Application of Machine Learning Techniques to Improve Multi-Radar Multi-Sensor (MRMS) Precipitation Estimates in the Western United States

ANDREW P. OSBORNE,[a,b] JIAN ZHANG,[b] MICHEAL J. SIMPSON,[a,b] KENNETH W. HOWARD,[b] AND
STEPHEN B. COCKS[a,b]

[a] *Cooperative Institute for Severe and High-Impact Weather Research and Operations, University of Oklahoma, Norman, Oklahoma*
[b] *NOAA/OAR National Severe Storms Laboratory, Norman, Oklahoma*

ABSTRACT: The Multi-Radar Multi-Sensor (MRMS) system produces a suite of hydrometeorological products that are widely used for applications such as flash flood warning operations, water resource management, and climatological studies. The MRMS radar-based quantitative precipitation estimation (QPE) products have greater challenges in the western United States compared to the eastern two-thirds of the CONUS due to terrain-related blockages and gaps in radar coverage. Further, orographic enhancement of precipitation often occurs, which is highly variable in space and time and difficult to accurately capture with physically based approaches. A deep learning approach was applied in this study to understand the correlations between several interacting variables and to obtain a more accurate precipitation estimation in these scenarios. The model presented here is a convolutional neural network (CNN), which uses spatial information from small grids of several radar variables to predict an estimated precipitation value at the central grid point. Several case analyses are presented along with a yearlong statistical evaluation. The CNN model 24-h QPE shows higher accuracy than the MRMS radar QPE for several cool season atmospheric river events. Areas of consistent improvement from the CNN model are highlighted in the discussion along with areas where the model can be further improved. The initial findings from this work help set the foundation for further exploration of machine learning techniques and products for precipitation estimation as part of the MRMS operational system.

SIGNIFICANCE STATEMENT: This study explores the development and use of a deep learning model to generate precipitation fields in the complex terrain of the western United States. Generally, the model is able to improve on the statistical performance of existing radar-based precipitation estimation methods for several case studies and over a long-term period in 2021. We explore the patterns associated with certain areas of strong performance and suggest potential means of improving areas with weaker performance. These initial results indicate the potential of deep learning to supplement radar-based approaches in areas with observational limitations.

KEYWORDS: Complex terrain; Precipitation; Deep learning

---

1. Introduction

The Multi-Radar Multi-Sensor (MRMS; Zhang et al. 2016) system provides operational high-resolution and rapid update quantitative precipitation estimates (QPEs) for hydrological situational awareness including flash flood warning operations, streamflow predictions, and water resource management. The current MRMS produces two classes of precipitation products: the first is based primarily on quality controlled radar data (Wang et al. 2019; Cocks et al. 2019; Zhang et al. 2020) while the second (Martinaitis et al. 2020) is a multisensor QPE that blends radar, hourly gauge observations, monthly precipitation climatologies from the Parameter-Elevation Regressions on Independent Slopes Model (PRISM; Daly et al. 2008, 1994; www.prism.oregonstate.edu), and quantitative precipitation forecasts (QPFs) from numerical weather prediction (NWP) models. The radar-based QPE is produced every 2 min at a latency of less than 90 s and demonstrates high accuracy when validated against gauges in areas with seamless radar coverage

in the lower atmosphere. However, the radar-based QPE still has challenges in areas of complex terrain, such as the western United States. The multisensor QPE has higher accuracy by comparison but has a 15–60-min latency due to the real-time gauge data availability.

The MRMS radar-based QPE was initially generated based on reflectivity–rain rate relationships (Zhang et al. 2016) and has evolved to use dual polarization (dual-pol) radar variables for attenuation-based rate estimation (Cocks et al. 2019; Zhang et al. 2020). The western United States has large voids in the low-level radar coverage due to beam blockages from high terrain and other radar siting issues (Germann et al. 2006; Maddox et al. 2002). Further, the west experiences unique microphysical processes, especially with atmospheric river events (Zhu and Newell 1998; Neiman et al. 2008; Chen et al. 2020) where substantial precipitation growth occurs in the lowest levels due to orographic forcing. The radar beam frequently scans above these processes and the high spatial variability of the rain rates in these scenarios prevents a simple correction factor from being applied. The current physically based MRMS approach may not fully encompass all the variables that are influencing the complex types of precipitation regimes seen in the western United States. Radar-based

Corresponding author: Andrew Osborne, andrew.osborne@noaa.gov

TABLE 1. List of input variables used in the CNN model setup and testing.

