Polarimetric Radar Quantitative Precipitation Estimation Using Deep Convolutional Neural Networks

Wenyuan Li, Haonan Chen, *Senior Member, IEEE*, Lei Han, *Member, IEEE*

***Abstract*—Accurate estimation of surface precipitation with high spatial and temporal resolution is critical for decision making regarding severe weather and water resources management. Polarimetric weather radar is the main operational instrument used for quantitative precipitation estimation (QPE). However, conventional parametric radar QPE algorithms such as the radar reflectivity (Z) and rain rate (R) relations can not fully represent clouds and precipitation dynamics due to their dependency on local raindrop size distributions and the inherent parameterization errors. This article develops four deep learning (DL) models for polarimetric radar QPE (i.e., RQPENet$_{\mathrm{D1}}$, RQPENet$_{\mathrm{D2}}$, RQPENet$_{\mathrm{V}}$, RQPENet$_{\mathrm{R}}$) using different core building blocks. In particular, multi-dimensional polarimetric radar observations are utilized as input and surface gauge measurements are used as training labels. The feasibility and performance of these DL models are demonstrated and quantified using U.S. Weather Surveillance Radar - 1988 Doppler (WSR-88D) observations near Melbourne, Florida. The experimental results show that the dense blocks-based models (i.e., RQPENet$_{\mathrm{D1}}$ and RQPENet$_{\mathrm{D2}}$) have better performance than residual blocks, RepVGG blocks-based models (i.e., RQPENet$_{\mathrm{R}}$ and RQPENet$_{\mathrm{V}}$) and five traditional Z-R relations. RQPENet$_{\mathrm{D1}}$ has the best quantitative performance scores, with a mean absolute error (MAE) of 1.58 mm, root mean squared error (RMSE) of 2.68 mm, normalized standard error (NSE) of 26%, and correlation of 0.92 for hourly rainfall estimates using independent rain gauge data as references. These results suggest that deep learning performs well in mapping the connection between polarimetric radar observations aloft and surface rainfall.**

***Index Terms*—Dual-polarization, weather radar, quantitative precipitation estimation, deep learning**

I. INTRODUCTION

Precipitation is a crucial component of the global and regional water cycles and playing an important role in weather-, water-, and climate-related research and applications. In the context of climate changes, the water cycle is accelerated, resulting in more frequent flash floods. In addition, owning to rapid urbanization, densely populated areas are more vulnerable to urban flash floods and other weather extremes [1–3]. Therefore, accurate and timely estimation of precipitation is critical to decision making regarding severe weather and water resources management.

Radars and rain gauges are the two main instruments used for precipitation monitoring and measurements. Since radars can provide spatially continuous observations over large areas with high temporal resolutions [3–5], they are considered the cornerstones of the national severe weather observation and forecast infrastructure in many countries. Rain gauges are often used for evaluation and bias correction of radar rainfall estimates since rain gauges provide direct measurements of rainfall, though at point locations [6–9]. Traditionally, the relationships between radar reflectivity (Z) and rainfall rate (R), also known as Z-R relations, are utilized for radar quantitative precipitation estimation (QPE), and these relations are still used for many operational radars. Since the introduction of radar polarization [10], a large amount of research effort has been devoted to polarimetric radar systems and QPE based on polarimetric observables, including radar reflectivity, differential reflectivity (Z_{dr}), and specific differential phase (K_{dp}) (e.g., [4][5][11][12]).

The polarimetric variables offer more microphysical information about precipitation such as raindrop shape and size distributions, leading to better rainfall estimates than conventional Z-R relations. Nevertheless, both Z-R and polarimetric radar rainfall relations are generally derived based on nonlinear regression (typically power-law relations) between radar observables and rainfall rates. This type of fixed-parameter power-law relations cannot represent all types of rainfall at multiple climate regimes, or even one climate region. Often, the parametric radar rainfall relations need to be adjusted for regional applications. In addition, the inherent parameterization error in these relations is large, especially in Z-R relations ([4][5]). Kirstetter et al. [13] considered the inherent uncertainty in the Z-R relations of different precipitation types based on the Multi-Radar Multi-Sensor (MRMS) QPE system, which improved the accuracy of precipitation estimation to a certain extent. Nevertheless, none of these power-law relations could incorporate the spatial features of precipitation in with terms of radar observations.

In recent years, due to the powerful data fitting and information extraction capabilities of deep learning (DL), it has been successfully applied to many research areas such as image classification, object detection, and image segmentation. For

Wenyuan Li and Lei Han were supported by the National Natural Science Foundation of China under Grant 42275003. The work of Haonan Chen was supported by the U.S. National Science Foundation (NSF) CAREER program. *(Corresponding author: Lei Han.)*

Wenyuan Li is with the Faculty of Information Science and Engineering, Ocean University of China, Qingdao 266100, China. She is currently a visiting scholar in the Department of Electrical and Computer Engineering at Colorado State University, Fort Collins, CO 80523, USA (email: wenyuan.li@colostate.edu).

Haonan Chen is with the Department of Electrical and Computer Engineering at Colorado State University, Fort Collins, CO 80523, USA (email: haonan.chen@colostate.edu).

Lei Han is with the Faculty of Information Science and Engineering, Ocean University of China, Qingdao 266100, China (email: hanlei@ouc.edu.cn).

example, the Convolutional Neural Networks (CNN) developed by the Visual Geometry Group (VGG) at University of Oxford [14] have achieved great success in image classification. In order to take into account the result accuracy and the inference speed of the model, RepVGG [15] uses a re-parameterization technique, which can use a multi-branch structure when the model training, and convert the multi-branch into a single-branch during model testing. In addition, the residual block and dense block in the ResNet [16] and DenseNet [17] networks have addressed the gradient vanishing problem to a large extent when the model layers are too deep: ResNet won the double champion of ImageNet Large Scale Visual Recognition Challenge (ILSVRC) classification and object detection task in 2015, and DenseNet won the best paper of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) in 2017. The three classic models mentioned above have far-reaching influence in the development of deep learning, partly because their core building blocks are highly innovative and practical, and are easy to implement. When applied to the QPE task, the core building blocks show promising performance by capturing more spatial correlated information from dual-polarization radar data and the rain-falling process.

In this article, we uses classical DL core building blocks (i.e., dense blocks, residual blocks, and RepVGG blocks) to address the identified issues in radar QPE. In particular, four DL QPE models are designed to extract precipitation information and quantify surface rainfall rate using 3-D polarimetric radar observations above the ground. The feasibility and performance of these DL models are demonstrated and quantified using the Weather Surveillance Radar - 1988 Doppler (WSR-88D) observations near Melbourne (KMLB), Florida, USA. To our knowledge, this article is among the first studies about using machine learning for quantitative applications of weather radars. Section II describes the study domain and data used in this article. Section III details the architecture of the designed DL models. Section IV presents the experimental results. Discussion and conclusions are provided in Section V.

II. STUDY DOMAIN AND DATASETS

A. *Study Domain*

Our study domain is located in the Florida Peninsula, the third most populous state in the United States. Fig. 1 shows the distribution of the rain gauges and the location of the KMLB dual-polarization radar station represented by the yellow dot (80.65°W, 28.11°N). Here, it should be noted that, due to the degradation of data quality for longer radar observation distance, we only select radar observations and rain gauge data within 150 km from the KMLB radar station as the dataset. Therefore, the red dots and black dots in Fig. 1 represent the distribution of rain gauges within and beyond 150 km from the radar site, respectively. All these rain gauges form the South Florida Water Management District (SFL) rain gauge network.

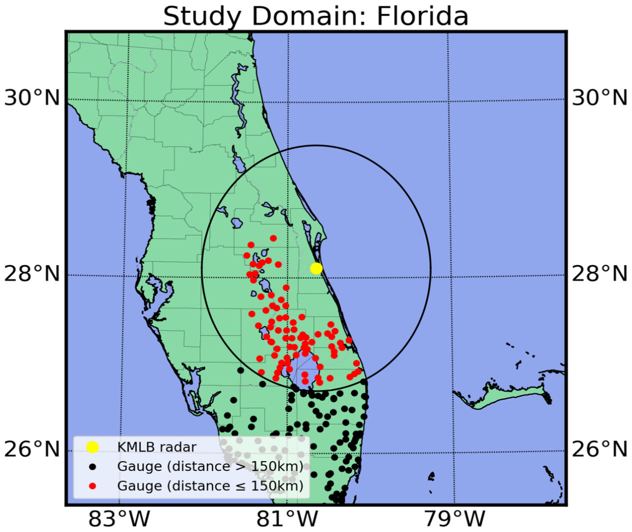


Fig. 1. The study domain (black circle with a radius of 150 km). The yellow dot represents the KMLB WSR-88D radar location (28.11°N, 80.65°W) in Melbourne, Florida. The red dots and black dots indicate distribution of rain gauges within and beyond 150 km from the radar site, respectively.

