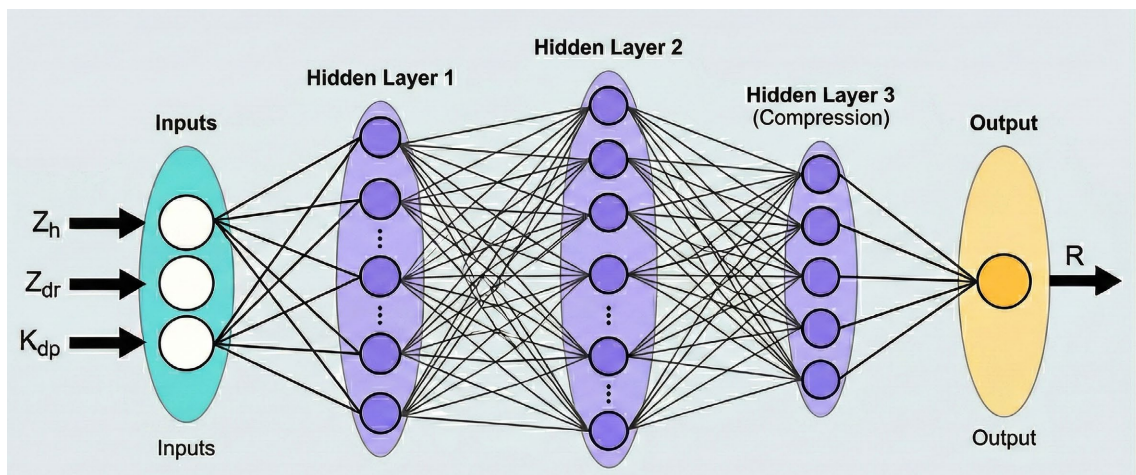
The dataset was partitioned into a training set (80%) and a testing set (20%) using a random sampling approach. It is important to distinguish that this study addresses Quantitative Precipitation Estimation (QPE), which is a contemporaneous regression task, as opposed to forecasting. Since our pre-processing ensured a unique 1-to-1 matchup between each rain gauge and its corresponding radar pixel at each time step, random shuffling does not introduce spatial data leakage. This method allows the model to encounter a balanced variety of precipitation microphysics from different stages of various convective systems, enhancing its generalization capabilities.

### 2.3.2. Model Architecture

The deep learning model developed in this study relies on a multilayer, fully connected neural network architecture, designed to perform Quantitative Precipita-

tion Estimation (QPE) from polarimetric radar data. The network is implemented using the Sequential API of the Keras library (Tensorflow backend) and comprises four fully connected layers (Figure 2). The architecture of the MLP model was structured as follows:

- Use a first hidden layer of 64 neurons, using a ReLU (Rectified Linear Unit) activation function combined with L2 regularization with a coefficient  $\lambda = 1 \times 10^{-4}$  to limit model complexity and prevent overfitting;
- Avoid a second, deeper hidden layer, comprising 128 neurons also activated by ReLU and subject to L2 regularization with the same coefficient  $\lambda = 1 \times 10^{-4}$ ;
- Do a third hidden layer of 16 neurons, still with ReLU activation, without explicit regularization. This layer plays a role in information compression, facilitating the synthesis of learned representations before the output;
- Use an output layer composed of a single neuron with linear activation, suitable for a regression task aiming to predict the continuous precipitation rate expressed in  $\text{mm}\cdot\text{h}^{-1}$ .



**Figure 2.** Architecture of the Multi-Layer Perceptron (MLP) model developed for rainfall estimation from radar polarimetric variables.

This architecture was selected after multiple tests for its ability to model highly non-linear relationships between polarimetric radar variables and ground-measured rainfall intensities, while maintaining controlled complexity through L2 regularization. The combination of these layers enables the network to learn both detailed and generalized representations.

### 2.3.3. Model Compilation and Training

The model was compiled with the Adam (Adaptive Moment Estimation) optimizer with a learning rate of  $10^{-4}$ . This choice, more conservative than the default value of  $10^{-3}$ , aimed to ensure a more stable gradient descent, which is particularly useful in the context of non-linear regression on noisy or highly variable data such as those derived from radar measurements, which are characteristic of our dataset. Our dataset also exhibited a strong imbalance, with a scarcity of high precipitation

rate values.

The loss function selected for training was the Mean Squared Error (MSE), which is appropriate for continuous regression tasks such as the precipitation rate estimation addressed in this study. In addition, the coefficient of determination ( $R^2$ ) was also used to evaluate the quality of the model's predictions. This latter statistical indicator quantifies the proportion of variance explained by the model, thus providing a global indication of the quality of the regression.

Training was performed over 100 epochs with a batch size of 32, representing an effective compromise between convergence speed and optimization stability. To control overfitting, 30% of the training data were set aside for validation. Furthermore, a set of callbacks (such as early stopping, learning rate reduction in case of stagnation, and saving the best model) was used to optimize the training process and prevent performance degradation on validation data.

#### 2.3.4. Heavy Rainfall Detection

Reliable detection of heavy rainfall episodes constitutes a major challenge in the operational evaluation of precipitation estimation models. In this study, the performance of the MLP model was examined not only in terms of mean error and overall correlation, but also through its ability to discriminate between intense and light rainfall events. This was evaluated using the Receiver Operating Characteristic (ROC) curve [37], which is a proven method for analyzing the detection of meteorological events.

The ROC curve illustrates the trade-off between the True Positive Rate (TPR), i.e., the proportion of heavy rainfall correctly identified, and the False Positive Rate (FPR), corresponding to light rainfall episodes misclassified as intense. Each point on the curve corresponds to a different probability threshold applied to the MLP model outputs. The closer the curve rises toward the upper left corner of the graph, the higher the model's discrimination capability.

Our case, heavy rainfall probabilities were derived from the continuous precipitation values estimated by the MLP model by defining a critical threshold  $R_c = 20 \text{ mm} \cdot \text{h}^{-1}$ . Predictions above this threshold were considered as heavy rainfall, and those below as moderate or light rainfall. For each threshold, the (TPR, FPR) pairs were calculated, allowing the corresponding ROC curve to be plotted.

The AUC (Area Under the Curve) score was selected as a key indicator to globally quantify the model's classification performance, regardless of the chosen decision threshold [38] [39]. The AUC provides an aggregate measure of class separability: a value of 1 indicates perfect discrimination between heavy and light rainfall episodes. Conversely, an AUC of 0.5 corresponds to performance equivalent to chance (random behavior), as demonstrated by [40]. A score below 0.5 would indicate a model performing worse than chance, generally suggesting a systematic inversion of predicted classes.

From an operational standpoint, the ROC curve also allows for tuning the optimal decision threshold according to the intended objective:

- Maximizing detection (increasing TPR) at the expense of a slight increase in

false alarms;

- Prioritizing reliability (reducing FPR) for sensitive applications such as flood forecasting;
- Do a third hidden layer of 16 neurons, still with ReLU activation, without explicit regularization. This layer plays a role in information compression, facilitating the synthesis of learned representations before the output;
- Use an output layer composed of a single neuron with linear activation, suitable for a regression task aiming to predict the continuous precipitation rate expressed in  $\text{mm}\cdot\text{h}^{-1}$ .

This flexibility is essential for calibrating the estimation system according to the local hydrometeorological context and the risk tolerance of end-users.

Thus, the ROC curve proves to be a valuable complementary tool for evaluating the MLP model, offering a more operational insight into its performance than continuous metrics alone (RMSE, MAE, or  $R^2$ ).

### 2.3.5. Comparative Study with Parametric Algorithms

The MLP deep learning model developed in the present study was compared to parametric rainfall estimation algorithms such as the  $Z_h - R$ ,  $R - K_{dp}$ ,  $R - Z_{dr}K_{dp}$ , and  $R - Z_h Z_{dr} K_{dp}$  estimators, all implemented in West Africa. The performance of the different models was evaluated using five metrics used to assess absolute error, relative error, systematic bias, and structural fidelity between observed and predicted values. The metrics used are defined according to the following formulas:

#### 1) Mean Absolute Error (MAE):

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) \quad (3)$$

MAE measures the mean absolute error between observations  $y_i$  and predictions  $\hat{y}_i$ . It is expressed in the same unit as the target variable. Less sensitive to extreme values than RMSE, it generally reflects a typical error without over-weighting large errors.

#### 2) Normalized Root Mean Square Error par l'écart-type (NRMSE)

$$\text{NRMSE} = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}}{\sigma_y} \quad (4)$$

where  $\sigma_y$  is the standard deviation of the observations. This metric expresses the root mean square error relative to the natural variability of the data. It is dimensionless and allows for the comparison of the performance of different models.

#### 3) Normalized Bias (NBias%):

$$\text{NBias}\% = 100 \times \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)}{\bar{y}} \quad (5)$$

The NBias% quantifies the systematic tendency of the model to overestimate (positive values) or underestimate (negative values) the observations. It is ex-

pressed as a percentage of the mean of the observations. A value close to 0% indicates a globally unbiased model.

#### 4) Pearson Correlation Coefficient ( $r$ ):

$$r = \frac{\sum_1^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_1^n (y_i - \bar{y})^2} \sqrt{\sum_1^n (\hat{y}_i - \bar{\hat{y}})^2}} \quad (6)$$

The coefficient  $r$  measures the strength and direction of the linear relationship between observed and predicted values. It ranges between  $-1$  and  $+1$ . A value close to  $1$  indicates a strong linear correlation, without necessarily implying a low error.

### 2.3.6. Statistical Calibration and Baseline Models

While the initial comparisons between our MLP model and parametric algorithms were conducted using existing relations from our study area, it is valuable to employ a complementary approach for a more extensive evaluation. Consequently, to ensure a more rigorous and fair assessment of the Multi-Layer Perceptron (MLP), the coefficients of the traditional parametric relationships ( $R - Z_h$ ,  $R - K_{dp}$ ,  $R - Z_h$ ,  $Z_{dr}$  and  $R - Z_h$ ,  $Z_{dr}$ ,  $K_{dp}$ ) were specifically re-optimized for our study site. Rather than relying solely on the standard coefficients previously utilized, we performed a local statistical calibration.

The optimization process was carried out as follows:

- Objective Function: A non-linear least-squares regression was employed to determine the optimal coefficients for the power-law equations.
- Data Isolation: Calibration was performed exclusively using the training dataset, ensuring that the baseline models were exposed to the same environmental variability as the MLP.
- Optimization Criterion: The coefficients were adjusted to minimize the Root Mean Square Error (RMSE) between radar-derived estimates and ground-based rain gauge measurements at an instantaneous time scale.