No.	Variables
1	Seamless hybrid scan reflectivity (SHSR)
2	SHSR height (SHSRH)
3	Vertically integrated liquid (VIL)
4	Radar quality index (RQI)
5	Composite reflectivity (CREF)
6	Reflectivity at lowest altitude (RALA)
7	Brightband bottom (BB_BOTTOM)
8	Brightband top (BB_TOP)
9	Reflectivity at 0°C
10	Reflectivity at −5°C
11	Reflectivity at −10°C
12	Reflectivity at −15°C
13	Reflectivity at −20°C

estimation methods use only the near-surface radar variables and a precomputed rain rate relationship for QPE generation, which tends to have a dry bias (Zhang et al. 2012a; Martinaitis et al. 2020) during heavy rain in these regions. To address these limitations, a statistically based, deep learning approach

that can incorporate a wide range of input variables is being explored in this work.

Deep learning (Chollet 2018) has become increasingly popular as its efficacy for solving problems across a wide range of scientific disciplines has been demonstrated. One type of deep learning model is a convolutional neural network (CNN), which uses spatial filters to parse hierarchical information from input data and propagate the learned patterns through several layers to come up with a prediction (LeCun et al. 1998). A CNN model can take the spatial fields of several interacting variables (predictors) and make a prediction for a target variable (predictand). Meteorological applications tend to have large datasets with spatially and temporally coherent information, and the pattern recognition abilities of CNN models fit well with this type of data (Sadeghi et al. 2019). The convolutional filters of the CNN model allow it to convolve spatial features necessary to make predictions, rather than using each grid point as a feature or relying on users for manual feature extraction (Pan et al. 2019). CNNs have been shown to provide value in the dynamical modeling space over traditional NWP approaches that run into limitations at the increasingly small scales of simulations (Pan et al. 2019). CNNs were also applied to precipitation estimation problems
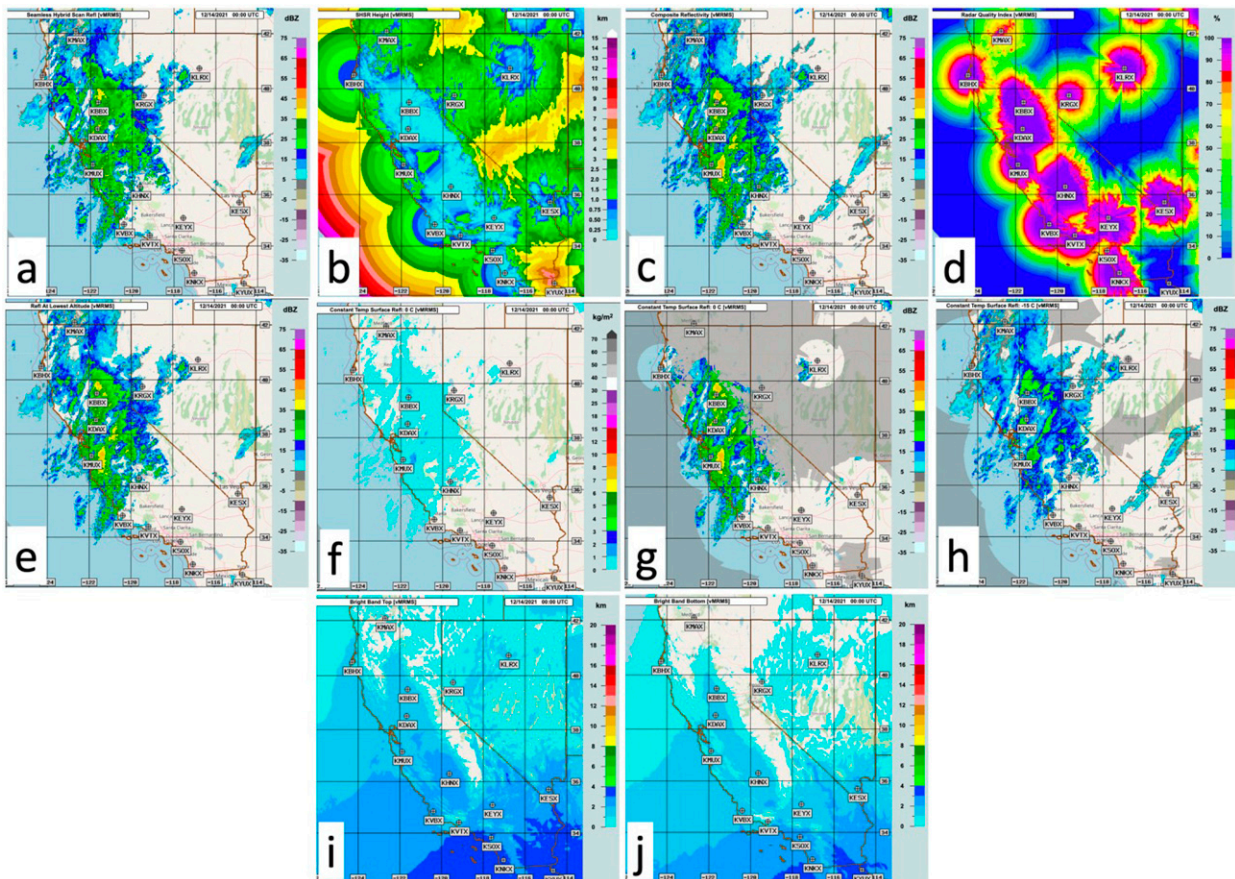