Florida has a subtropical and tropical rainforest climate, with a peculiar peninsula geographical environment. It has more thunderstorm activity than other parts of U.S. Specifically, the frequency of summer thunderstorms in central Florida, our study domain, equals that of the world's maximum thunderstorm areas, i.e., Lake Victoria in Africa and the Amazon basin[1]. Florida often has minor flash floods that may pose a serious threat to human life. In addition to this, most of the study domain is flat and relatively low in elevation, making it less obscuring and more visible for radar observations. Therefore, this is an ideal study domain for QPE.

B. *Datasets*

In this study, dual-polarization radar data are collected from the KMLB Weather Surveillance Radar – 1988 Doppler (WSR-88D) near Melbourne, Florida, USA. The KMLB radar is an S-band radar that emits electromagnetic waves with a wavelength of 10.0-11.1 cm. The radar has a maximum detection range of 460 kilometers and an azimuth range of 0-360°. Its antenna scanning elevation angle ranges from 0.5° to 19.5°, and can be as low as 0.2° or even lower in coastal areas. It takes about 5-7 minutes to complete a volume scan. The KMLB radar range resolution is 250m, and radar azimuth resolution is 0.5°. The rain gauge data used in this article are collected from the SFL tipping bucket rain gauge network in Florida, which has undergone an extensive data quality. These data are also used as the verification data for the Global Precipitation Measurement (GPM) mission of NASA. The above radar and rain gauge data from 2016-2019 are used in this study. Specifically, we use the data from 2016 to 2018 for training (80% of the data) and validation (20% of the data), and the data from 2019 as independent test data.

In general, ground clutter occurs on WSR-88D radar products when the returns from stationary or nearly stationary ground targets are not filtered out. This clutter may be caused by building and wind farms and can be mistakenly considered

---

[1] https://climatecenter.fsu.edu/topics/thunderstorms

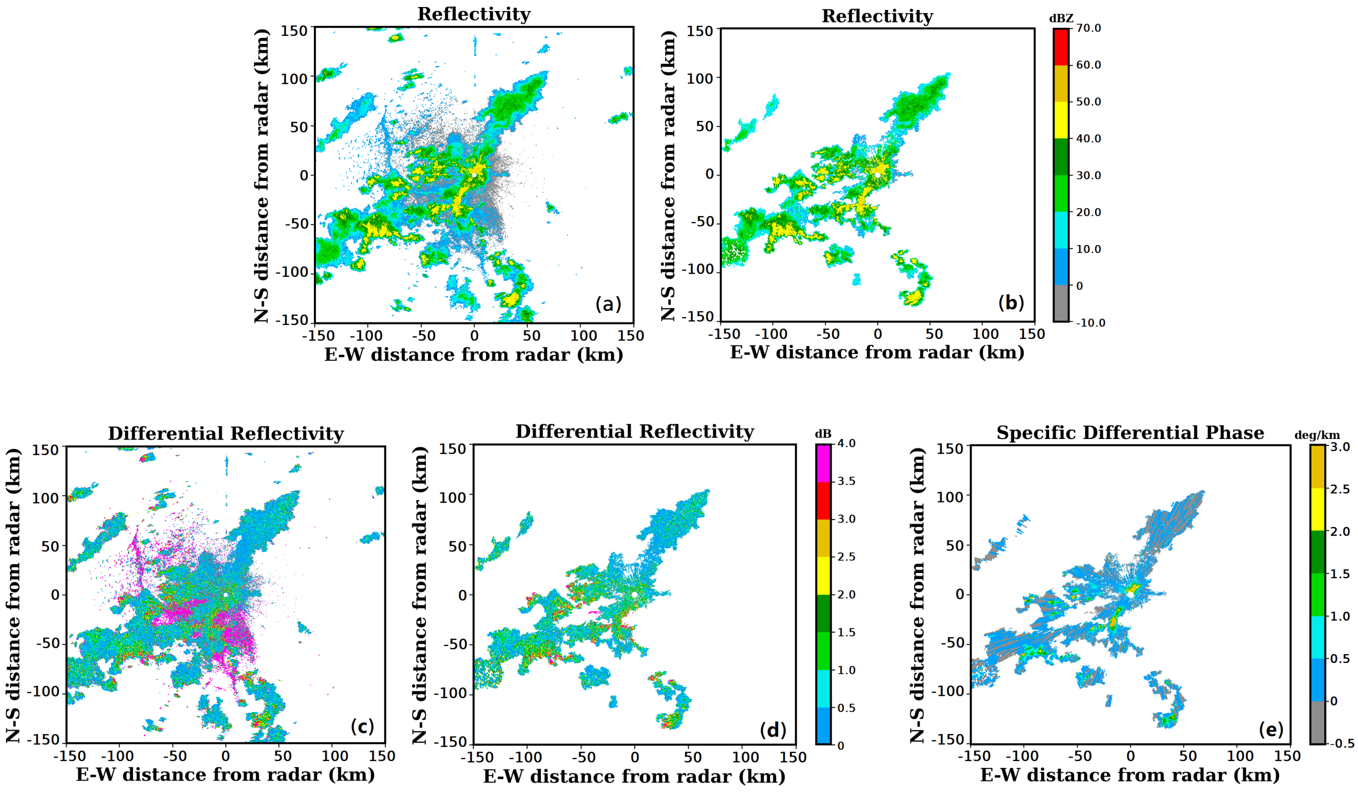


Fig. 2. Example KMLB WSR-88D radar observations at 20:34 UTC, 19 July 2017: (a) radar reflectivity Z before quality control; (b) radar reflectivity Z after quality control; (c) differential reflectivity Z_{dr} before quality control; (d) differential reflectivity Z_{dr} after quality control; (e) specific differential phase K_{dp} estimated using DROPS2.0 [5].

as real precipitation. Therefore, they need to be filtered out before further processing. The DROPS2.0 algorithm is used in our study to filter the non-precipitation echoes and estimate the specific differential phase K_{dp} [5]. Fig. 2 shows the radar reflectivity Z and differential reflectivity Z_{dr} before and after quality control, as well as the estimated specific differential phase K_{dp} at 20:34 UTC on 19 July 2017. It should be mentioned that the rain gauge data with rainfall amount less than 2.5 mm and those with rainfall rate less than 1 mm/h are also excluded.

III. METHODOLOGY

Figure 3 illustrates the pipeline of our DL framework for QPE. It is divided into two independent parts: (a) model training and (b) estimation. In the model training period, the historical 3-D radar and rain gauge data are fed directly into the DL network to train a regressor. The trained regressor is then used to make estimation for precipitation using newly arrived radar data. Next, we will introduce how to establish the training data set followed by the model architecture introduction.

A. Establishing the Training Data Set

Because dual-polarization radar observations have been widely used for quantitative precipitation estimation, here we use three observables Z, Z_{dr} and K_{dp} as model's input, where Z is radar reflectivity, Z_{dr} is the difference between horizontal radar reflectivity and vertical radar reflectivity, K_{dp} is one-half the range derivative of the differential phase shift (ϕ_{dp}) [4]. Meanwhile, since rain gauge observations (i.e. rainfall rate) are the most accurate measurement of precipitation, we use them as the ground truth for model training.

As the temporal resolution of the two kinds of data are different, we need to align radar observations and rain gauge data. The KMLB radar observations with the temporal resolution of 5-7 minutes, whereas rain gauges are sparsely located and produce precipitation observations with the temporal resolution of 1 minute. In this study, the average value of rain gauge data within the interval of two adjacent radar observations is used as the ground truth. Next, we will introduce how to construct one training data sample.