This approach aims to confirm that the MLP's ability to handle non-linearities and multi-parameter inputs provides a significant advantage, even when compared to the most highly optimized traditional equations.

### 2.3.7. Diebold-Mariano Test

In order to rigorously assess whether the observed performance differences between the MLP model and the reference parametric algorithms are statistically significant, we used the Diebold-Mariano test or DM test [41]. This test constitutes a standard statistical tool for comparing the predictive accuracy of two competing models on the same dataset, while accounting for potential correlation between errors and the forecast horizon.

The principle of the DM test is based on the analysis of the time series of loss differentials between two models, defined as:

$$d_t = L(e_{1,t}) - L(e_{2,t}) \quad (7)$$

where  $e_{1,t}$  and  $e_{2,t}$  represent, respectively, the forecast errors of model 1 (here, the neural model) and model 2 (here, the other traditional rainfall estimation algorithms) at time  $t$ , and  $L(\cdot)$  is an appropriate loss function (here, the squared error  $e^2$  to evaluate accuracy based on RMSE).

Under the null hypothesis  $H_0$  (both models have the same average accuracy), the expected value of the loss differentials is zero ( $E[d_t] = 0$ ), which implies equivalent predictive accuracy between the two models. The test statistic is given by:

$$DM = \frac{\bar{d}}{\sqrt{\frac{\hat{\gamma}_0 + 2 \sum_{k=1}^{h-1} \hat{\gamma}_k}{T}}} \quad (8)$$

where  $\bar{d}$  is the sample mean of  $d_t$ ,  $\hat{\gamma}_k$  denotes the autocovariance at lag  $k$  of the series  $d_t$ ,  $h$  is the forecast horizon (here  $h = 1$ ), and  $T$  is the sample size. Under the null hypothesis  $H_0$ , the DM statistic asymptotically follows a standard normal distribution.

A high absolute value of the DM statistic, accompanied by a p-value below the chosen significance level (generally  $\alpha = 5\%$ ), leads to the rejection of  $H_0$  in favor of the alternative hypothesis  $H_1$ , and indicates that one of the two models exhibits significantly superior predictive accuracy. The interpretation of the sign of the DM statistic is straightforward: a negative value indicates that Model 1 has a lower mean loss (and thus better performance) than Model 2, while a positive value indicates the reverse.

The application of this test in our study enabled us not only to confirm the superiority of the MLP model over parametric approaches, but also to ensure that the observed differences are not attributable to chance, thereby strengthening the robustness of the conclusions.

Recent applications of the test in the hydrometeorological field have confirmed its relevance for objectively comparing performance between physical models and machine learning models. For instance, while [41] established the statistical framework, recent studies such as [42]-[44] have employed the DM test to validate the statistical superiority of deep learning models in rainfall estimation compared to conventional parametric methods.

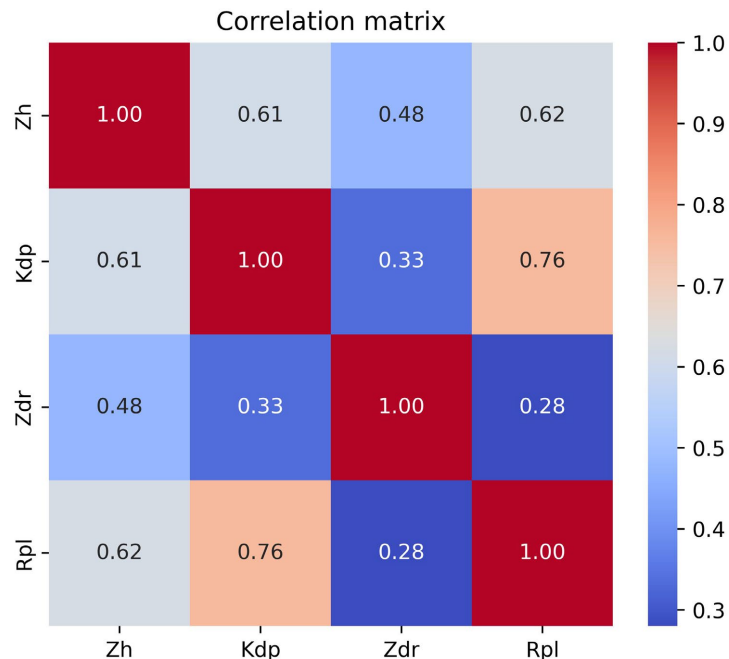
### 3. Results and Discussion

#### 3.1. Preliminary Data Analysis

A preliminary data analysis was conducted to identify the main statistical characteristics of the variables used for model training.

**Figure 3** presents the correlation matrix between the radar variables ( $Z_h$ ,  $Z_{dr}$ ,  $K_{dp}$ ) and the observed rainfall ( $R_{pl}$ ). The results highlighted a particularly high linear correlation between the specific differential phase ( $K_{dp}$ ) and  $R_{pl}$  with  $r = 0.75$ , followed by horizontal reflectivity ( $Z_h$ ) with a moderate correlation, while differential reflectivity ( $Z_{dr}$ ) showed a weaker correlation. These observa-

tions confirm the value of combining multiple radar parameters to improve precipitation estimation, with  $K_{dp}$  playing a central role, particularly for high intensities. Furthermore, the moderate cross-correlations between  $Z_h$ ,  $Z_{dr}$  and  $K_{dp}$  suggest a complementarity that can be exploited in a deep learning context.

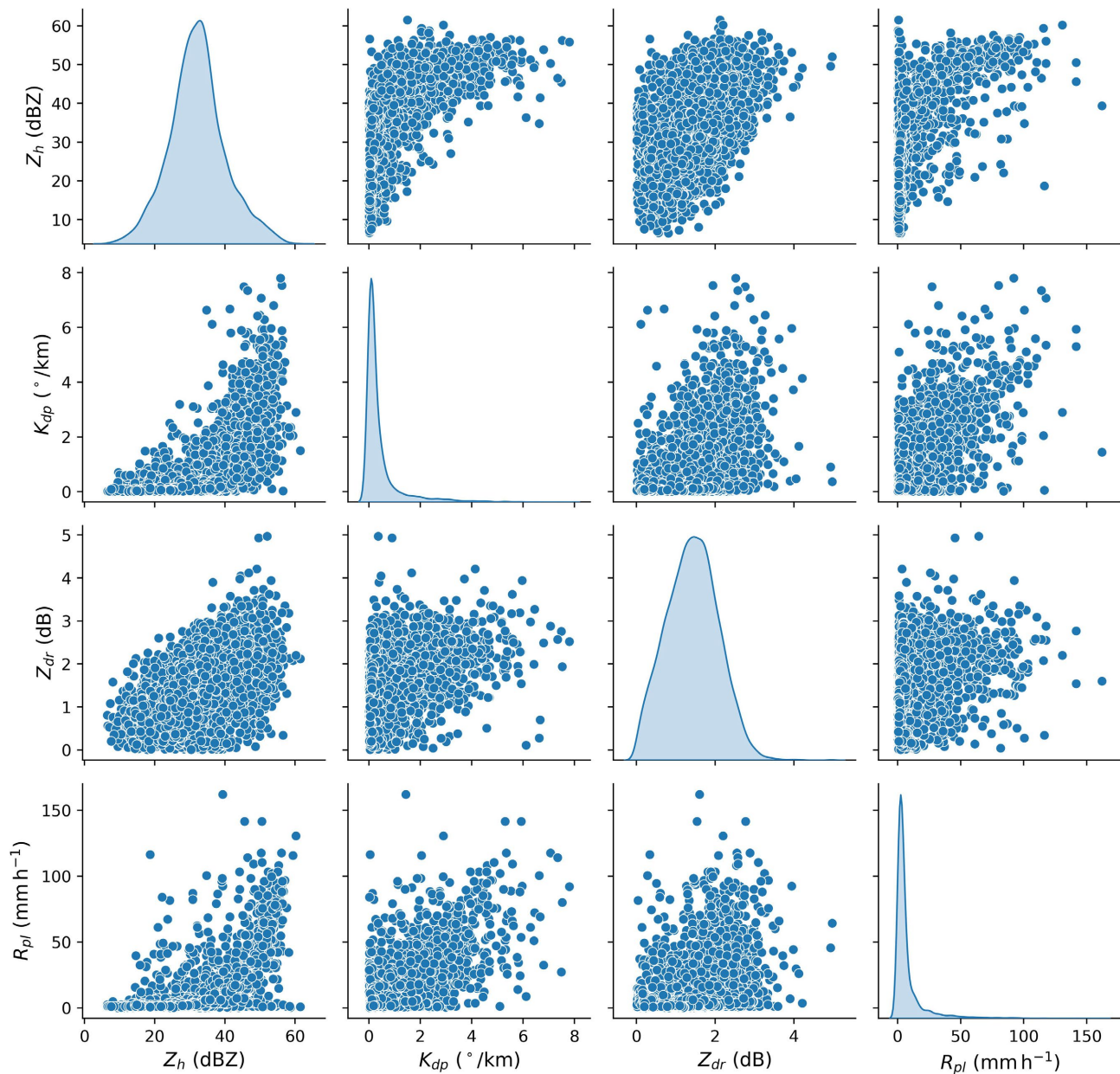


**Figure 3.** Correlation matrix.

These observations derived from the correlation matrix analysis corroborate the results of [21], which show that a polarimetric X-band radar, despite its attenuation constraints, can provide reliable precipitation estimates by relying on differential variables such as  $K_{dp}$ . Furthermore, [27] also analyzed the respective contributions of  $Z_h$ ,  $Z_{dr}$  and  $K_{dp}$  to rainfall rate estimation from simulated DSD data, highlighting the superiority of  $K_{dp}$  in many cases. Moreover, the work of [45] remains a classic reference in this field, already illustrating how radar relationships ( $Z_h$ ,  $Z_{dr}$ ,  $K_{dp}$ ) can be combined for precipitation estimation. Finally, [32] emphasized that in tropical environments, polarimetric observables such as  $K_{dp}$  and  $Z_{dr}$  are particularly valuable for rainfall estimation from X-band radars.

**Figure 4** presents a radar-rain pair plot, allowing for the simultaneous visualization of marginal distributions (on the diagonal) and bivariate relationships (off-diagonal) between all variable pairs. Examination of these distributions revealed a strong skewness in precipitation, a marked correlation between  $K_{dp}$  and  $R_{pl}$  as well as significant scatter for heavy rainfall, particularly between  $Z_h$  and  $R_{pl}$ . These findings reinforce the relevance of a deep learning-based approach, capable of modeling non-linear relationships and leveraging the complementarity of input variables.