FIG. 1. Example input fields valid at 0000 UTC 14 Dec 2021: (a) SHSR, (b) SHSRH, (c) CREF, (d) RQI, (e) RALA, (f) VIL, (g) reflectivity at 0°C height, (h) reflectivity at −15°C height, (i) height of brightband top, and (j) height of brightband bottom.
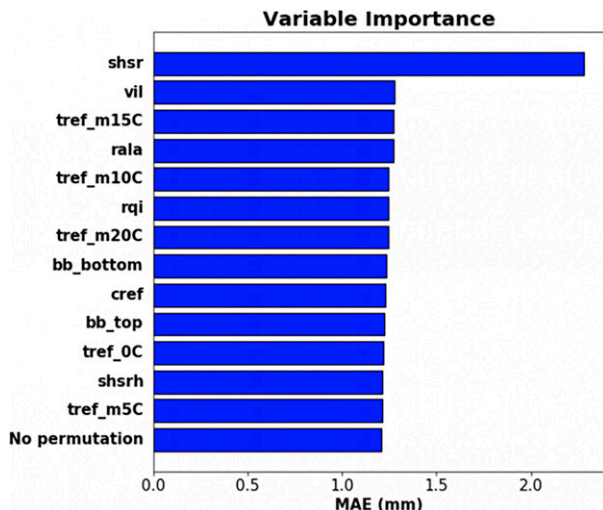
FIG. 2. Permutation importance testing results, where each input variable is randomly permuted separately and the resulting impact on the model averaged MAE (mm) is shown. The model is run on the full training dataset for these permutation experiments.

TABLE 2. Time periods of the training, validation, and simulation data.

Training	Validation	Simulation
		6 Mar 2019
February 2019	February 2020	27 Nov 2019
December 2019		17 Jan 2020
December 2020		13 Mar 2020
		All precipitation days 2021

height (SHSRH); radar QPE quality index (RQI)] and the melting layer (brightband top and bottom). The reflectivities at different temperature levels and the VIL field are included as they can provide information regarding the microphysical processes that contributed to precipitation. All input fields are 2D grids that come from the MRMS mosaic dataset with horizontal grid spacing of 1 km (Zhang et al. 2016), and Fig. 1 shows a few examples of the variables. The SHSR (Fig. 1a), which is the lowest level reflectivity corrected for range/height variations of reflectivity and blockage effects, is the main variable used in the physically based radar QPE and thus a key input variable for the CNN model. Permutation testing (Fig. 2) confirms the SHSR variable is the most important to the CNN model as the model error increases the most when SHSR is randomly permuted. Since the SHSR observations come from different heights at different ranges from the radar, the SHSRH (Fig. 1b) field is included as an input variable to help mitigate range dependent errors in the QPE. Other variables include composite reflectivity (CREF; Fig. 1c), RQI (Fig. 1d), reflectivity at lowest altitude (RALA; Fig. 1e), VIL (Fig. 1f), reflectivity at several temperature levels (Figs. 1g,h) and brightband top (Fig. 1i) and bottom (Fig. 1j). CREF is derived from the MRMS 3D reflectivity mosaic and is the maximum value within each grid column. RALA is also derived from the MRMS 3D reflectivity mosaic and is the nonmissing reflectivity at the lowest altitude within each grid column. RALA is similar to SHSR but without the brightband corrections. The RQI field (Martinaitis et al. 2020; Zhang et al. 2012b) is a function of the SHSRH and radar beam blockages and represents the relative correlations of the SHSR observations with the surface precipitation. The brightband top and bottom height fields are a blend of the radar-derived brightband top and bottom heights (Zhang et al. 2008) and the 2D freezing-level height from NWP models.