As shown in Fig. 4, the input of the model is a tensor $T_{6\times9\times9}$ containing the 3D radar data of the above 3 variables in a 9×9 pixel window centered in pixel (5, 5). The window size 9×9 is chosen according to the optimal results of our experiments. It should be mentioned that only the lowest and second lowest elevation angles of radar observation data are used in this study. The main reason is that, although the radar beams at high elevation angles can detect rain aloft, due to the ongoing evaporation while falling, some of the rain is evaporated before it reaches the ground, so the rainfall may

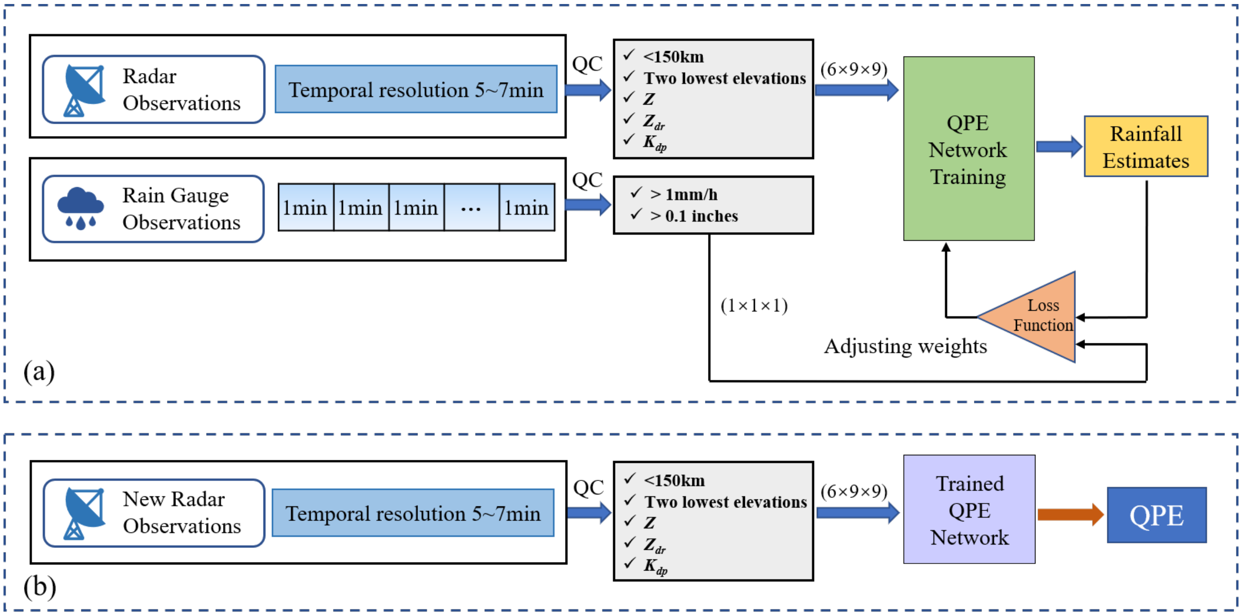


Fig. 3. Flowchart of the deep learning method for polarimetric radar QPE: (a) training stage; (b) application of the trained models. The historical 3-D radar and rain gauge data are fed directly into the DL network to train a regressor during the training period. The trained regressor is then used to estimate QPE based on new radar data. QC stands for data quality control.

be overestimated. Therefore, there are 3 variables and each has 2 elevation angles, this is how the number 6 (3 × 2) comes about. For the learning process, the learning target is the nearest rain gauge data (i.e. rainfall rate) corresponding to the centered pixel (5, 5). To produce a data pool, the 9 × 9 pixel window is moved across the radar image sequentially, shifting location one pixel at a time on the radar image generated every 5-7 minutes. Finally, the total number of samples is 101337, including training, validation and testing data.

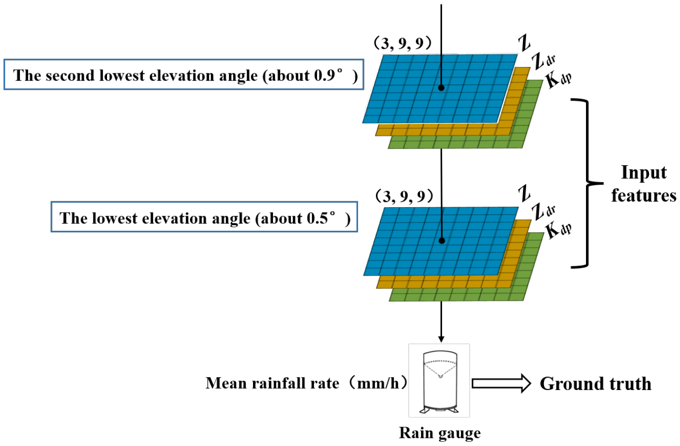


Fig. 4. Illustration of a training data sample for the devised deep learning models.

B. Model Architecture

Considering residual blocks, dense blocks, and RepVGG blocks are flexible, efficient, and generalizable, we use them as the core building blocks of our radar quantitative precipitation estimation neural networks (RQPENet). RQPENet$_{\mathrm{D1}}$ and RQPENet$_{\mathrm{D2}}$ both use dense blocks as the core building blocks, but differ in the number of model layers. RQPENet$_{\mathrm{R}}$ and RQPENet$_{\mathrm{V}}$ use residual blocks and RepVGG blocks as the core building blocks, respectively. Among them, we give a detailed introduction of how to use the dense block to build our RQPENet$_{\mathrm{D1}}$. The other three models can be achieved similarly. In this article, four DL models are evaluated to find which core building block is most suitable for QPE.

Figure 5 shows the overall RQPENet architecture using the dense blocks, i.e., RQPENet$_{\mathrm{D1}}$. At the beginning of RQPENet$_{\mathrm{D1}}$, the input is fed into one convolution layer with 96 kernels of size 3 × 3 followed by four dense blocks and three transition layers. To the end, the rainfall rate (R; units: mm/h-1) is obtained through an adaptive average pooling layer and a fully connected (FC) layer. The dense blocks and transition layers are the two primary components which are used recursively to form the model structure. The detailed structures of the dense blocks and transition layers are introduced below.

Each dense block is composed of several bottleneck layers. Every bottleneck layer has batch normalization, ReLU and convolution sub-layers, i.e., BN-ReLU-Conv(1 × 1) and BN-ReLU-Conv(3 × 3). The four dense blocks contains 6, 12, 36 and 24 bottleneck layers respectively. The dense block enables feature reuse through copy and concatenate operations, which increase direct connections between feature maps within a block and ensure maximum information exchange between various layers of the model. The transition layer consists of a 1 × 1 Conv layer that keeps the same number of channels, followed by an average pooling layer with stride 2. The benefit of the transition layer is to decrease both the model weight parameters and the feature maps size while extracting high-level implicit features.

In addition, RQPENet$_{\mathrm{D1}}$ involves a hyper-parameter k,

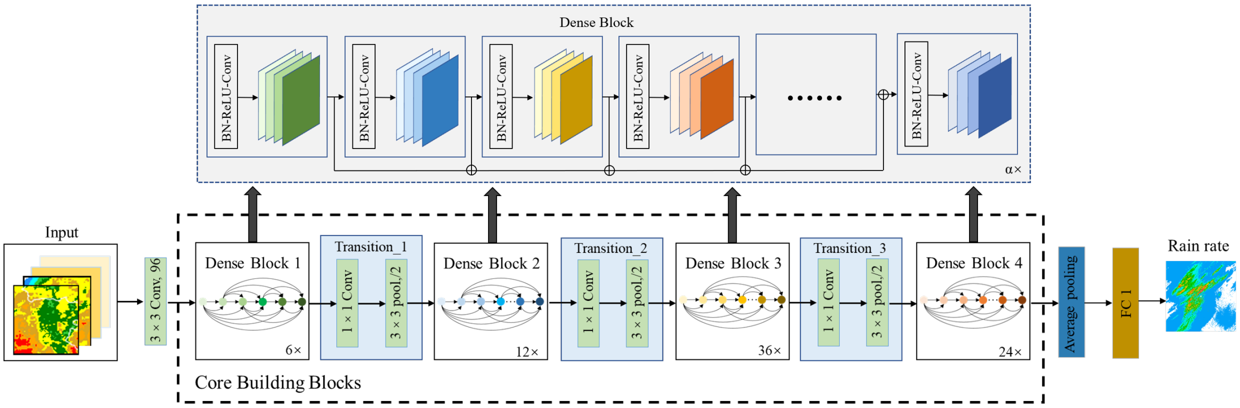


Fig. 5. The RQPENet$_{\mathrm{D1}}$ model architecture with four Dense Blocks, including the detailed structure of each Dense Block. Note that each Dense Block uses different number of bottleneck layers, although the overall structure of the four Dense Blocks is the same.