In the recent field of deep learning applied to Quantitative Precipitation



**Figure 4.** Pairwise relationships between radar polarimetric variables and rainfall rate.

Estimation (QPE) from polarimetric radars, several studies have confirmed this potential. For example, [46] proposed the DQPENet model, an architecture based on dense blocks for radar QPE, using variables  $Z_h$ ,  $Z_{dr}$  and  $K_{dp}$ , with model interpretations via permutation tests. [47] used a multiscale Convolutional Neural Network (CNN) on radar variables to estimate precipitation during typhoons, leading to a significant improvement in bias and RMSE compared to classical algorithms based on DSD. Finally, [48] proposed a deep learning architecture for rainfall estimation from simulated dual-polarization radars, confirming that non-linear models can outperform fixed  $Z_h$ - $R$  relationships in certain scenarios.

However, the highly imbalanced distribution of rainfall intensities, dominated

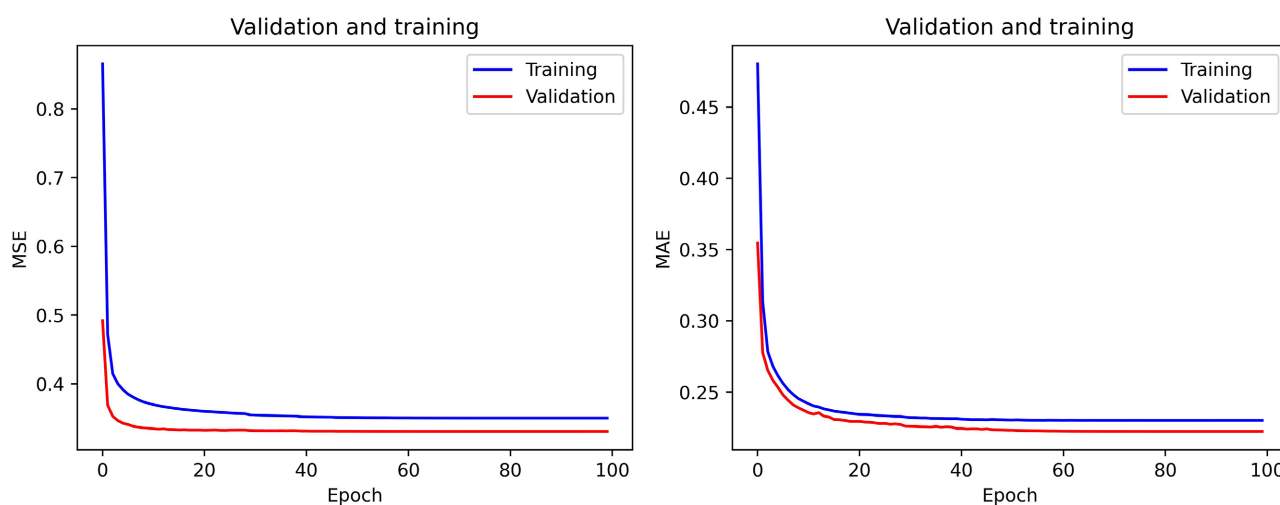
by light rainfall, limits the model's ability to predict heavy precipitation. A targeted data augmentation strategy for the latter was therefore tested. Nevertheless, this resampling did not yield the expected gains: statistical indicators (MAE and MSE) generally deteriorated, with the exception of the coefficient of determination ( $R^2$ ), which recorded a slight improvement of  $\Delta R = +0.03$ . In the literature on machine learning applied to rare phenomena (e.g., in hydrology or meteorology), this type of class imbalance is well known to introduce significant biases in Mean Squared Errors (MSE), while global metrics like  $R^2$  can be subject to partial overfitting effects on rare events. We can cite general works [49] on handling imbalance in machine learning (oversampling, weighted loss, etc.) in the context of time series or extreme events, although such studies remain scarce in radar Quantitative Precipitation Estimation (QPE). They suggested that simple oversampling can produce artifacts or amplify noise in rare data.

Thus, model training was conducted using real data retrieved by the Xport radar for polarimetric variables and data measured by rain gauges for target variables.

### 3.2. Model Training

**Figure 5** displays the model's learning curves for two metrics: the Mean Squared Error MSE (**Figure 5**, left panel) and the Mean Absolute Error MAE (**Figure 5**, right panel) as a function of the number of epochs for both the training and validation sets.

The learning curves indicate that both metrics decreased rapidly during the initial iterations before stabilizing at approximately the 40th epoch. This trend was observed for both the training and validation sets, with comparable values between the two. This behavior indicates that the model converges towards an optimal solution while maintaining good generalization capability. The absence of significant divergence between the validation and training curves suggests that the model suffers from neither underfitting nor overfitting.



**Figure 5.** Evolution of training and validation losses during the MLP model learning process. The plots show the evolution of the Mean Squared Error (MSE) (left) and Mean Absolute Error (MAE) (right) over 100 training.

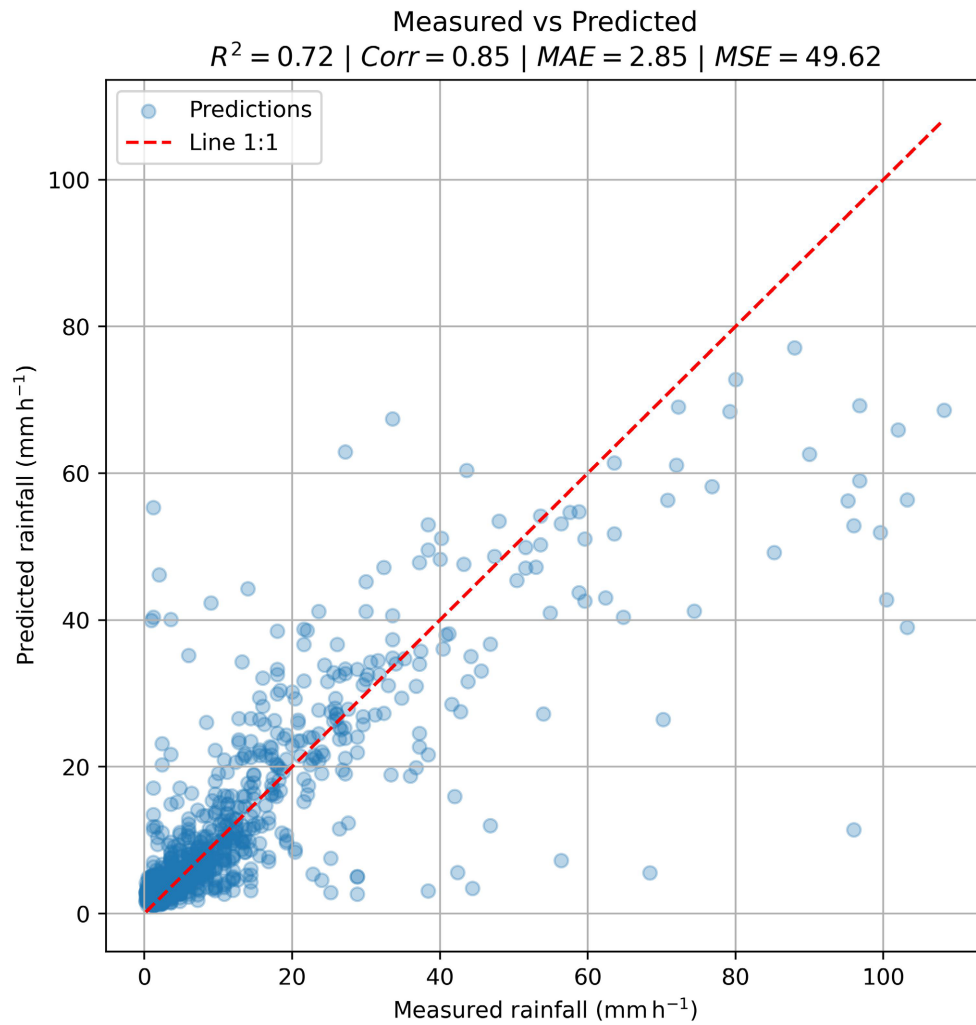
This trend of rapid convergence followed by the stabilization of error metrics aligns with behaviors frequently observed in radar QPE studies relying on deep neural networks. For instance, [50] demonstrated that, in a radar-precipitation application, errors (MSE or MAE) decreased rapidly during the initial epochs and then plateaued towards the end of training, reflecting efficient and well-regularized learning.

Furthermore, [51] in the context of radar bias correction using deep learning, showed that their learning curves remained relatively close between the training and validation sets. This supported the hypothesis that the model was not overfitting for heavy rainfall events, despite class imbalance. Thus, the observed behavior (rapid error descent followed by a stable plateau, with close training/validation curves) is consistent with expectations for a well-trained and well-regularized model in the context of deep learning-based QPE. Finally, the variation in MSE and MAE during training and validation attests to the robustness of the chosen architecture and the effectiveness of the adopted training strategy.

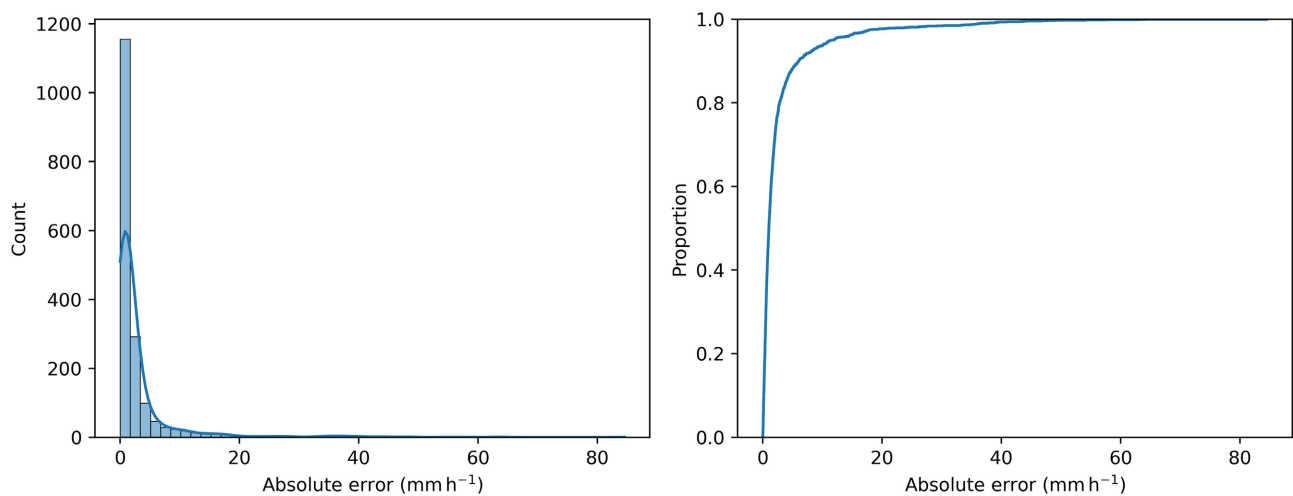
### 3.3. Model Evaluation

**Figure 6** displays a scatter plot illustrating the relationship between precipitation rates measured by rain gauges and those predicted by the MLP model (in  $\text{mm}\cdot\text{h}^{-1}$ ) relative to the 1:1 line. The results indicated a Pearson correlation of  $r = 0.85$  and a coefficient of determination of  $R^2 = 0.72$ , demonstrating the model's good overall accuracy. The Mean Absolute Error (MAE) was  $2.85 \text{ mm}\cdot\text{h}^{-1}$ , whereas the Mean Squared Error (MSE) was  $46.62 \text{ mm}^2\cdot\text{h}^{-2}$ . Most data points clustered near the 1:1 line for intensities below  $20 \text{ mm}\cdot\text{h}^{-1}$  (light to moderate intensities). Consequently, the majority of estimates aligned well with observations within this intensity range. This reflects the model's capability to effectively capture common rainfall events, consistent with numerous studies on radar Quantitative Precipitation Estimation (QPE), where alignment typically improves for light to moderate intensities.