in different climate regions (Sadeghi et al. 2019; Miao et al. 2019; Sha et al. 2021; Yo et al. 2021). Initial work in this area from the MRMS research group looked at the precipitation estimation capability of a CNN model over Taiwan using single-radar, dual-pol variables as input data. The results from the study showed some statistical improvement in CNN QPE compared to the radar-based, dual-pol synthetic QPE used operationally in Taiwan, especially in mountainous terrain. The research presented here will examine how a similar CNN model setup performs in the mountainous area of the western United States. An expanded set of input variables, relative to that study, will be explored including radar observation heights, melting layer heights, reflectivities at different temperature levels, and vertically integrated liquid (VIL).

Section 2 of this paper will include a brief description of the CNN model, the processes for the model training, and the experimental setup. Then in section 3 we will present the CNN precipitation estimates for several cases, comparing the CNN output to the physically based MRMS radar QPEs and discussing the model's performance. The final section will summarize our conclusions and suggest potential extensions of this work. A list of acronyms and definitions are presented in the appendix.

2. Methodology

a. Input data

For the CNN model presented here, the input fields consist of 5 km × 5 km grids of 13 different radar-based variables (Table 1) and the CNN output is a precipitation value at the central grid point. Several different reflectivity fields are used along with additional variables related to the radar sampling characteristics [seamless hybrid scan reflectivity (SHSR)

TABLE 3. Overview of the precipitation size bins used to separate training examples. For each of these bins, the value to the left is exclusive (>) and the value to the right is inclusive (≤).

Precipitation size bins		
0–1 mm	10–12 mm	22–24 mm
1–2 mm	12–14 mm	24–26 mm
2–4 mm	14–16 mm	26–28 mm
4–6 mm	16–18 mm	28–30 mm
6–8 mm	18–20 mm	>30 mm
8–10 mm	20–22 mm	

TABLE 4. List of maximum and minimum values from the training dataset for each input variable. These values represent the summation of the 30 radar input field files available over an hour. These maximum and minimum values are used to scale the respective input variable for each data sample.

Variable name	Max value	Min value
SHSR	1443.5 dBZ	0 dBZ
VIL	107.8 kg m^{-2}	0 kg m^{-2}
RQI	30	0
CREF	1543 dBZ	0 dBZ
RALA	1411 dBZ	0 dBZ
BB_BOTTOM	97 707 m	0 m
BB_TOP	118 707 m	0 m
SHSRH	161.8 km	0.03 km
Reflectivity at 0°C	1393 dBZ	0 dBZ
Reflectivity at −5°C	1218 dBZ	0 dBZ
Reflectivity at −10°C	1130.5 dBZ	0 dBZ
Reflectivity at −15°C	1081 dBZ	0 dBZ
Reflectivity at −20°C	991 dBZ	0 dBZ

The truth dataset used in the training process was initially from hourly precipitation gauge observations in the Meteorological Assimilation Data Ingest System (MADIS; https://madis.ncep.noaa.gov/). The gauge data were quality controlled based on a comparison to the MRMS radar QPEs (Qi et al. 2016; Martinaitis et al. 2021). The model was later tested using the MRMS multisensor hourly accumulated QPE ("Q3MS"; Martinaitis et al. 2020) as a gridded truth dataset with more spatially continuous coverage compared to the irregularity of the gauge observations. Each grid point was treated as a separate training example with its QPE value being used to constrain the model. The resulting model output was similar whether the gauge truth or gridded truth datasets were used on their own or combined in a single training dataset. For the CNN model presented here, 3 months of training data were used with 2 months of hourly gauge truth data (February and December 2019) and 1 month of gridded hourly Q3MS truth data (December 2020) (Table 2). An additional month of hourly gauge truth data (February 2020) was used for the model validation process. The input variables (listed in Table 1) are available every 2 min, but both of the truth datasets contain hourly precipitation information available at the top of each hour. This necessitated an accumulation of the 2-min input variable fields to hourly values for the training process. This approach does not fully leverage the temporal variability information available in the radar fields and a long short-term memory (LSTM; Hochreiter and Schmidhuber 1997; Yang et al. 2020) model that can account for the temporal patterns in the data will be explored in future work.