called growth rate, which sets the number of channels added at each layer within a dense block. In our experiments, k is set to 48. Design as described above strengthens feature transfer, effectively utilizes features in the spatial dimension, helps alleviate the vanishing-gradient problem, and reduces the number of model parameters. The loss function used by RQPENet$_{\mathrm{D1}}$ is defined as the mean square error $\mathcal{L}$ of radar rainfall estimates retrievals with respect to rain gauge rainfall observations:

$$\mathcal{L} = \frac{1}{N}\sum\left(Z - \hat{Z}\right)^2 \tag{1}$$

where Z and $\hat{Z}$ represent the rainfall rate estimated by the model and the training target label, respectively, and N is the total number of samples of the testing dataset. The model weighs are updated through the following equation:

$$w_{i,j,k(\mathrm{new})} = w_{i,j,k(\mathrm{old})} - \rho\frac{\partial\mathcal{L}}{\partial w_{i,j,k(\mathrm{old})}} \tag{2}$$

where $w_{i,j,k}$ represents the weight associated with the kth features of the jth bottleneck layer within dense block i, and ρ is learning rate.

After a detailed description of RQPENet$_{\mathrm{D1}}$, the other three methods are briefly described below.

(1) RQPENet$_{\mathrm{D2}}$. It has similar architecture to RQPENet$_{\mathrm{D1}}$, but differs in the third and fourth dense blocks. The bottleneck layers in dense block 3 and dense block 4 in RQPENet$_{\mathrm{D2}}$ and RQPENet$_{\mathrm{D1}}$ are (24, 16) and (36, 24) respectively, which means RQPENet$_{\mathrm{D2}}$ is a smaller network compare to RQPENet$_{\mathrm{D1}}$ and it has fewer parameters and less computational complexity.

(2) RQPENet$_{\mathrm{R}}$. It uses the residual module as the core component to form the model. Residual blocks consist of skip connections that allow inputs to skip one or more convolutional layers and are added to the output of the last convolutional layer in the module. Residual learning can improve attention to residual or hard-to-learn information in radar data. In RQPENet$_{\mathrm{R}}$, four similar residual blocks are placed in series, with each residual block containing 3, 4, 23 and 3 bottleneck layers, respectively.

(3) RQPENet$_{\mathrm{V}}$. The RepVGG block is used as the core building block in this model. RepVGG adds multiple parallel skip connections (i.e., multi-brunch structure) in the training stage based on the residual module. Each block has three parallel branches: a main branch with a convolution kernel size of 3x3, a shortcut branch with a convolution kernel size of 1x1, and a BN-only shortcut branch. In addition, it is converted to a single-brunch structure during testing stage to reduce the number of parameters used by the model. RQPENet$_{\mathrm{V}}$ has five stages, and each stage contains 1, 4, 6, 16 and 1 RepVGG blocks.

IV. EXPERIMENTS AND RESULTS

A. *Comparison Experiments*

To cross-compare the performance of DL models in precipitation estimation, this study also adopts five Z-R relations, including three commonly used parametric relations (Eq. 3a-3c) [18, 19], and two Z-R relations used in Florida (Eq. 3d-3e) [20], as follows:

$$R_a(Z) = 1.70 \times 10^{-2} \times Z^{0.714} \tag{3a}$$

$$R_b(Z) = 3.64 \times 10^{-2} \times Z^{0.625} \tag{3b}$$

$$R(Z, Z_{dr}) = 1.42 \times 10^{-2} \times Z^{0.770} \times (10^{\frac{Z_{dr}}{10}})^{-1.670} \tag{3c}$$

$$R_d(Z) = 2.62 \times 10^{-2} \times Z^{0.687} \tag{3d}$$

$$R(K_{dp}) = sign(K_{dp}) \times 54.3|K_{dp}|^{0.806} \tag{3e}$$

where Z is in mm^6m^{-3}; R denotes the rainfall rate in mm/h; K_{dp} is in $^\circ km^{-1}$ and $sign(K_{dp})$ item allows R to be negative. The subscripts a, b, and d of R correspond to the formula number, which is convenient for distinguishing in the tables.

B. *Evaluation Metrics*

In this article, the rainfall rate over an hour is accumulated to obtain the hourly radar rainfall amount. For evaluating RQPENet performance on radar precipitation estimating, the root mean squared error (RMSE), mean absolute error (MAE), the correlation coefficient (CC) and normalized standard error (NSE) are computed respectively as follow:

$$\mathrm{RMSE}=\sqrt{\frac{\sum_{n=1}^{N}\left(R_n\ -G_n\right)^2}{N}} \tag{4}$$

$$\mathrm{MAE}=\frac{\sum_{n=1}^{N}\left|R_n-G_n\right|}{N} \tag{5}$$

$$\mathrm{CC}=\frac{\sum_{n=1}^{N}\left(R_n-\bar{R_n}\right)\left(G_n-\bar{G_n}\right)}{\sqrt{\sum_{n=1}^{N}\left(R_n-\bar{R_n}\right)^2}\sqrt{\sum_{n=1}^{N}\left(G_n-\bar{G_n}\right)^2}} \tag{6}$$

$$\mathrm{NSE}=\frac{\sum_{n=1}^{N}\left|R_n-G_n\right|}{\sum_{n=1}^{N}G_n} \tag{7}$$

$$\mathrm{BIAS}=\frac{\sum_{n=1}^{N}R_n}{\sum_{n=1}^{N}G_n} \tag{8}$$

where R_n and G_n denote the radar precipitation estimates and rain gauge observations (mm) at the nth hour, respectively. $\bar{R_n}$ and $\bar{G_n}$ represent the mean of radar precipitation estimates and rain gauge observations (mm) in the test dataset. N is the total number of hours, which is 3950 in our experiments. Higher CC and BIAS ratio close to 1, or lower RMSE, MAE, NSE indicate better estimation performance.

In addition, this article introduces five widely used evaluation metrics in atmospheric science, namely, probability of detection (POD), false alarm ratio (FAR), critical success index (CSI), Heidke skill score (HSS), and Gilbert skill score (GSS). They are respectively defined as:

$$\mathrm{POD}=\frac{TP}{TP+FN} \tag{9}$$

$$\mathrm{FAR}=\frac{FP}{TP+FP} \tag{10}$$

$$\mathrm{CSI}=\frac{TP}{TP+FN+FP} \tag{11}$$

$$\mathrm{HSS}=\frac{2\times(TP\times TN-FP\times FN)}{(TP+FN)\times(FN+TN)+(\mathrm{TP}+\mathrm{FP})\times(FP+TN)} \tag{12}$$

$$\mathrm{GSS}=\frac{TP-Q}{TP+FN+FP-Q} \tag{13}$$

$$\mathrm{Q}=\frac{(TP+FN)\times(TP+FP)}{TP+FN+FP+TN} \tag{14}$$

where TP, TN, FN and FP denote true positive, true negative, false negative and false positive, respectively. If the radar precipitation estimates and ground truth are greater than a threshold, they are considered positive, otherwise negative. TP means that the estimates and ground truth are both positive, TN means that they are both negative, FN means that the estimate is negative while the ground truth is positive, whereas FP means that the estimate is positive while the ground truth is negative. Gilbert Skill Score (GSS) has another name "Equitable Threat Score (ETS)", which measures the fraction of events that the model is able to correctly estimate after removing the odds of random hits (i.e., Q). GSS is an improvement to CSI, which can balance misses and false alarms, and the evaluation score is fairer. Higher POD, CSI, HSS, GSS, or lower FAR indicate better estimation performance.

C. Results and Analysis

TABLE I: Evaluation results of hourly rainfall estimates from nine QPE methods.