However, underestimation was observed for high intensities, with greater scatter of predictions relative to the 1:1 line beyond  $40 \text{ mm}\cdot\text{h}^{-1}$ . This behavior is typical when nonlinear regression models encounter marked dataset imbalance and the high variability of intense precipitation. Several studies [13] [21] [52] among others, have reported negative biases and increased scatter for heavy rain in radar-gauge validations within tropical environments. These studies highlight the key role of polarimetric observables, particularly  $K_{dp}$ , in improving robustness without entirely eliminating difficulties at the extremes. The distribution of absolute errors (**Figure 7**, left panel) revealed a strong concentration of predictions around small deviations from observations, with a marked peak for errors below  $5 \text{ mm}\cdot\text{h}^{-1}$ . Beyond this threshold, frequency decreased rapidly, and errors exceeding  $20 \text{ mm}\cdot\text{h}^{-1}$  appeared as isolated cases, likely attributable to extreme precipitation episodes or specific measurement conditions. The corresponding Cumulative Distribution Function (CDF) (**Figure 7**, right panel) shows that



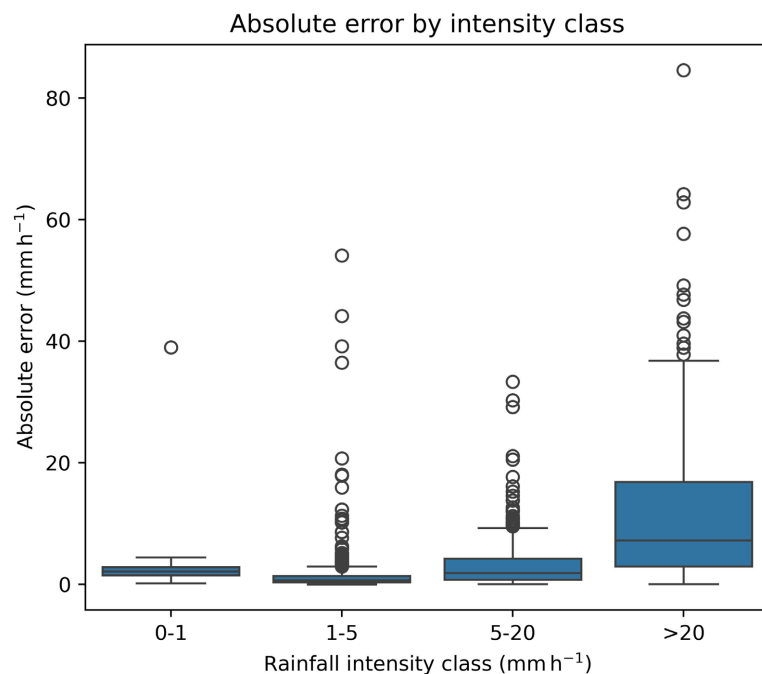
**Figure 6.** Comparison between measured and MLP-predicted rainfall rates. The scatter plot compares observed rainfall rates (x-axis) and those predicted by the MLP model (y-axis). The red dashed line represents the 1:1 reference line.



**Figure 7.** Distribution and cumulative distribution of absolute rainfall estimation errors from the MLP model. The left panel shows the histogram of absolute errors (in  $\text{mm}\cdot\text{h}^{-1}$ ) between predicted and observed rainfall rates, while the right panel presents the corresponding cumulative distribution function (CDF).

approximately 80% of predictions exhibited an error lower than roughly 5 mm·h<sup>-1</sup>, and over 95% remained below a threshold of approximately 10 mm·h<sup>-1</sup>. The rapid stabilization of the curve beyond 20 mm·h<sup>-1</sup> confirms the rarity of large errors.

The evaluation of absolute errors by rainfall intensity class (Figure 8) highlights the model's distinct behavior depending on precipitation magnitude. For light intensities (0 - 1 mm·h<sup>-1</sup> and 1 - 5 mm·h<sup>-1</sup>), the median error remained below 2 mm·h<sup>-1</sup> with limited scatter, demonstrating the model's satisfactory ability to reproduce these common situations. Errors for the intermediate class (5 - 20 mm·h<sup>-1</sup>) remained moderate but exhibited increased variability, reflecting a substantial capability in modeling these events. However, for very heavy rainfall (>20 mm·h<sup>-1</sup>), performance deteriorated significantly, with an error median exceeding 10 mm·h<sup>-1</sup> and extreme values surpassing 80 mm·h<sup>-1</sup>. This heteroscedasticity highlights the recurrent difficulty for learning-based approaches in correctly capturing extreme episodes, which are generally under-represented in training data.



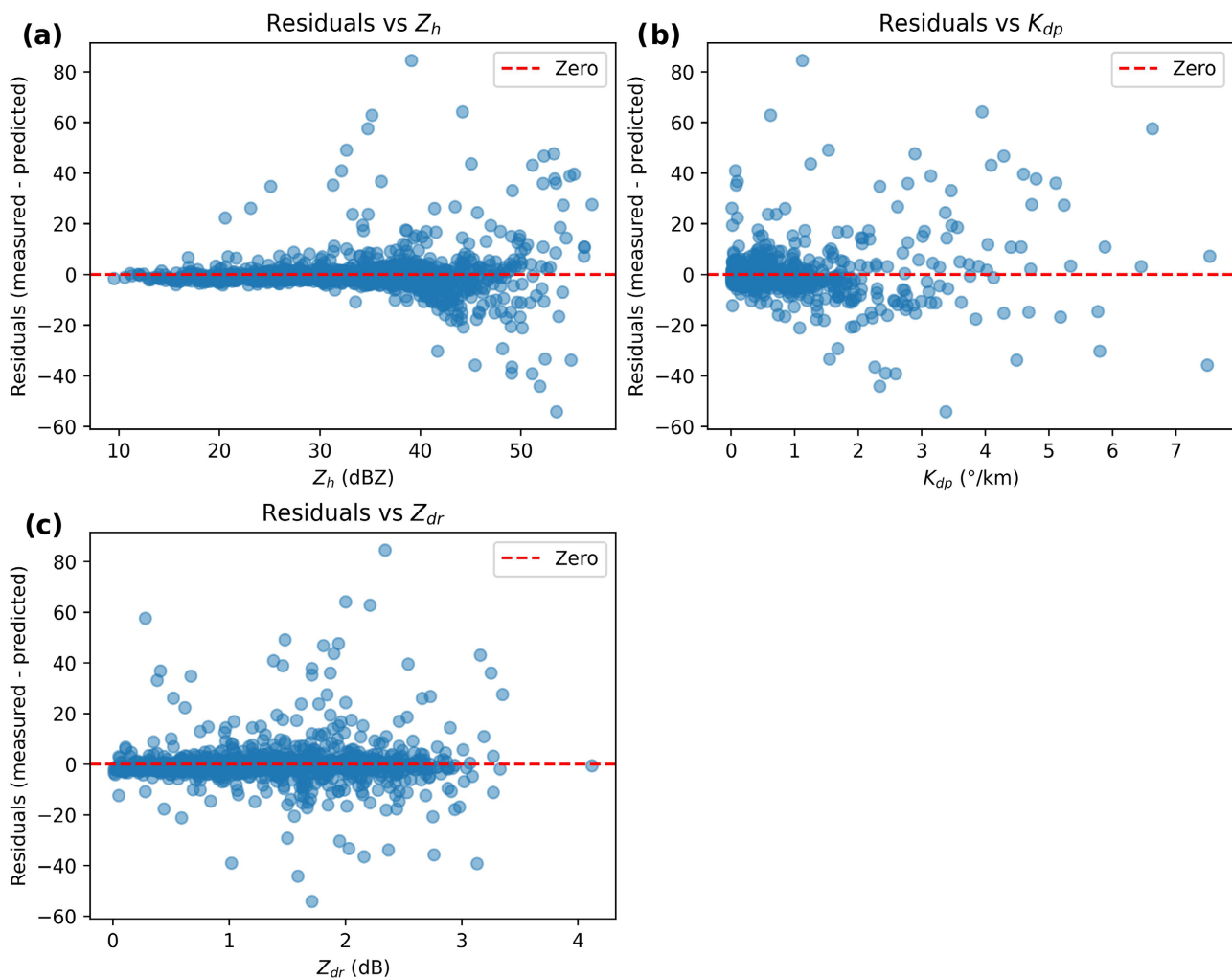
**Figure 8.** Distribution of absolute rainfall estimation errors by rainfall intensity class.

Figure 9 presents the analysis of residuals ( $R_{measured} - R_{predicted}$ ) as a function of each radar variable ( $Z_h$ ,  $Z_{dr}$ ,  $K_{dp}$ ). On average, residuals were centered around zero, confirming the absence of a global systematic bias. For  $Z_h$ , residuals remained constrained up to 35 dBZ, but a notable increase in variability appeared beyond 40 dBZ. Regarding  $K_{dp}$ , the majority of data points were concentrated in the range below 2°/km, with moderate deviations, whereas high values (>4°/km) were associated with increased scatter, which is characteristic of intense rainfall. Conversely, for  $Z_{dr}$ , residuals remained symmetrically distributed around zero, indicating that the model effectively accounts for this parameter. These observa-

tions corroborate the previous diagnostics, namely an underestimation of extreme rainfall episodes.