b. CNN model setup

There were several preprocessing steps applied to the input data to improve the model performance, including binning the data samples by precipitation amount (full list of data bins shown in Table 3) and balancing the number of training samples per precipitation bin. The number of samples in the 4–6-mm size bin was used as the uniform sample size drawn from each bin. This intermediate bin was chosen for the sample size to avoid overfitting to more numerous samples in the lowest precipitation bins, while still maintaining an adequate number of samples for the model to learn from. A rotation was also applied to the input grids through different angles. This is a common data augmentation technique used within CNN models to increase the number of samples available to the model. The training samples were also shuffled to avoid any temporal correlations in the input data. With wide ranging values in the different input variables, scaling was needed to avoid having the model only focus on those variables with high magnitude variances (Lagerquist et al. 2020). The scaler operation applied here transforms each input variable field to take on a value between 0 and 1, based on where the raw
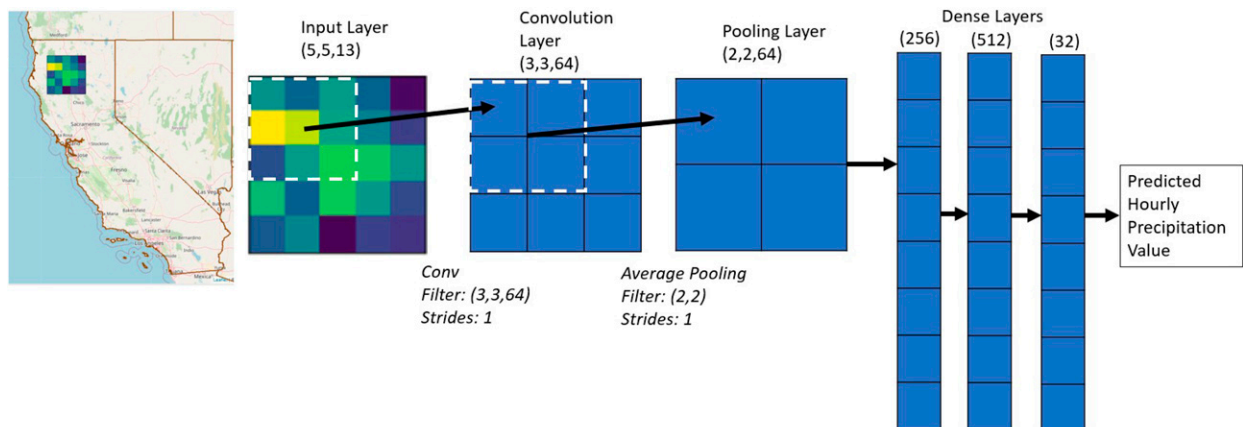