Model	MAE (mm)	RMSE (mm)	CC	NSE (%)	BIAS (ratio)
RQPENet$_{\mathrm{D1}}$	**1.58**	**2.68**	**0.92**	**26**	**0.91**
RQPENet$_{\mathrm{D2}}$	1.66	2.83	**0.92**	27	0.85
RQPENet$_{\mathrm{V}}$	1.79	3.01	**0.92**	29	0.82
RQPENet$_{\mathrm{R}}$	1.75	2.98	0.91	29	0.85
$R_a(Z)$	3.26	5.07	0.76	54	0.54
$R_b(Z)$	3.39	5.34	0.77	56	0.48
$R(Z, Z_{dr})$	3.34	5.15	0.78	55	0.49
$R_d(Z)$	2.98	4.79	0.76	49	0.63
$R(K_{dp})$	2.42	3.74	0.85	40	0.81

TABLE II: Evaluation results of hourly rainfall estimates from nine QPE methods. A threshold of 2 mm is applied.

Model	POD	FAR	CSI	HSS	GSS
RQPENet$_{\mathrm{D1}}$	0.94	0.07	**0.87**	**0.75**	**0.60**
RQPENet$_{\mathrm{D2}}$	0.94	0.08	**0.87**	0.73	0.58
RQPENet$_{\mathrm{V}}$	0.93	0.08	0.86	0.72	0.56
RQPENet$_{\mathrm{R}}$	**0.95**	0.10	0.86	0.70	0.53
$R_a(Z)$	0.60	0.03	0.59	0.42	0.26
$R_b(Z)$	0.62	0.03	0.61	0.43	0.28
$R(Z, Z_{dr})$	0.58	**0.02**	0.58	0.41	0.26
$R_d(Z)$	0.67	0.04	0.65	0.46	0.30
$R(K_{dp})$	0.79	0.07	0.74	0.56	0.39

TABLE III: Evaluation results of hourly rainfall estimates from nine QPE methods. A threshold of 5 mm is applied.

Model	POD	FAR	CSI	HSS	GSS
RQPENet$_{\mathrm{D1}}$	**0.78**	0.08	**0.73**	**0.74**	**0.59**
RQPENet$_{\mathrm{D2}}$	0.75	0.07	0.71	0.72	0.57
RQPENet$_{\mathrm{V}}$	0.71	0.06	0.68	0.69	0.53
RQPENet$_{\mathrm{R}}$	0.75	0.07	0.71	0.72	0.56
$R_a(Z)$	0.45	0.06	0.44	0.45	0.29
$R_b(Z)$	0.40	0.05	0.39	0.41	0.26
$R(Z, Z_{dr})$	0.41	**0.04**	0.40	0.42	0.27
$R_d(Z)$	0.53	0.08	0.51	0.52	0.36
$R(K_{dp})$	0.70	0.12	0.64	0.64	0.47

We use the 2019 radar observations and rain gauge data as the testing dataset. Firstly, we evaluate the overall performance of the nine QPE methods, including four DL methods and five Z-R relations in terms of MAE, RMSE, CC, NSE and BIAS. Then, we derived the evaluation scores (i.e., POD, FAR, CSI,

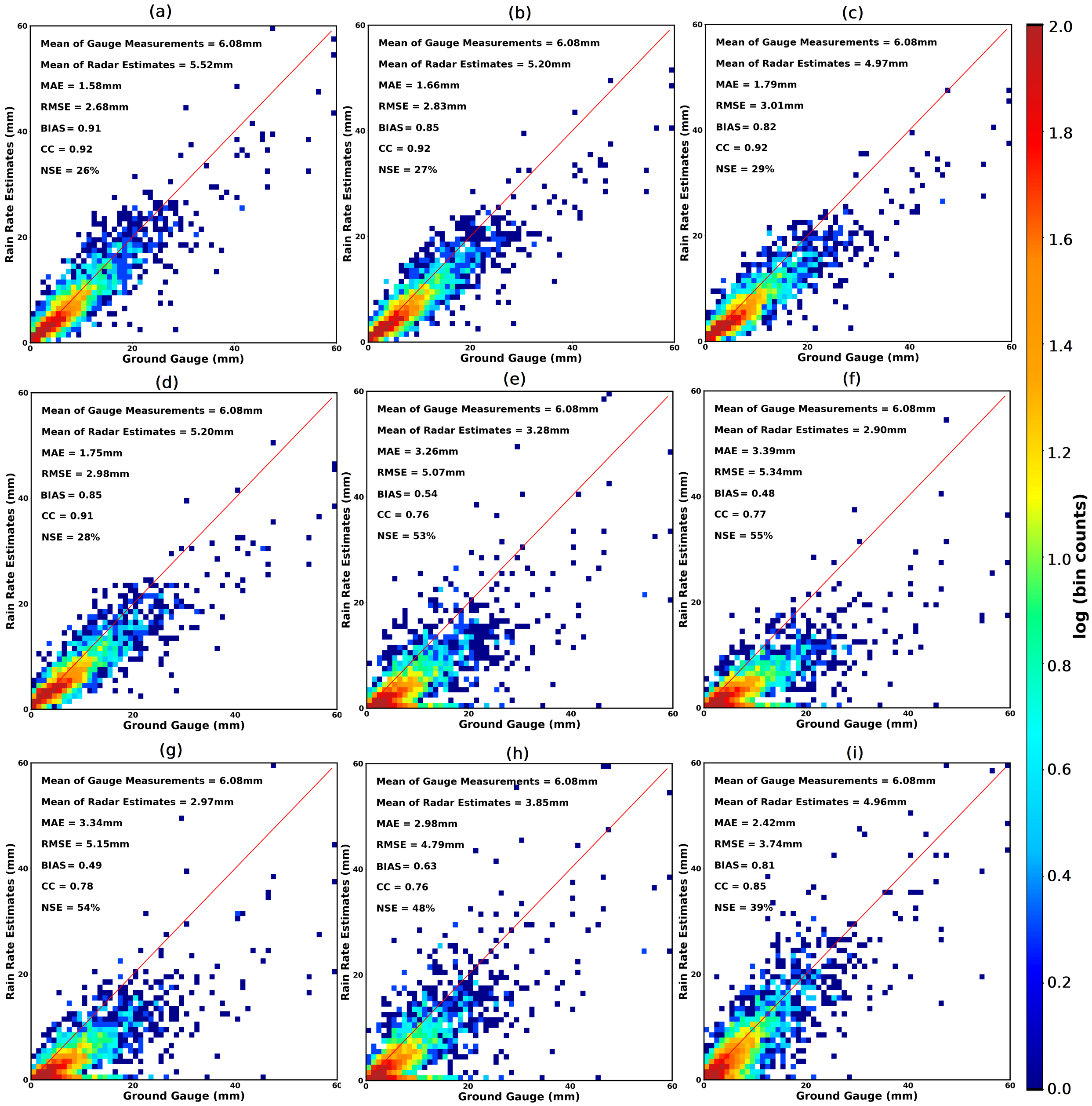


Fig. 6. Scatter plots of hourly precipitation estimates from different deep learning models versus evaluation gauge observations: (a)-(d) represent RQPENet$_{\mathrm{D1}}$, RQPENet$_{\mathrm{D2}}$, RQPENet$_{\mathrm{V}}$ and RQPENet$_{\mathrm{R}}$ respectively, and (e)-(i) correspond to $R_a(Z)$, $R_b(Z)$, $R(Z, Z_{dr})$, $R_d(Z)$ and $R(K_{dp})$, respectively.

HSS and GSS) for various RQPENet and Z-R relations under different precipitation thresholds (2 mm, 5 mm and 10 mm). In addition, three cases studies are selected to illustrate the performance of four RQPENet methods.

Table I shows the MAE, RMSE, CC, NSE and BIAS ratio values of the nine QPE methods. Overall, RQPENet$_{\mathrm{D1}}$ achieves superior estimation performance than other eight approaches. It has the lowest MAE of 1.58 mm, RMSE of 2.68 mm, NSE of 26%, highest CC of 0.92 and BIAS ratio of 0.91. RQPENet$_{\mathrm{D2}}$ achieves the second best performance. This shows that the dense blocks are more suitable than residual and RepVGG blocks on quantitative precipitation estimation task.

Figure 6 shows the scatter plots of precipitation estimation error for different QPE methods, where subplots (a)-(d) correspond to RQPENet$_{\mathrm{D1}}$, RQPENet$_{\mathrm{D2}}$, RQPENet$_{\mathrm{V}}$ and RQPENet$_{\mathrm{R}}$ respectively, and (e)-(i) correspond to $R_a(Z)$, $R_b(Z)$, $R(Z, Z_{dr})$, $R_d(Z)$ and $R(K_{dp})$ respectively. The horizontal axis is the hourly rainfall amount observed by the rain gauge, and the vertical axis is the precipitation estimated by the RQPENet. This plot uses the $y = log_{10}x$ function to represent the number of samples per scatter point, where x is the actual number of samples at a given scatter point, and y is the color-coded value in the color range. The closer the distribution of scatter points to the red diagonal, the more similar the quantitative precipitation estimates to the ground truth. Fig.