The results presented in **Figures 6-9** suggest that, while the model demonstrates robust overall performance, specific optimizations are necessary to improve prediction under extreme conditions, particularly in the presence of high reflectivity and intense convective events.

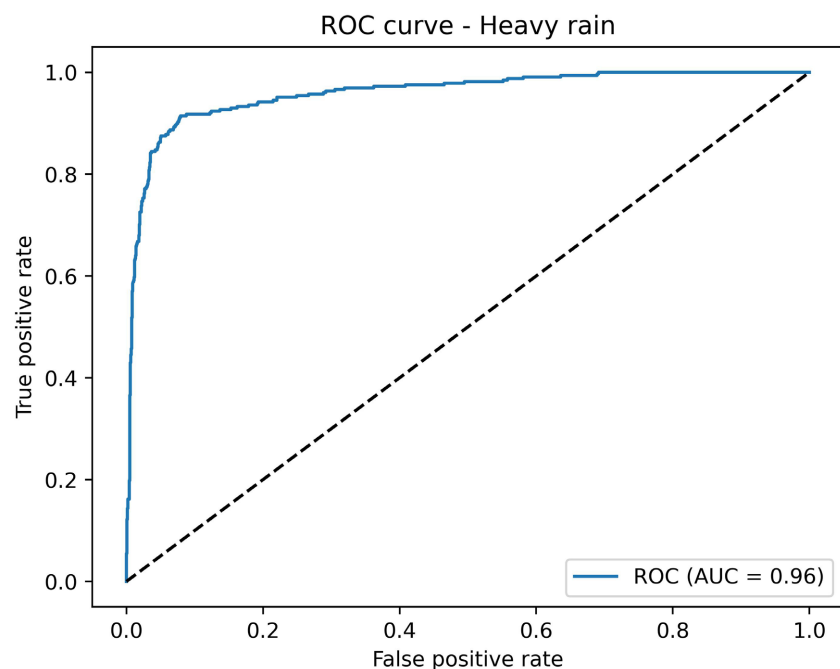
Finally, although neural networks (ANN/MLP, CNN) applied to polarimetric variables ( $Z_h$ ,  $Z_{dr}$ ,  $K_{dp}$ ) improve QPE compared to parametric schemes, recent literature focusing on West Africa [53]-[56] indicates that extremes remain more difficult to predict without bias. This observation is likely attributable to the statistical rarity of high rainfall rates, sampling imbalance, and measurement uncertainties. These conclusions align with our findings and reinforce the premise that the MLP provides reliable estimates across the dominant intensity range, despite a tendency to underestimate the highest intensities.



**Figure 9.** Residuals of the MLP model as a function of radar polarimetric variables. The plots show the distribution of residuals (measured - predicted rainfall, in  $\text{mm}\cdot\text{h}^{-1}$ ) with respect to (a) horizontal reflectivity, (b) specific differential phase, and (c) differential reflectivity. The red dashed line represents the zero-error reference.

### 3.4. Heavy Rainfall Detection

**Figure 10** presents the Receiver Operating Characteristic (ROC) curve, which evaluated the model's ability to detect intense rainfall episodes ( $\geq 20 \text{ mm}\cdot\text{h}^{-1}$ ). The y-axis (TPR, or True Positive Rate) indicates the proportion of correct detections among actual positive events, while the x-axis (FPR, or False Positive Rate) measures the proportion of false alarms relative to negative cases. The black dashed diagonal represents a random classifier, serving as a minimal performance baseline. The results highlighted a curve situated distinctly above the random diagonal, with an Area Under the Curve (AUC) of 0.96. Such a value reflects a very strong discriminatory capacity: the model has a 96% probability of assigning a higher score to a truly intense event than to a non-intense one. The steep initial slope of the curve indicates that a TPR greater than 0.9 can be achieved for an FPR lower than 0.1, reflecting an optimal trade-off between detection and false alarm limitation.



**Figure 10.** Receiver Operating Characteristic (ROC) curve for rainfall detection above  $20 \text{ mm}\cdot\text{h}^{-1}$ .

This observation aligns with the use of ROC/AUC for evaluating classification models (or extreme event discrimination) in meteorology and hydrology. For instance, [37] noted that the AUC ranges from 0 (model worse than chance) to 1 (perfect model), and that a high AUC indicates a good ability to separate positive and negative classes, particularly in the presence of class imbalance; they applied this to radar data to improve precipitation forecast performance.

Furthermore, in climatological or hydrological studies, the use of the ROC curve is often recommended to address the issue of class imbalance (such as in our dataset). For example, [57] reported that ROC is a more robust evaluation

method that is not (or less) affected by sample imbalance when predicting extreme events (e.g., for seasonal extreme precipitation prediction).

Thus, obtaining an AUC of 0.96, coupled with a favorable slope at low FPR, indicates that our MLP model is particularly effective at discriminating intense rainfall episodes, offering an excellent trade-off between sensitivity and specificity. This performance is particularly noteworthy given that in many QPE or nowcasting studies, models achieve good results for light to moderate rainfall but often struggle to maintain strong discriminatory capability for extreme events, making this result highly encouraging.

Furthermore, the joint evaluation of residuals (**Figure 9**) and the ROC curve highlighted the complementarity between quantitative accuracy and event detection capability. The residual analysis, performed as a function of polarimetric radar variables ( $Z_h$ ,  $Z_{dr}$ ,  $K_{dp}$ ), showed that the model exhibits no significant global bias, although a tendency toward underestimation appears for high  $Z_h$  and  $K_{dp}$  values, which are typical of intense convective episodes. In parallel, the ROC curve for intense rainfall detection ( $R \geq 20 \text{ mm} \cdot \text{h}^{-1}$ ) displayed an Area Under the Curve (AUC) of 0.96, reflecting an exceptional discriminatory capability in capturing the signatures associated with intense precipitation episodes. This performance indicates that, despite room for improvement in refining the quantitative prediction of extreme intensities, the detection of their occurrence remains highly reliable. This dual analysis confirms the model's suitability for operational applications by combining robustness in distinguishing extreme events with satisfactory accuracy across the entire dataset.

### 3.5. Analysis of Extreme Event Underestimation

Although the MLP model demonstrates superior overall performance, a persistent underestimation of extreme rainfall events (intensities exceeding  $40 \text{ mm} \cdot \text{h}^{-1}$ ) is observed. This deviation can be attributed to both statistical and physical factors. Statistically, the dataset imbalance, common in tropical climates like West Africa, results in a lower frequency of high-intensity events during the training phase, which may constrain the model's ability to fully capture the peak of the distribution. Physically, these extreme intensities are often associated with larger raindrops where the transition from Rayleigh to Mie scattering occurs, introducing non-linearities that are challenging to encapsulate perfectly. Furthermore, despite the polarimetric corrections applied, residuals from attenuation correction in high-reflectivity convective cells and the variability of the Drop Size Distribution (DSD) during intense squall lines can introduce biases. These physical limitations, combined with the saturation effects in some radar variables, contribute to the observed spread at the higher end of the rainfall spectrum.

### 3.6. Comparison with Other Rainfall Estimation Algorithms

The Deep Learning MLP model developed in this study was compared to rainfall estimation algorithms implemented for our study area.

**Table 1** compares the performance of various precipitation estimation models based on polarimetric radar data, including: a Deep Learning MLP model; a  $Z_h - R$  estimator obtained via non-linear least squares optimization accounting for microphysical processes [52],  $Z_h - R$  estimators derived from radar-gauge adjustments; and estimators derived from disdrometric data simulations using the T-Matrix method with the Andsager microphysical model [58] for drop shape assumptions (DSD).

The results indicated that the MLP achieved the best performance across all indicators, with an RMSE of  $7.04 \text{ mm}\cdot\text{h}^{-1}$ , an NRMSE of 0.53, and a MAE of  $2.85 \text{ mm}\cdot\text{h}^{-1}$ , values significantly lower than those of the other models. The near-zero bias (0.88%) demonstrates good neutrality in the estimates, while the coefficient of determination ( $R^2 = 0.72$ ) confirms the model's strong explanatory power.

Approaches based on non-linear least squares optimization [52] or on radar-gauge regressions [20] [21] [59] exhibited higher errors and marked negative

**Table 1.** Performance comparison between the MLP deep learning model and classical parametric radar-rainfall estimators.

Model	Coefficients		RMSE (mm/h)	NRMSE	MAE (mm/h)	NBias (%)	$r$	$R^2$
	$a$	$b, c, d$						
MLP Deep Learning	N/A		<b>7.04</b>	<b>0.53</b>	<b>2.85</b>	<b>0.88</b>	<b>0.85</b>	<b>0.72</b>
Optimization								
$R = aZ_h^b$ [52]	253	$b: 1.66$	7.95	0.60	2.92	-22.34	0.81	0.64
Regression (Radar-raingauges observations)								
$R = aZ_h^b$ [20]	439	$b: 1.38$	9.06	0.68	3.17	-11.19	0.80	0.53
$R = aZ_h^b$ [21]	655	$b: 1.29$	9.34	0.70	3.56	-19.77	0.79	0.50
$R = aZ_h^b$ [59]	460	$b: 1.36$	9.32	0.70	3.24	-10.40	0.80	0.51
$R = aK_{dp}^b$ [20]	12.38	$b: 0.83$	8.56	0.65	4.04	-26.82	0.80	0.58
$R = aK_{dp}^b$ [21]	21.03	$b: 0.57$	8.90	0.67	5.51	36.32	0.77	0.55
$R = aK_{dp}^b Z_{dr}^c$ [21]	20.09	$b: 0.59, c: -0.05$	9.92	0.64	6.01	12.21	0.78	0.59
$R = aZ_h^b Z_{dr}^c K_{dp}^d$ [21]	3.16	$b: 0.18, c: -0.05, d: 0.38$	8.83	0.57	4.67	-3.39	0.85	0.67
T-Matrix Simulation (DSD Andsager Model)								
$R = aZ_h^b$ [5]	424	$b: 1.40$	8.79	0.66	3.11	-12.67	0.80	0.56
$R = aK_{dp}^b$ [5]	12.28	$b: 0.825$	8.60	0.65	4.07	-27.37	0.80	0.58
$R = aK_{dp}^b$ [21]	13.64	$b: 0.83$	8.25	0.62	4.00	-19.38	0.80	0.61
$R = aK_{dp}^b Z_{dr}^c$ [5]	15.47	$b: 0.97, c: -0.47$	9.04	0.68	4.43	-22.33	0.74	0.54
$R = aK_{dp}^b Z_{dr}^c$ [21]	15.13	$b: 0.94, c: -0.29$	8.79	0.66	5.29	31.56	0.77	0.56
$R = aZ_h^b Z_{dr}^c K_{dp}^d$ [21]	9.42	$b: 0.05, c: -0.34, d: 0.89$	7.72	0.58	3.88	9.31	0.83	0.66

biases, revealing a tendency to underestimate high intensities. Models derived from T-Matrix microphysical simulations [5] [21], although performing better than those based on radar-gauge regressions, remained inferior to the MLP model in terms of both accuracy and correlation.