FIG. 3. Example schematic of how an input field progresses through the different layers of the CNN model. The white dashed box represents the convolutional or pooling filters being applied to the respective layers in a moving window. The filter size and stride length are indicated to the bottom and the dimension information for each model layer is shown to the top. At the far left, the location of the training example in Northern California is shown (magnified to be larger than the actual 5 km × 5 km size for visualization purposes). A simulated QPE field is generated by applying this process over every grid point of the western U.S. domain.
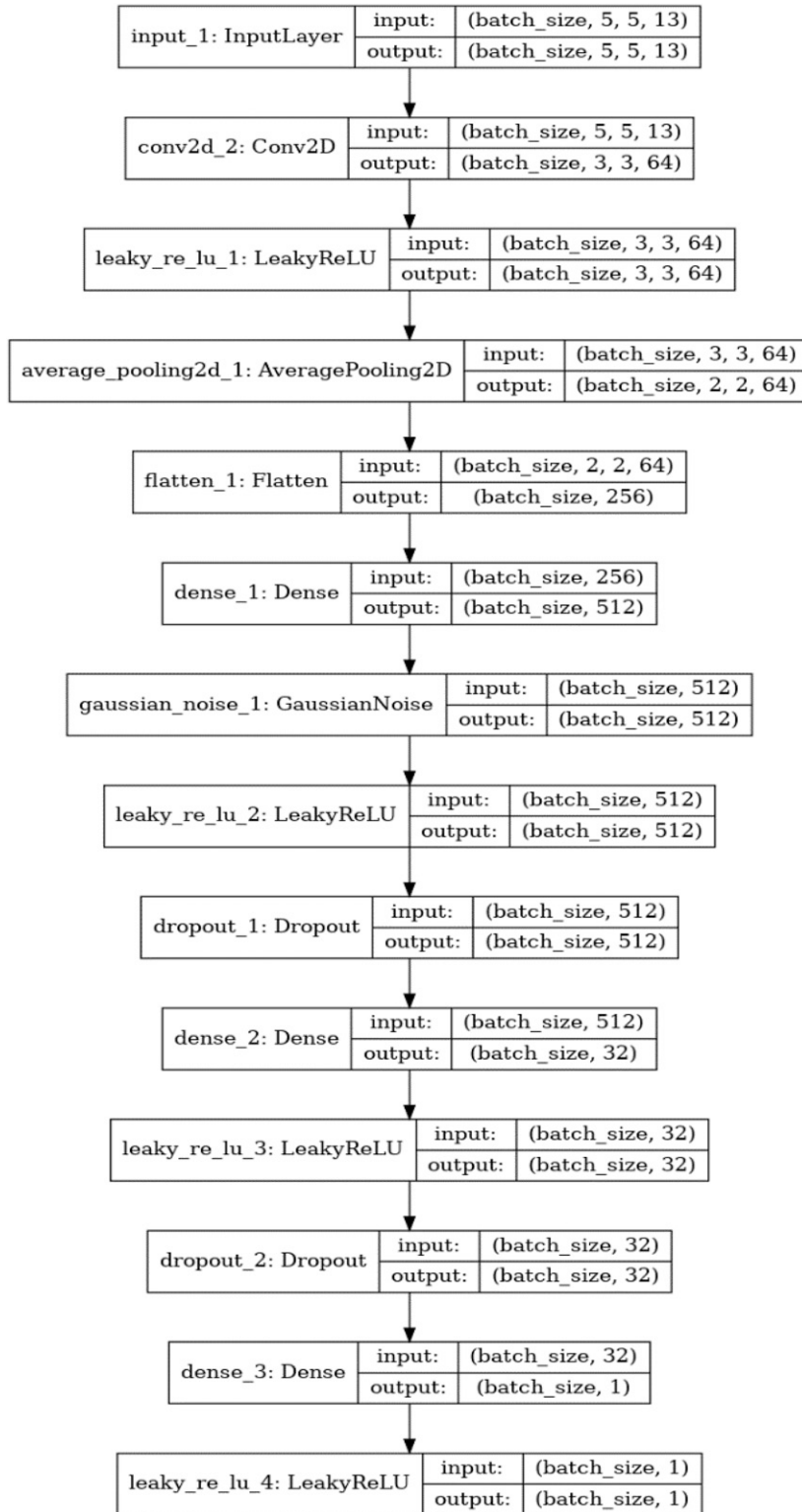
input_1: InputLayer	input:	(batch_size, 5, 5, 13)
	output:	(batch_size, 5, 5, 13)

conv2d_2: Conv2D	input:	(batch_size, 5, 5, 13)
	output:	(batch_size, 3, 3, 64)

leaky_re_lu_1: LeakyReLU	input:	(batch_size, 3, 3, 64)
	output:	(batch_size, 3, 3, 64)

average_pooling2d_1: AveragePooling2D	input:	(batch_size, 3, 3, 64)
	output:	(batch_size, 2, 2, 64)

flatten_1: Flatten	input:	(batch_size, 2, 2, 64)
	output:	(batch_size, 256)

dense_1: Dense	input:	(batch_size, 256)
	output:	(batch_size, 512)

gaussian_noise_1: GaussianNoise	input:	(batch_size, 512)
	output:	(batch_size, 512)

leaky_re_lu_2: LeakyReLU	input:	(batch_size, 512)
	output:	(batch_size, 512)

dropout_1: Dropout	input:	(batch_size, 512)
	output:	(batch_size, 512)

dense_2: Dense	input:	(batch_size, 512)
	output:	(batch_size, 32)

leaky_re_lu_3: LeakyReLU	input:	(batch_size, 32)
	output:	(batch_size, 32)

dropout_2: Dropout	input:	(batch_size, 32)
	output:	(batch_size, 32)

dense_3: Dense	input:	(batch_size, 32)
	output:	(batch_size, 1)

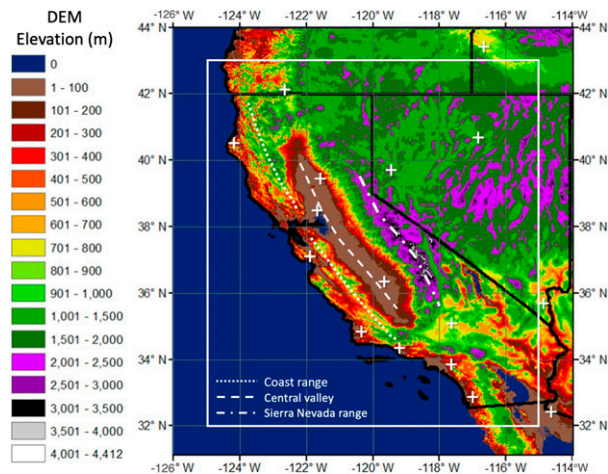| leaky_re_lu_4: LeakyReLU	input:	(batch_size, 1)
	output:	(batch_size, 1)

FIG. 4. CNN model architecture.