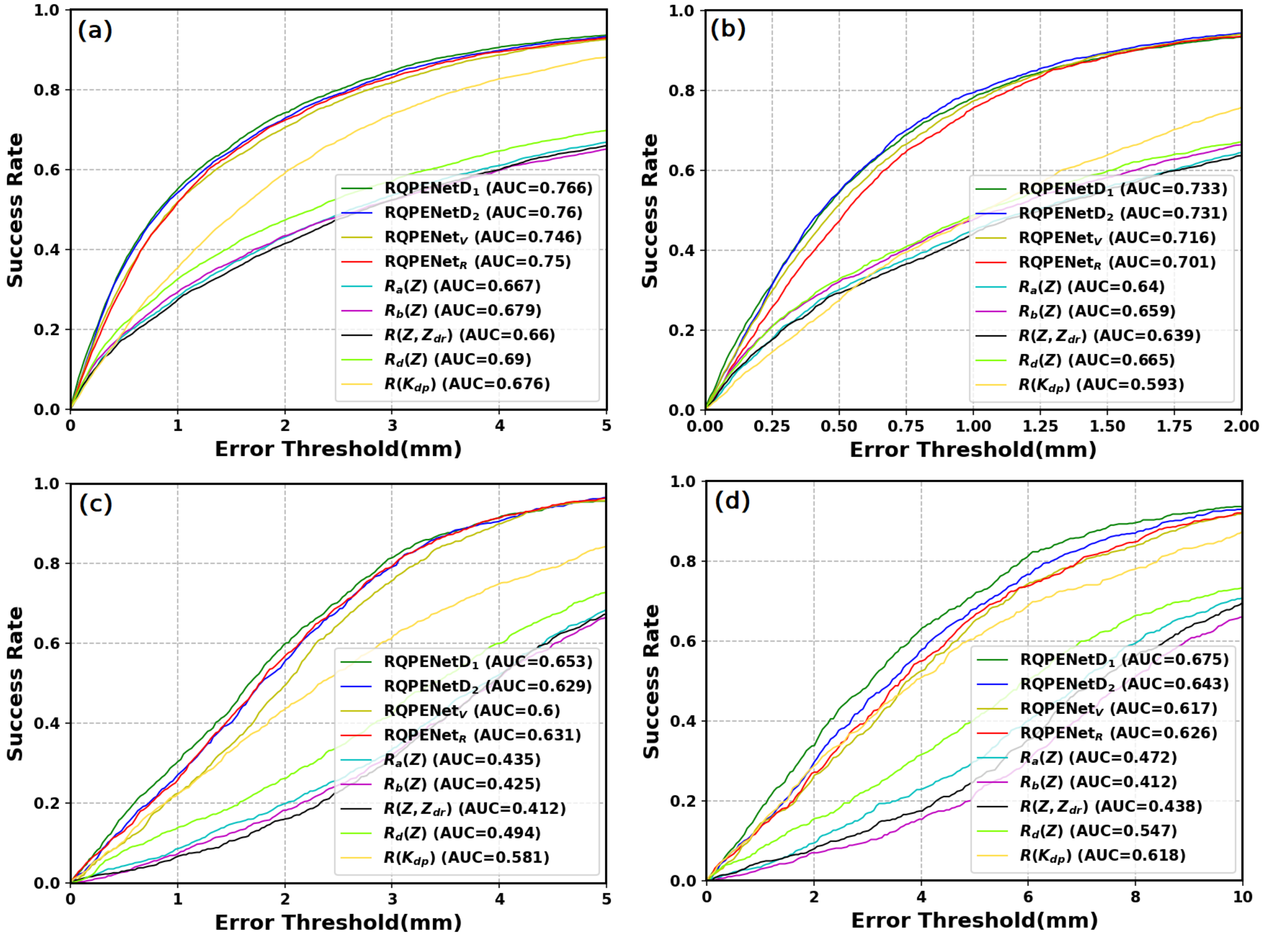


Fig. 7. The area under the curve (AUC) for the four deep learning methods and five Z-R relations: (a) is the evaluation result for all the precipitation samples; (b) is the evaluation result for the samples with precipitation less than 5 mm; (c) is the evaluation result for the samples with precipitation between 5 mm and 10 mm; (d) is the evaluation result for the samples with precipitation larger than 10 mm.

TABLE IV: Evaluation results of hourly rainfall estimates from nine QPE methods. A threshold of 10 mm is applied.

Model	POD	FAR	CSI	HSS	GSS
RQPENet$_{\mathrm{D1}}$	**0.73**	0.14	**0.65**	**0.75**	**0.59**
RQPENet$_{\mathrm{D2}}$	0.67	0.11	0.62	0.72	0.57
RQPENet$_{\mathrm{V}}$	0.63	0.10	0.59	0.70	0.54
RQPENet$_{\mathrm{R}}$	0.65	0.10	0.60	0.71	0.55
$R_a(Z)$	0.39	0.13	0.37	0.48	0.32
$R_b(Z)$	0.28	**0.07**	0.27	0.38	0.23
$R(Z, Z_{dr}))$	0.32	0.08	0.31	0.43	0.27
$R_d(Z)$	0.48	0.17	0.44	0.55	0.38
$R(K_{dp})$	0.67	0.19	0.58	0.69	0.52

6(a)-6(i) indicate that all nine methods generally underestimate the precipitation, especially at heavy precipitation events. As can be seen from Fig. 6(a), the precipitation estimates of RQPENet$_{\mathrm{D1}}$ are the closest to the rain gauge observations, with an average radar precipitation estimate is 5.52mm. This shows that the dense blocks-based models can alleviate the problem of precipitation underestimation to a certain extent, while RQPENet$_{\mathrm{D1}}$ is more effective than RQPENet$_{\mathrm{D2}}$. Fig. 6(e)-6(i) are the scatter plots of five traditional methods for precipitation estimation. It can be seen that the estimation results are not as good as those from the machine learning models.

Next, we demonstrate the QPE performance for different precipitation intensities based on three hourly rainfall accumulation thresholds, i.e., 2 mm, 5 mm and 10 mm respectively. Tables II, III and IV show the quantitative evaluation results of POD, FAR, CSI, HSS and GSS using nine methods for QPE task. For all thresholds, RQPENet$_{\mathrm{D1}}$ achieves the highest CSI, HSS and GSS values which represent it has the best comprehensive performance. For the 2 mm threshold (Tables II), RQPENet$_{\mathrm{R}}$ achieves the highest POD value. For the 5 and 10 mm threshold (Tables III and IV), $R(Z, Z_{dr})$ and $R_b(Z)$ achieve the lowest FAR values, which mean less false alarms. But it should be mentioned that, although POD and FAR are important evaluation metrics, they cannot be used separately to evaluate a method. If a method has only a high POD value or a low FAR value, it does not mean that the method is good. It is only meaningful when a method has a high POD and a low FAR simultaneously. In general, RQPENet$_{\mathrm{D1}}$ has relatively