These results confirmed the superiority of the MLP for rainfall estimation using polarimetric radar data, particularly due to its capability to model complex non-linear relationships and reduce the systematic biases observed in traditional parametric approaches.

### 3.7. Comparative Analysis with Optimized Baseline Models

To provide a more exhaustive validation of the proposed Multilayer Perceptron (MLP), its performance was analyzed against locally optimized versions of the most commonly used polarimetric relations. As shown in **Table 2**, each traditional model was reduced to its fundamental power-law form and specifically calibrated for our study area. More precisely, the single-parameter models  $R(Z_h)$  and  $R(K_{dp})$  were optimized to determine their respective coefficients  $a'$  and  $b'$ . For more complex relations, such as  $R(K_{dp}, Z_{dr})$  and the full polarimetric relation  $R(Z_h, Z_{dr}, K_{dp})$ , the optimization process identified the optimal sets of three ( $a', b', c'$ ) and four ( $a', b', c', d'$ ) coefficients, respectively.

Presenting the results in this manner allows us to observe the evolution of error metrics as more physical variables are integrated into the equations. The results presented in **Table 2** indicate an improvement in rainfall estimation, as evidenced by the progression of the performance indicators (RMSE, NRMSE, MAE, NBias,  $r$ , and  $R^2$ ) compared to parametric models that were not optimized for our training dataset. Thus, while local optimization significantly enhances the accuracy of traditional power laws compared to their original versions found in the literature, the MLP consistently yields the best scores. This demonstrates that the neural network does more than simply “fit” the local data better than a standard regression; it captures higher-order dependencies

### 3.8. Diebold-Mariano Test

**Table 3** compares the performance of the Deep Learning MLP model to that of

**Table 2.** Performance metrics of optimized parametric models compared to the proposed MLP model for rainfall estimation.

Model	Coefficients		RMSE (mm/h)	NRMSE	MAE (mm/h)	NBias (%)	$r$	$R^2$
	$a'$	$b'(c', d')$						
MLP Deep Learning		N/A	<b>7.04</b>	<b>0.53</b>	<b>2.85</b>	<b>0.88</b>	<b>0.85</b>	<b>0.72</b>
$R = a'Z_h^{b'}$	0.1577	$b' : 0.4752$	7.77	0.59	3.43	4.57	0.82	0.66
$R = a'K_{dp}^{b'}$	15.8732	$b' : 0.8228$	8.00	0.60	4.01	-6.09	0.80	0.64
$R = a'K_{dp}^{b'}Z_{dr}^{c'}$	13.6624	$(b' : 0.7943, c' : 0.3685)$	8.23	0.62	4.12	-3.38	0.79	0.61
$R = a'Z_h^{b'}Z_{dr}^{c'}K_{dp}^{d'}$	1.1521	$(b' : 0.2959, c' : -0.0707, d' : 0.3014)$	7.20	0.54	3.59	17.86	<b>0.85</b>	0.70

**Table 3.** Results of the Diebold-Mariano (DM) test comparing the MLP deep learning model with classical radar-rainfall estimators.

Compared model	DM	<i>p</i> -value
$Z = 424R^{1.40}$ Simulation [5]	-3.304	$1.0 \times 10^{-3}$
$Z = 439R^{1.38}$ Observation [20]	-3.451	$6.0 \times 10^{-4}$
$Z = 655R^{1.29}$ Observation [21]	-3.726	$2.0 \times 10^{-4}$
$Z = 460R^{1.36}$ Observation [59]	-3.573	$2.0 \times 10^{-4}$
$Z = 253R^{1.66}$ Optimization [52]	-2.858	$4.3 \times 10^{-3}$
$R = 12.28K_{dp}^{0.825}$ Simulation [5]	-4.978	$p \leq 10^{-4}$
$R = 12.38K_{dp}^{0.83}$ Observation [20]	-4.930	$p \leq 10^{-4}$
$R = 21.03K_{dp}^{0.57}$ Observation [21]	-7.908	$p \leq 10^{-4}$
$R = 13.64K_{dp}^{0.83}$ Simulation [21]	-4.469	$p \leq 10^{-4}$
$R = 20.09K_{dp}^{0.59}Z_{dr}^{-0.05}$ Observation [21]	-7.102	$p \leq 10^{-4}$
$R = 3.16Z_h^{0.18}K_{dp}^{0.38}Z_{dr}^{-0.05}$ Observation [21]	-3.520	$4.0 \times 10^{-4}$
$R = 9.42Z_h^{0.05}K_{dp}^{0.89}Z_{dr}^{-0.34}$ Simulation [21]	-4.339	$p \leq 10^{-4}$
$R = 15.47K_{dp}^{0.97}Z_{dr}^{-0.47}$ Simulation [5]	-5.822	$p \leq 10^{-4}$
$R = 15.13K_{dp}^{0.94}Z_{dr}^{-0.29}$ Simulation [21]	-4.916	$p \leq 10^{-4}$

various parametric estimators commonly used for precipitation estimation based on polarimetric radar variables. The Diebold-Mariano (DM) test evaluated whether the performance difference between two methods is statistically significant.

The Diebold-Mariano test results, presented in **Table 3**, showed that the Deep Learning MLP model systematically outperforms all tested parametric precipitation estimation algorithms. Indeed, all DM values were negative and highly significant, with *p*-values less than 0.005 and the majority less than or equal to  $10^{-4}$ , corresponding to a confidence level exceeding 99.99%. Improvements were particularly pronounced compared to relations derived from radar-gauge observations (  $DM = -7.908$  ) and simulated microphysical formulas, confirming the MLP model’s capability to reduce errors even when reference models rely on detailed physical assumptions.

This generalized superiority, combined with the high significance of the *p*-values, demonstrated that the MLP’s performance gain is not due to chance but to better exploitation of non-linear relationships between polarimetric radar variables and precipitation. Even the optimized model by [52], although designed to minimize errors, remained significantly less performant (  $DM = -2.858$ ,  $p = 0.0043$  ). These results also confirm that the MLP’s advantage is not merely a

marginal improvement in statistical indicators (RMSE, MAE,  $R^2$ , NBias%), but is also validated by a formal model comparison test, ensuring that the observed performance gains are highly statistically significant.

These results are consistent with the works of [41] and [60] who established that a significantly negative DM statistic confirms that the first model (here, the MLP) offers predictive accuracy superior to that of the second, beyond random sampling fluctuations.

The results of the Diebold-Mariano (DM) test applied to the locally optimized parametric models (Table 4) further consolidate the generalized superiority of the Deep Learning approach. While local optimization, as discussed in the previous section, significantly reduces systematic bias and improves the statistical indicators of traditional power laws compared to their literature-derived versions, it remains statistically insufficient to match the predictive accuracy of the MLP.

**Table 4.** Results of the Diebold-Mariano (DM) test comparing the MLP deep learning model with parametric models were locally optimized using non-linear least-squares regression on the training dataset.

Compared model	Coefficients		DM	p-value
	$a'$	$b'(c', d')$		
$R = a'Z_h^{b'}$	0.1577	$b' : 0.4752$	-2.624	0.0087
$R = a'K_{dp}^{b'}$	15.8732	$b' : 0.8228$	-3.903	0.0001
$R = a'K_{dp}^{b'}Z_{dr}^{c'}$	13.6624	$(b' : 0.7943, c' : 0.3685)$	-4.400	$p \leq 10^{-4}$
$R = a'Z_h^{b'}Z_{dr}^{c'}K_{dp}^{d'}$	1.1521	$(b' : 0.2959, c' : -0.0707, d' : 0.3014)$	-1.076	0.2818

The DM statistics in Table 4 reveal several key points:

- Persistence of MLP superiority: For the standard  $R(Z_h)$  and  $R(K_{dp})$  relations, the DM values remain negative and highly significant (DM = -2.624,  $p = 0.0087$  and DM = -3.903,  $p = 0.0001$ , respectively). This confirms that even when these models are tuned with the optimal coefficients for the specific study area ( $a' = 0.1577$ ,  $b' = 0.4752$  for  $Z_h$ ), the MLP still provides a statistically significant gain in precision;
- Impact of multi-parameter optimization: The most notable contrast appears with the optimized  $R(K_{dp}, Z_{dr})$  relation. Despite being calibrated to minimize errors on the training dataset, it exhibits a highly significant negative DM value (DM = -4.400,  $p \leq 10^{-4}$ ), indicating that the MLP's ability to model non-linear interactions between  $K_{dp}$  and  $Z_{dr}$  is fundamentally superior to a fixed power-law structure;
- Convergence with the full polarimetric Relation: Interestingly, the comparison between the MLP and the fully optimized  $R(Z_h, Z_{dr}, K_{dp})$  relation shows a DM value of -1.076 with a p-value of 0.2818. While the MLP still maintains a lower error rate (as shown by the negative DM), the difference is no longer

statistically significant at the 5% threshold. This suggests that when a parametric model integrates the full set of polarimetric variables and is perfectly optimized locally, it begins to approach the performance of the neural network, although it still cannot surpass it.

This transition from a highly significant superiority (against single-parameter models) to a more marginal advantage (against the optimized full  $Z_h / Z_{dr} / K_{dp}$  model) confirms the analysis in **Table 2**. The MLP effectively acts as a “universal optimizer” that captures the physical information contained within the polarimetric variables more efficiently than traditional regressions. These results validate that the observed performance gains are not merely due to local calibration, but to the inherent capacity of Deep Learning to encapsulate the complex microphysics of tropical precipitation.

## 4. Conclusions

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your journal for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper.

This study demonstrates the relevance and effectiveness of a Deep Learning MLP model for the quantitative estimation of tropical precipitation using polarimetric radar data from the AMMA campaign. The developed architecture, jointly exploiting the variables  $Z_h$ ,  $Z_{dr}$  and  $K_{dp}$ , successfully captured the complex non-linear relationships between radar observables and ground-measured rainfall rates, systematically outperforming traditional parametric approaches, whether derived from observational regressions, optimizations, or microphysical simulations.