FIG. 5. The analysis domain (white box) of the current study and a digital elevation model (DEM). The white dotted line indicates the California Coast Ranges, the dashed line indicates the Central Valley, and the dot–dashed line indicates the Sierra Nevada range. The white "+" symbols represent radar sites.

value is located relative to the minimum and maximum values of that variable (Table 4) within the training dataset.

The CNN model used here is an example of a deep learning scheme with several model layers including a convolution layer for delineating spatial features, a pooling layer to down-sample the gridded information for faster processing, and dense layers (i.e., standard neural network layers) to come up with the predicted value (Lagerquist et al. 2020). The schematic in Fig. 3 outlines how a training example over Northern California would progress through the model layers, starting with the thirteen 5 km × 5 km input variable grids (with 1-km MRMS pixel size) and ending with the predicted precipitation value at the center point of that 5 km × 5 km grid.

The CNN starts with a convolutional layer that applies 64 distinct 3 km × 3 km convolutional filters to the 5 km × 5 km input variable grids. The convolutional filters are applied in a moving window centered over each of the nonboundary grid points of the input grid. The resultant output from the matrix operations in the convolutional layer is a 3 km × 3 km feature map for all 64 filters. A leaky rectified linear unit (ReLU) activation function (Maas et al. 2013) is applied after the convolutional layer to avoid the vanishing gradients problem, to allow the model to learn nonlinear relationships, and to avoid losing information from neurons with values below 0. A pooling layer downscales the 3 km × 3 km feature maps to 2 km × 2 km fields by performing an average value operation over the four 2 × 2 boxes of the 3 × 3 grids. The resultant 64 channels of 2 km × 2 km grids are then flattened into a 256 × 1 vector that is sent through three successive 1D neural network (dense) layers with a single predicted precipitation value as output from the final layer. The predicted value is compared to the quality-controlled rain gauge (or gridded truth) precipitation value at the center of the 5 km × 5 km input grid, with the goal of minimizing the mean absolute error
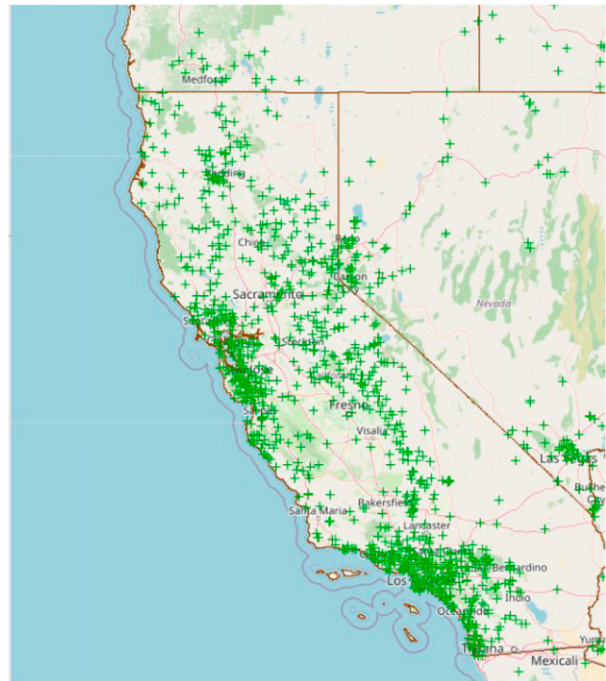

FIG. 6. Distribution of hourly MADIS gauges (green "+" symbols) used in the training dataset from 1300 UTC 18 Feb 2019.

(MAE) between those values after many data samples and training iterations. The model parameters are adjusted via backpropagation after each batch of training data, with the gradient descent method used to minimize the MAE value. A validation step is applied to assess the model performance on an independent dataset during the training process. After this training and validation process, the CNN model is used to make inferences at each grid point of the western U.S. domain and generate a "simulated" CNN QPE for selected cases independent of the training and validation datasets.