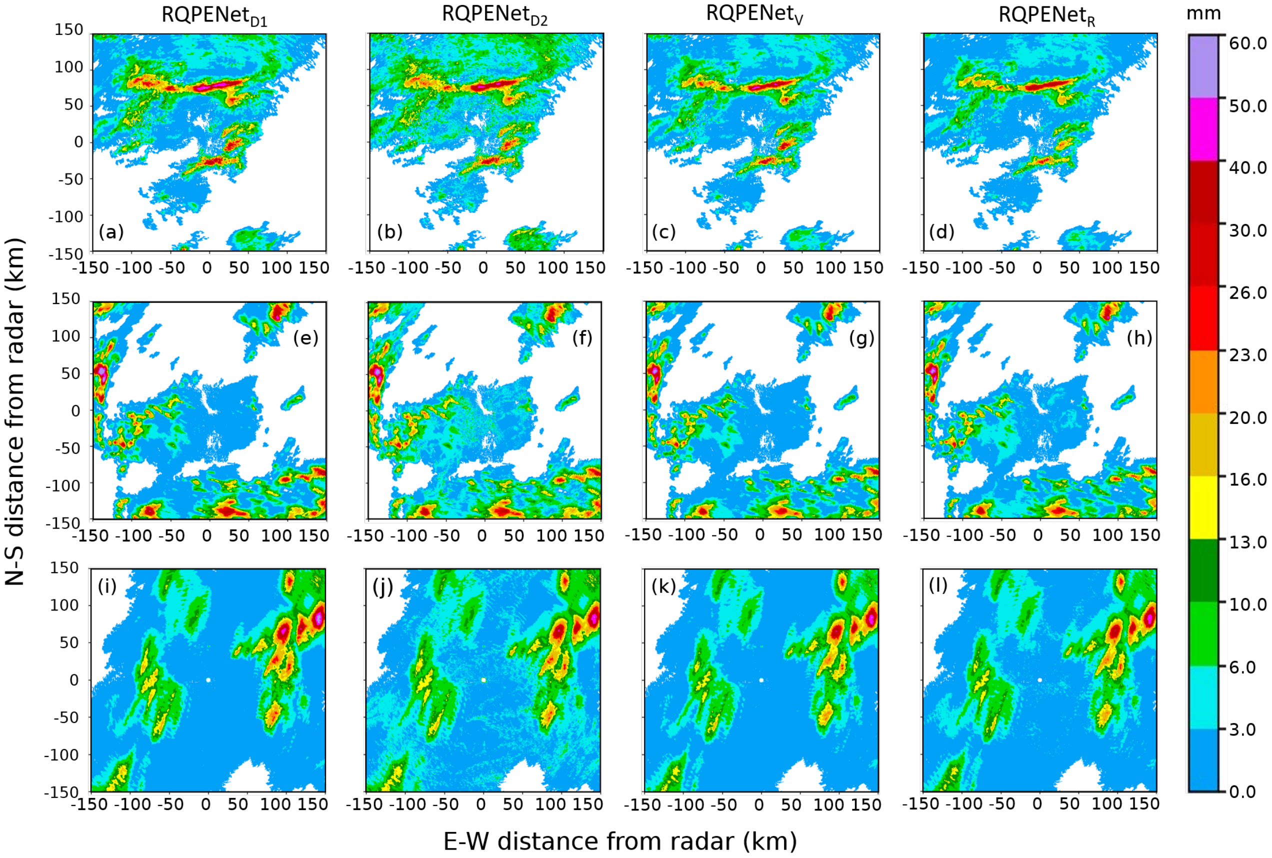


Fig. 8. Rainfall maps derived from four deep learning QPE models: (a)-(d) represent the precipitation maps from RQPENet$_{\mathrm{D1}}$, RQPENet$_{\mathrm{D2}}$, RQPENet$_{\mathrm{V}}$ and RQPENet$_{\mathrm{R}}$ at 21:00 UTC on 19 June 2019, respectively; (e)-(h) are the same as (a)-(d), but at 18:00 UTC on 1 August 2019; (i)-(l) are the same as (a)-(d), but at 14:00 UTC on 19 October 2019.

high POD values and relatively low FAR values, which make it achieve the best overall performance.

Figure 7 shows the area under the curve (AUC) for the nine QPE approaches. The horizontal axis is the error threshold between the precipitation estimates and the ground truth. The vertical axis is the estimated success rate when the error value between the precipitation estimates and the ground truth is less than this threshold. The higher the AUC, the better the model is at precipitation estimation. As shown in Fig. 7(a), RQPENet$_{\mathrm{D1}}$ outperforms the other three models considering all testing samples; RQPENet$_{\mathrm{D2}}$ achieves the second best performance. Then, we separate all testing samples into three categories to calculate the AUC values: < 5 mm, [5 mm, 10 mm], > 10 mm. Fig. 7(b) to (d) show that RQPENet$_{\mathrm{D1}}$ achieves the best performance in weak to heavy precipitation cases. In general, the results reveal that the dense blocks-based models, i.e. RQPENet$_{\mathrm{D1}}$ and RQPENet$_{\mathrm{D2}}$, perform best in the QPE task in both heavy rain and weak rain cases, and the RQPENet$_{\mathrm{D1}}$ model has the best comprehensive performance.

Finally, we select three real cases to demonstrate the performance of four RQPENet models. Fig. 8 shows the precipitation maps derived from the four models at 21:00 UTC on 19 June 2019, 18:00 UTC on 1 August 2019, and 14:00 UTC on 19 October 2019. Tables V, VI, and VII are the quantitative evaluation results of these three cases, respectively. Scrutinizing the rainfall maps in Fig. 8 (a)-(d), which are derived from RQPENet$_{\mathrm{D1}}$, RQPENet$_{\mathrm{D2}}$, RQPENet$_{\mathrm{V}}$ and RQPENet$_{\mathrm{R}}$, respectively, at 21:00 UTC on 19 June 2019. It can be seen that compared with RQPENet$_{\mathrm{D1}}$, RQPENet$_{\mathrm{R}}$ and RQPENet$_{\mathrm{V}}$ underestimate precipitation in heavy precipitation areas, which is consistent with the analysis in Fig. 6. The evaluation results in Table V show that RQPENet$_{\mathrm{D1}}$ achieves the best performance in this case. Similar results can be found for the other two cases.

TABLE V: Evaluation results of four QPE algorithms for hourly rainfall estimates at 21:00 UTC, 19 June 2019.

Model	MAE (mm)	RMSE (mm)	CC	NSE (%)	BIAS (ratio)
RQPENet$_{\mathrm{D1}}$	**1.32**	**1.87**	0.66	**32**	**0.80**
RQPENet$_{\mathrm{D2}}$	1.44	1.98	0.65	35	0.74
RQPENet$_{\mathrm{V}}$	1.51	2.07	**0.69**	37	0.68
RQPENet$_{\mathrm{R}}$	1.65	2.19	0.63	40	0.67

V. CONCLUSION AND DISCUSSION

In this article, we designed four deep learning models for quantitative precipitation estimation using dense blocks,

TABLE VI: Evaluation results of four QPE algorithms for hourly rainfall estimates at 18:00 UTC, 1 August 2019.

Model	MAE (mm)	RMSE (mm)	CC	NSE (%)	BIAS (ratio)
$\text{RQPENet}_{\text{D1}}$	**1.63**	**2.26**	0.90	**23**	**0.90**
$\text{RQPENet}_{\text{D2}}$	1.76	2.38	**0.91**	25	0.87
$\text{RQPENet}_{\text{V}}$	1.98	2.74	**0.91**	28	0.80
$\text{RQPENet}_{\text{R}}$	2.03	2.71	0.89	29	0.84

TABLE VII: Evaluation results of four QPE algorithms for hourly rainfall estimates at 14:00 UTC, 19 October 2019.

Model	MAE (mm)	RMSE (mm)	CC	NSE (%)	BIAS (ratio)
$\text{RQPENet}_{\text{D1}}$	**2.25**	**2.98**	**0.78**	**38**	**0.71**
$\text{RQPENet}_{\text{D2}}$	2.36	3.11	**0.78**	40	0.67
$\text{RQPENet}_{\text{V}}$	2.67	3.41	0.75	46	0.62
$\text{RQPENet}_{\text{R}}$	2.59	3.34	0.76	44	0.63

RepVGG blocks and residual blocks, namely, $\text{RQPENet}_{\text{D1}}$, $\text{RQPENet}_{\text{D2}}$, $\text{RQPENet}_{\text{V}}$ and $\text{RQPENet}_{\text{R}}$. Dual-polarization radar observations and rain gauge data are used in these deep leaning models. In the data preprocessing stage, we filter the non-precipitation echoes from radar observations data and align the radar observations and rain gauge observations in the spatial and temporal domain. The training, validation and testing dataset are collected in Florida during 2016-2019. Ten evaluation metrics are used in our study for quantitative analysis. The experimental results show that dense blocks-based models achieve higher accuracy results in QPE and alleviates the problem of rainfall underestimation in heavy precipitation area compared to the other methods. This is because the unique dense connection structure in dense blocks strengthens the information flow and transmission efficiency across feature maps, encourages feature maps reuse and fully utilizes spatiotemporal information. In addition, $\text{RQPENet}_{\text{D1}}$ with deeper layers performs better than $\text{RQPENet}_{\text{D2}}$ in terms of overall evaluation metrics and specific precipitation events.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for providing careful review and comments on this article. We also acknowledge NVIDIA Corporation for donating the graphics processing units (GPUs) used in this research. The radar data used in this article were obtained from the National Centers for Environmental Information (NCEI). The intermediate products, such as the rainfall features used in training the deep learning models and the model parameters are available upon request.