The performance achieved, characterized by minimal errors (RMSE, MAE, NRMSE), a near-zero global bias, and a high coefficient of determination ( $R^2 = 0.72$ ), confirms the model’s robustness and accuracy for the majority of rainy situations, while revealing room for improvement regarding extreme episodes. The excellent discriminatory capability for heavy rainfall detection (AUC = 0.96) and the statistical significance of the observed gains, validated by the Diebold-Mariano test, reinforce the reliability of the results and their operational scope.

By combining quantitative accuracy with event detection capability, the developed MLP positions itself as a high-performance tool for hydrometeorological prediction and risk management in tropical zones. Future research perspectives include enriching the dataset with additional observations, implementing learning strategies adapted to extreme events, and exploring advanced neural architectures (e.g., convolutional and recurrent networks) to further improve generalization capability and the representation of complex microphysical processes.

## Acknowledgements

This research was conducted under the auspices of AMMA. Based on the French

initiative, AMMA was built by an international scientific group. A large number of agencies, especially from France, the United Kingdom, the USA and Africa, currently fund it. It has been the beneficiary of a major financial contribution from the European Community's Sixth Framework Research Program. Detailed information on scientific coordination and funding is available on the AMMA international website: <http://www.amma-international.org>.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- [1] Marshall, J.S. and Palmer, W.M.K. (1948) The Distribution of Raindrops with Size. *Journal of Meteorology*, **5**, 165-166. [https://doi.org/10.1175/1520-0469\(1948\)005<0165:tdorws>2.0.co;2](https://doi.org/10.1175/1520-0469(1948)005<0165:tdorws>2.0.co;2)
- [2] Battan, L.J. (1973) Radar Observation of the Atmosphere. University of Chicago Press.
- [3] Yuter, S.E. and Houze, R.A. (1997) Measurements of Raindrop Size Distributions over the Pacific Warm Pool and Implications for  $Z$ - $R$  Relations. *Journal of Applied Meteorology*, **36**, 847-867. [https://doi.org/10.1175/1520-0450\(1997\)036<0847:morsdo>2.0.co;2](https://doi.org/10.1175/1520-0450(1997)036<0847:morsdo>2.0.co;2)
- [4] Atlas, D., Ulbrich, C.W., Marks, F.D., Amitai, E. and Williams, C.R. (1999) Systematic Variation of Drop Size and Radar-Rainfall Relations. *Journal of Geophysical Research: Atmospheres*, **104**, 6155-6169. <https://doi.org/10.1029/1998jd200098>
- [5] Maki, M., Keenan, T.D., Sasaki, Y. and Nakamura, K. (2001) Characteristics of the Raindrop Size Distribution in Tropical Continental Squall Lines Observed in Darwin, Australia. *Journal of Applied Meteorology*, **40**, 1393-1412. [https://doi.org/10.1175/1520-0450\(2001\)040<1393:cotrsd>2.0.co;2](https://doi.org/10.1175/1520-0450(2001)040<1393:cotrsd>2.0.co;2)
- [6] Rao, T.N., Rao, D.N., Mohan, K. and Raghavan, S. (2001) Classification of Tropical Precipitating Systems and Associated  $Z$ - $R$  Relationships. *Journal of Geophysical Research: Atmospheres*, **106**, 17699-17711. <https://doi.org/10.1029/2000jd900836>
- [7] Uijlenhoet, R. (2001) Raindrop Size Distributions and Radar Reflectivity-Rain Rate Relationships for Radar Hydrology. *Hydrology and Earth System Sciences*, **5**, 615-628. <https://doi.org/10.5194/hess-5-615-2001>
- [8] Lee, G. and Zawadzki, I. (2006) Radar Calibration by Gage, Disdrometer, and Polarimetry: Theoretical Limit Caused by the Variability of Drop Size Distribution and Application to Fast Scanning Operational Radar Data. *Journal of Hydrology*, **328**, 83-97. <https://doi.org/10.1016/j.jhydrol.2005.11.046>
- [9] Bournas, A. and Baltas, E. (2022) Determination of the  $Z$ - $R$  Relationship through Spatial Analysis of X-Band Weather Radar and Rain Gauge Data. *Hydrology*, **9**, Article 137. <https://doi.org/10.3390/hydrology9080137>
- [10] Sauvageot, H. and Lacaux, J. (1995) The Shape of Averaged Drop Size Distributions. *Journal of the Atmospheric Sciences*, **52**, 1070-1083. [https://doi.org/10.1175/1520-0469\(1995\)052<1070:tsoads>2.0.co;2](https://doi.org/10.1175/1520-0469(1995)052<1070:tsoads>2.0.co;2)
- [11] Nzeukou, A., Sauvageot, H., Ochou, A.D. and Kebe, C.M.F. (2004) Raindrop Size Distribution and Radar Parameters at Cape Verde. *Journal of Applied Meteorology*, **43**, 90-105. [https://doi.org/10.1175/1520-0450\(2004\)043<0090:rsdarp>2.0.co;2](https://doi.org/10.1175/1520-0450(2004)043<0090:rsdarp>2.0.co;2)

- [12] Moumouni, S., Gosset, M. and Houngninou, E. (2008) Main Features of Rain Drop Size Distributions Observed in Benin, West Africa, with Optical Disdrometers. *Geophysical Research Letters*, **35**, L23807. <https://doi.org/10.1029/2008gl035755>
- [13] Russell, B., Williams, E.R., Gosset, M., Cazenave, F., Descroix, L., Guy, N., *et al.* (2010) Radar/Rain-Gauge Comparisons on Squall Lines in Niamey, Niger for the Amma. *Quarterly Journal of the Royal Meteorological Society*, **136**, 289-303. <https://doi.org/10.1002/qj.548>
- [14] Bringi, V.N., Tang, T. and Chandrasekar, V. (2004) Evaluation of a New Polarimetrically Based  $Z$ - $R$  Relation. *Journal of Atmospheric and Oceanic Technology*, **21**, 612-623. [https://doi.org/10.1175/1520-0426\(2004\)021<0612:eoanpb>2.0.co;2](https://doi.org/10.1175/1520-0426(2004)021<0612:eoanpb>2.0.co;2)
- [15] Anagnostou, M.N., Kalogiros, J., Anagnostou, E.N., Tarolli, M., Papadopoulos, A. and Borga, M. (2010) Performance Evaluation of High-Resolution Rainfall Estimation by X-Band Dual-Polarization Radar for Flash Flood Applications in Mountainous Basins. *Journal of Hydrology*, **394**, 4-16. <https://doi.org/10.1016/j.jhydrol.2010.06.026>
- [16] You, C.H., Kang, M.Y., Hwang, Y., Yee, J.J., Jang, M. and Lee, D.I. (2018) A Statistical Approach to Radar Rainfall Estimates Using Polarimetric Variables. *Atmospheric Research*, **209**, 65-75. <https://doi.org/10.1016/j.atmosres.2018.03.016>
- [17] Bringi, V.N. and Chandrasekar, V. (2001) Polarimetric Doppler Weather Radar: Principles and Applications. Cambridge University Press. <https://doi.org/10.1017/cbo9780511541094>
- [18] Seliga, T.A. and Bringi, V.N. (1976) Potential Use of Radar Differential Reflectivity Measurements at Orthogonal Polarizations for Measuring Precipitation. *Journal of Applied Meteorology*, **15**, 69-76. [https://doi.org/10.1175/1520-0450\(1976\)015<0069:puordr>2.0.co;2](https://doi.org/10.1175/1520-0450(1976)015<0069:puordr>2.0.co;2)
- [19] Sachidananda, M. and Zrnić, D.S. (1987) Rain Rate Estimates from Differential Polarization Measurements. *Journal of Atmospheric and Oceanic Technology*, **4**, 588-598. [https://doi.org/10.1175/1520-0426\(1987\)004<0588:rrefdp>2.0.co;2](https://doi.org/10.1175/1520-0426(1987)004<0588:rrefdp>2.0.co;2)
- [20] Gosset, M., Zahiri, E. and Moumouni, S. (2010) Rain Drop Size Distribution Variability and Impact on X-Band Polarimetric Radar Retrieval: Results from the AMMA Campaign in Benin. *Quarterly Journal of the Royal Meteorological Society*, **136**, 243-256. <https://doi.org/10.1002/qj.556>
- [21] Koffi, A.K., Gosset, M., Zahiri, E., Ochou, A.D., Kacou, M., Cazenave, F., *et al.* (2014) Evaluation of X-Band Polarimetric Radar Estimation of Rainfall and Rain Drop Size Distribution Parameters in West Africa. *Atmospheric Research*, **143**, 438-461. <https://doi.org/10.1016/j.atmosres.2014.03.009>
- [22] Ryzhkov, A.V., Giangrande, S.E. and Schuur, T.J. (2005) Rainfall Estimation with a Polarimetric Prototype of WSR-88D. *Journal of Applied Meteorology*, **44**, 502-515. <https://doi.org/10.1175/jam2213.1>
- [23] Giangrande, S.E. and Ryzhkov, A.V. (2008) Estimation of Rainfall Based on the Results of Polarimetric Echo Classification. *Journal of Applied Meteorology and Climatology*, **47**, 2445-2462. <https://doi.org/10.1175/2008jamc1753.1>
- [24] Uijlenhoet, R., Steiner, M. and Smith, J.A. (2003) Variability of Raindrop Size Distributions in a Squall Line and Implications for Radar Rainfall Estimation. *Journal of Hydrometeorology*, **4**, 43-61. [https://doi.org/10.1175/1525-7541\(2003\)004<0043:vorsdi>2.0.co;2](https://doi.org/10.1175/1525-7541(2003)004<0043:vorsdi>2.0.co;2)
- [25] Alfieri, L., Claps, P. and Laio, F. (2010) Time-Dependent  $Z$ - $R$  Relationships for Estimating Rainfall Fields from Radar Measurements. *Natural Hazards and Earth System*