REFERENCES

[1] V. Chandrasekar, H. Chen, and B. Philips, "Principles of high-resolution radar network for hazard mitigation and disaster management in an urban environment," *Journal of the Meteorological Society of Japan. Ser. II*, vol. 96, pp. 119–139, 2018.

[2] R. Cifelli, V. Chandrasekar, L. Herdman, D. D. Turner, A. B. White, T. I. Alcott, M. Anderson, P. Barnard, S. K. Biswas, M. Boucher, J. Bytheway, H. Chen, H. Cutler, J. M. English, L. Erikson, F. Junyent, D. J. Gottas, J. Jasperse, L. E. Johnson, J. Krebs, J. van de Lindt, J. Kim, M. Leon, Y. Ma, M. Marquis, W. Moninger, G. Pratt, C. Radhakrishnan, M. Shields, J. Spaulding, B. Tehranirad, and R. Webb, "Advanced quantitative precipitation information: Improving monitoring and forecasts of precipitation, streamflow, and coastal flooding in the san francisco bay area," *Bulletin of the American Meteorological Society*, 2022. [Online]. Available: https://journals.ametsoc.org/view/journals/bams/aop/BAMS-D-21-0121.1/BAMS-D-21-0121.1.xml

[3] H. Chen and V. Chandrasekar, "The quantitative precipitation estimation system for dallas–fort worth (dfw) urban remote sensing network," *Journal of Hydrology*, vol. 531, pp. 259–271, 2015.

[4] V. Bringi and V. Chandrasekar, *Polarimetric Doppler weather radar: principles and applications*. Cambridge university press, 2001.

[5] H. Chen, V. Chandrasekar, and R. Bechini, "An improved dual-polarization radar rainfall algorithm (drops2. 0): Application in nasa ifloods field campaign," *Journal of Hydrometeorology*, vol. 18, no. 4, pp. 917–937, 2017.

[6] D. Willie, H. Chen, V. Chandrasekar, R. Cifelli, C. Campbell, and D. Reynolds, "Evaluation of multisensor quantitative precipitation estimation methodologies," in *2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS*. IEEE, 2013, pp. 2219–2221.

[7] J. Zhang, K. Howard, C. Langston, S. Vasiloff, B. Kaney, A. Arthur, S. Van Cooten, K. Kelleher, D. Kitzmiller, F. Ding *et al.*, "National mosaic and multi-sensor qpe (nmq) system: Description, results, and future plans," *Bulletin of the American Meteorological Society*, vol. 92, no. 10, pp. 1321–1338, 2011.

[8] D.-J. Seo and J. Breidenbach, "Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements," *Journal of Hydrometeorology*, vol. 3, no. 2, pp. 93–111, 2002.

[9] D.-J. Seo, J. Breidenbach, and E. Johnson, "Real-time estimation of mean field bias in radar rainfall data," *Journal of Hydrology*, vol. 223, no. 3-4, pp. 131–147, 1999.

[10] T. A. Seliga and V. Bringi, "Potential use of radar differential reflectivity measurements at orthogonal polarizations for measuring precipitation," *Journal of Applied Meteorology and Climatology*, vol. 15, no. 1, pp. 69–76, 1976.

[11] A. Ryzhkov and D. Zrnić, "Comparison of dual-polarization radar estimators of rain," *Journal of Atmospheric and Oceanic Technology*, vol. 12, no. 2, pp. 249–256, 1995.

[12] R. Cifelli and V. Chandrasekar, "Dual-polarization radar rainfall estimation," *Washington DC American Geophysical Union Geophysical Monograph Series*, vol. 191, pp.

105–125, 2010.

[13] P.-E. Kirstetter, J. J. Gourley, Y. Hong, J. Zhang, S. Moazamigoodarzi, C. Langston, and A. Arthur, “Probabilistic precipitation rate estimates with ground-based radar networks,” *Water Resources Research*, vol. 51, no. 3, pp. 1422–1442, 2015. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014WR015672

[14] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.

[15] X. Ding, X. Zhang, N. Ma, J. Han, G. Ding, and J. Sun, “Repvgg: Making vgg-style convnets great again,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 13 733–13 742.

[16] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

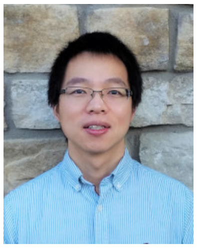
[17] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[18] S. E. Giangrande and A. V. Ryzhkov, “Estimation of rainfall based on the results of polarimetric echo classification,” *Journal of Applied Meteorology and Climatology*, vol. 47, no. 9, pp. 2445 – 2462, 2008. [Online]. Available: https://journals.ametsoc.org/view/journals/apme/47/9/2008jamc1753.1.xml

[19] M. Jang, D.-I. Lee, C.-H. You, D.-S. Kim, M. Maki, J.-H. Jeong, and H. Uyeda, “Quantitative precipitation estimates from radar reflectivity corrected by the sma method,” *Atmospheric Research*, vol. 104-105, pp. 111–118, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169809511002547

[20] E. A. Brandes, G. Zhang, and J. Vivekanandan, “Experiments in rainfall estimation with a polarimetric radar in a subtropical environment,” *Journal of Applied Meteorology*, vol. 41, no. 6, pp. 674 – 685, 2002. [Online]. Available: https://journals.ametsoc.org/view/journals/apme/41/6/1520-0450_2002_041_0674_eirewa_2.0.co_2.xml

Wenyuan Li received the B.Sc. and M.Sc. degree in electrical engineering from Shandong University of Science and Technology, Qingdao, China, in 2017 and 2020, respectively. She is currently pursuing the Ph.D. degree in electrical engineering at the Ocean University of China, Qingdao, China. She is also a visiting scholar in the Department of Electrical and Computer Engineering at Colorado State University, Fort Collins, CO, USA. Her research interests include atmospheric remote sensing, machine learning, and explainable learning.



Haonan Chen (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from Colorado State University (CSU), Fort Collins, CO, USA, in 2017. He has been an Assistant Professor in the Department of Electrical and Computer Engineering at CSU since August 2020. He is also an Affiliate Faculty with the Data Science Research Institute (DSRI) at CSU. Before joining the CSU faculty, he worked with the Cooperative Institute for Research in the Atmosphere (CIRA) and the National Oceanic and Atmospheric Administration (NOAA) Physical Sciences Laboratory, Boulder, CO, USA, from 2012 to 2020, first as a Research Student, then a National Research Council Research Associate and a Radar, Satellite, and Precipitation Research Scientist. His research interests span a broad range of remote sensing and multidisciplinary data science, including radar and satellite remote sensing of natural disasters, polarimetric radar systems and networking, clouds and precipitation observations and processes, big data analytics, and deep learning.

Dr. Chen was a recipient of the National Science Foundation (NSF) Faculty Early Career Development Program (CAREER) Award, Ralph E. Powe Junior Faculty Enhancement Award, IEEE Geoscience and Remote Sensing Society (GRSS) Early Career Award, American Meteorological Society (AMS) Editor’s Award, and CIRA Research and Service Initiative Award. He serves as an Associate Editor for the Journal of Atmospheric and Oceanic Technology, URSI Radio Science Bulletin, and the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATION AND REMOTE SENSING.

Lei Han received the B.Sc. degree in engineering mechanics and the M.Sc. degree in automatic control from Harbin Engineering University, Harbin, China, in 1998 and 2001, respectively, and the Ph.D. degree in atmospheric remote sensing from Beijing University, Beijing, China, in 2008. From 2001 to 2004, he was a Software Engineer with Lucent R&D, Qingdao, China. From 2013 to 2014, he was Visiting Scholar with the Earth Observation Laboratory (EOL), National Center for Atmospheric Research (NCAR), Boulder, CO, USA. He is currently a Professor with the College of Information Science and Engineering, Ocean University of China, Qingdao. His research interests include radar meteorology and machine learning.