- Sciences*, **10**, 149-158. <https://doi.org/10.5194/nhess-10-149-2010>
- [26] Munchak, S.J., Kummerow, C.D. and Elsaesser, G. (2012) Relationships between the Raindrop Size Distribution and Properties of the Environment and Clouds Inferred from TRMM. *Journal of Climate*, **25**, 2963-2978. <https://doi.org/10.1175/jcli-d-11-00274.1>
- [27] Simpson, M.J. and Fox, N.I. (2018) Dual-polarized Quantitative Precipitation Estimation as a Function of Range. *Hydrology and Earth System Sciences*, **22**, 3375-3389. <https://doi.org/10.5194/hess-22-3375-2018>
- [28] Chen, H., Chandrasekar, V., Cifelli, R. and Xie, P. (2020) A Machine Learning System for Precipitation Estimation Using Satellite and Ground Radar Network Observations. *IEEE Transactions on Geoscience and Remote Sensing*, **58**, 982-994. <https://doi.org/10.1109/tgrs.2019.2942280>
- [29] Barrera-Animas, A.Y., Oyedele, L.O., Bilal, M., Akinosho, T.D., Delgado, J.M.D. and Akanbi, L.A. (2022) Rainfall Prediction: A Comparative Analysis of Modern Machine Learning Algorithms for Time-Series Forecasting. *Machine Learning with Applications*, **7**, Article ID: 100204. <https://doi.org/10.1016/j.mlwa.2021.100204>
- [30] Huangfu, J., Hu, Z., Zheng, J., Wang, L. and Zhu, Y. (2024) Study on Quantitative Precipitation Estimation by Polarimetric Radar Using Deep Learning. *Advances in Atmospheric Sciences*, **41**, 1147-1160. <https://doi.org/10.1007/s00376-023-3039-0>
- [31] Li, W., Chen, H. and Han, L. (2024) Improving Explainability of Deep Learning for Polarimetric Radar Rainfall Estimation. *Geophysical Research Letters*, **51**, e2023GL107898. <https://doi.org/10.1029/2023gl107898>
- [32] Akponi, F.P., Moumouni, S., Zahiri, E., Kacou, M. and Gosset, M. (2025) Using an Artificial Neural Network to Assess Several Rainfall Estimation Algorithms Based on X-Band Polarimetric Variables in West Africa. *Atmosphere*, **16**, Article 371. <https://doi.org/10.3390/atmos16040371>
- [33] Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G. and Yacalis, G. (2018) Could Machine Learning Break the Convection Parameterization Deadlock? *Geophysical Research Letters*, **45**, 5742-5751. <https://doi.org/10.1029/2018gl078202>
- [34] Depraetere, C., Gosset, M., Ploix, S. and Laurent, H. (2009) The Organization and Kinematics of Tropical Rainfall Systems Ground Tracked at Mesoscale with Gages: First Results from the Campaigns 1999-2006 on the Upper Ouémé Valley (Benin). *Journal of Hydrology*, **375**, 143-160. <https://doi.org/10.1016/j.jhydrol.2009.01.011>
- [35] Carey, L.D., Rutledge, S.A., Ahijevych, D.A. and Keenan, T.D. (2000) Correcting Propagation Effects in C-Band Polarimetric Radar Observations of Tropical Convection Using Differential Propagation Phase. *Journal of Applied Meteorology*, **39**, 1405-1433. [https://doi.org/10.1175/1520-0450\(2000\)039<1405:cpeicb>2.0.co;2](https://doi.org/10.1175/1520-0450(2000)039<1405:cpeicb>2.0.co;2)
- [36] Milligan, G.W. and Cooper, M.C. (1988) A Study of Standardization of Variables in Cluster Analysis. *Journal of Classification*, **5**, 181-204. <https://doi.org/10.1007/bf01897163>
- [37] Mao, Y. and Sorteberg, A. (2020) Improving Radar-Based Precipitation Nowcasts with Machine Learning Using an Approach Based on Random Forest. *Weather and Forecasting*, **35**, 2461-2478. <https://doi.org/10.1175/waf-d-20-0080.1>
- [38] Bradley, A.P. (1997) The Use of the Area under the ROC Curve in the Evaluation of Machine Learning Algorithms. *Pattern Recognition*, **30**, 1145-1159. [https://doi.org/10.1016/s0031-3203\(96\)00142-2](https://doi.org/10.1016/s0031-3203(96)00142-2)
- [39] Fawcett, T. (2006) An Introduction to ROC Analysis. *Pattern Recognition Letters*, **27**, 861-874. <https://doi.org/10.1016/j.patrec.2005.10.010>

- [40] Hanley, J.A. and McNeil, B.J. (1982) The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology*, **143**, 29-36. <https://doi.org/10.1148/radiology.143.1.7063747>
- [41] Diebold, F.X. and Mariano, R.S. (1995) Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, **13**, 253-263. <https://doi.org/10.1080/07350015.1995.10524599>
- [42] Shabbir, M., Chand, S. and Iqbal, F. (2024) A Novel Hybrid Framework to Model the Relationship of Daily River Discharge with Meteorological Variables. *Meteorology Hydrology and Water Management*, **11**, 70-94. <https://doi.org/10.26491/mhwm/187899>
- [43] Bell, P., Catto, J., Jones, A. and Siegert, S. (2025) Developing and Evaluating a Bayesian Weather Generator for UK Precipitation Conditioned on Discrete Storm Types. <https://doi.org/10.5194/egusphere-2025-4110-supplement>
- [44] Pavan Kumar, S.T., Chotikarn, P., de Oliveira Júnior, J.F. and Lalremsang, P. (2025) Bayesian Change Point Estimation and Evaluation of Machine Learning Models for Multi-Scale SPI Prediction for Assam-Meghalaya Region in North-Eastern India. *International Journal of Climatology*, **45**, e70061. <https://doi.org/10.1002/joc.70061>
- [45] Gorgucci, E., Scarchilli, G., Chandrasekar, V. and Bringi, V.N. (2001) Rainfall Estimation from Polarimetric Radar Measurements: Composite Algorithms Immune to Variability in Raindrop Shape-Size Relation. *Journal of Atmospheric and Oceanic Technology*, **18**, 1773-1786. [https://doi.org/10.1175/1520-0426\(2001\)018<1773:refprm>2.0.co;2](https://doi.org/10.1175/1520-0426(2001)018<1773:refprm>2.0.co;2)
- [46] Li, W., Chen, H. and Han, L. (2023) Deep Learning for Polarimetric Radar Quantitative Precipitation Estimation: Model Interpretation and Explainability. *IGARSS 2023—2023 IEEE International Geoscience and Remote Sensing Symposium*, Pasadena, 16-21 July 2023, 5194-5197. <https://doi.org/10.1109/igarss52108.2023.10283431>
- [47] Zhang, Y., Bi, S., Liu, L., Chen, H., Zhang, Y., Shen, P., *et al.* (2021) Deep Learning for Polarimetric Radar Quantitative Precipitation Estimation during Landfalling Typhoons in South China. *Remote Sensing*, **13**, Article 3157. <https://doi.org/10.3390/rs13163157>
- [48] Chen, H., Chandrasekar, V. and Cifelli, R. (2019) A Deep Learning Approach to Dual-Polarization Radar Rainfall Estimation. 2019 *URSI Asia-Pacific Radio Science Conference (AP-RASC)*, New Delhi, 9-15 March 2019, 1-2. <https://doi.org/10.23919/ursiap-rasc.2019.8738337>
- [49] Rudy, S.H. and Sapsis, T.P. (2023) Output-Weighted and Relative Entropy Loss Functions for Deep Learning Precursors of Extreme Events. *Physica D: Nonlinear Phenomena*, **443**, Article ID: 133570. <https://doi.org/10.1016/j.physd.2022.133570>
- [50] Tian, W., Yi, L., Liu, W., Huang, W., Ma, G. and Zhang, Y. (2020) Ground Radar Precipitation Estimation with Deep Learning Approaches in Meteorological Private Cloud. *Journal of Cloud Computing*, **9**, Article No. 22. <https://doi.org/10.1186/s13677-020-00167-w>
- [51] Rosenhoover, M., Rushing, J., Beck, J., White, K. and Graves, S. (2025) Improving Doppler Radar Precipitation Prediction with Citizen Science Rain Gauges and Deep Learning. *Sensors*, **25**, Article 3719. <https://doi.org/10.3390/s25123719>
- [52] Kouadio, G., Zahiri, E., Kacou, M., Koffi, A.K., Ochou, A.D. and Assamoi, P. (2021) Improving Radar Rainfall Estimation by Accounting for Microphysical Processes Using a Micro Rain Radar in West Africa. *Atmospheric and Climate Sciences*, **11**, 658-688. <https://doi.org/10.4236/acs.2021.114039>

- [53] Dinh, T., Phung, D., Kim, S. and Bae, D. (2023) A New Approach for Quantitative Precipitation Estimation from Radar Reflectivity Using a Gated Recurrent Unit Network. *Journal of Hydrology*, **624**, Article ID: 129887. <https://doi.org/10.1016/j.jhydrol.2023.129887>
- [54] González-Abad, J. and Gutiérrez, J.M. (2025) Are Deep Learning Methods Suitable for Downscaling Global Climate Projections? An Intercomparison for Temperature and Precipitation over Spain. *Artificial Intelligence for the Earth Systems*, **4**, 56-77. <https://doi.org/10.1175/aies-d-24-0121.1>
- [55] Neuper, M. and Ehret, U. (2019) Quantitative Precipitation Estimation with Weather Radar Using a Data- and Information-Based Approach. *Hydrology and Earth System Sciences*, **23**, 3711-3733. <https://doi.org/10.5194/hess-23-3711-2019>
- [56] Verdelho, F.F., Beneti, C., Pavam, L.G., Calvetti, L., Oliveira, L.E.S. and Zanata Alves, M.A. (2024) Quantitative Precipitation Estimation Using Weather Radar Data and Machine Learning Algorithms for the Southern Region of Brazil. *Remote Sensing*, **16**, Article 1971. <https://doi.org/10.3390/rs16111971>
- [57] Wei, W., Yan, Z., Tong, X., Han, Z., Ma, M., Yu, S., *et al.* (2022) Seasonal Prediction of Summer Extreme Precipitation over the Yangtze River Based on Random Forest. *Weather and Climate Extremes*, **37**, Article ID: 100477. <https://doi.org/10.1016/j.wace.2022.100477>
- [58] Andsager, K., Beard, K.V. and Laird, N.F. (1999) Laboratory Measurements of Axis Ratios for Large Raindrops. *Journal of the Atmospheric Sciences*, **56**, 2673-2683. [https://doi.org/10.1175/1520-0469\(1999\)056<2673:lmoarf>2.0.co;2](https://doi.org/10.1175/1520-0469(1999)056<2673:lmoarf>2.0.co;2)
- [59] Kacou, M.H.A. (2014) Analyse des précipitations en zone sahélienne à partir d'un radar bande X polarimétrique. Ph.D. Thesis, Université Toulouse III—Paul Sabatier.
- [60] Harvey, D., Leybourne, S. and Newbold, P. (1997) Testing the Equality of Prediction Mean Squared Errors. *International Journal of Forecasting*, **13**, 281-291. [https://doi.org/10.1016/s0169-2070\(96\)00719-4](https://doi.org/10.1016/s0169-2070(96)00719-4)