The model is trained and tested using Keras within Python (Chollet et al. 2020). For the trained model presented here, the Adam optimizer was used along with a batch size of 1000 samples and 10 epochs. To help the model generalize to unseen data and avoid overfitting, a Gaussian noise layer is applied after the first dense layer and a dropout procedure (Hinton et al. 2012) is applied between some of the dense layers. The full CNN model setup can be seen in Fig. 4.

c. Experimental setup

The trained CNN model was applied to several cases in the cool season of 2019 and 2020 as well as all precipitation days during the year of 2021 (the "simulation data" in Table 2) over a regional domain in the western United States. The domain (Fig. 5) has a resolution of 0.01° latitude × 0.01° longitude (~1 km × 1 km) and covers California where the coastal and inland mountains often result in uplift and orographic enhancement of precipitation. The typical distribution of hourly
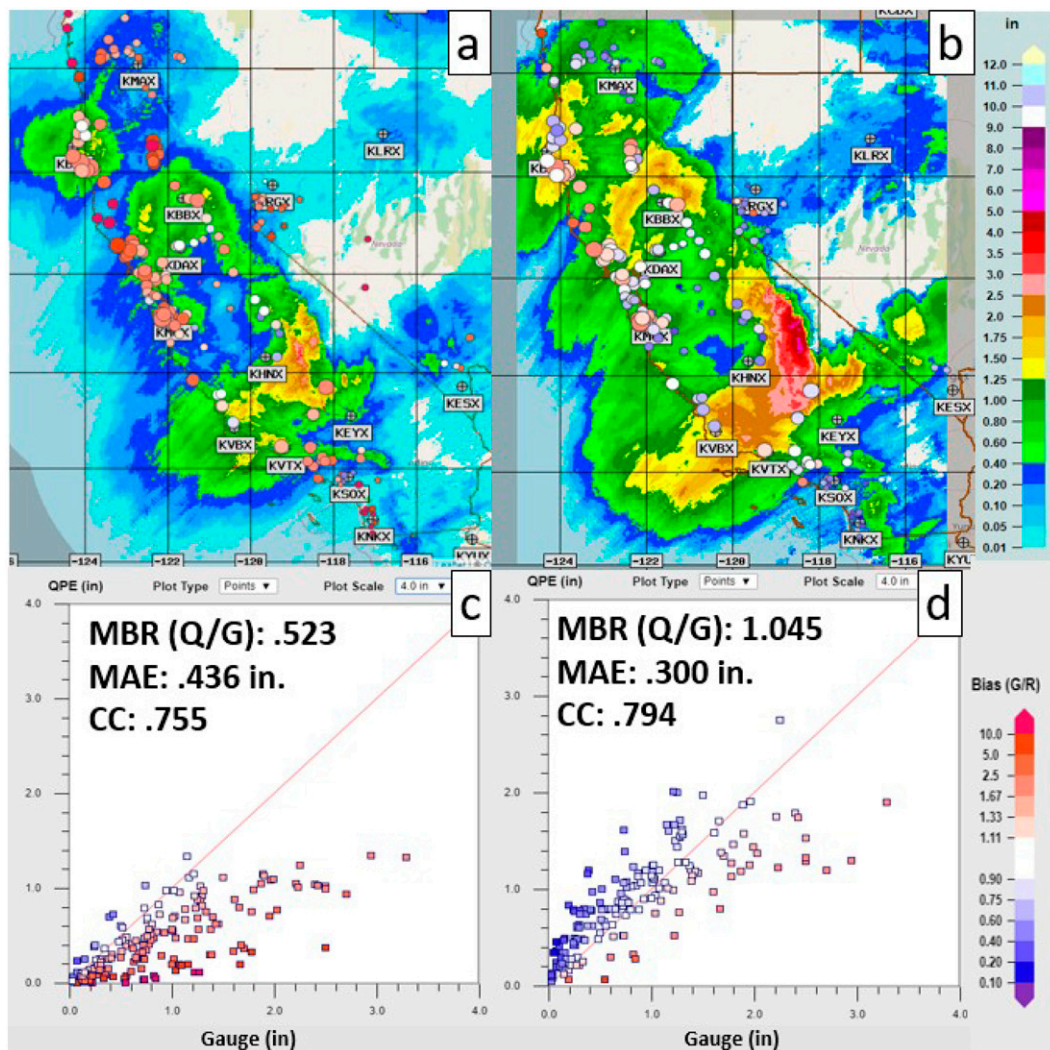
FIG. 7. (a) The MRMS Q3EVAP and (b) CNN model 24-h QPEs ending at 1500 UTC 6 Mar 2019 with validation against CoCoRaHS gauges. The solid circles in (a) and (b) indicate gauge locations. The size of the circles is proportional to the gauge amounts, and the color represents the QPE/gauge ratio. Cool (blue) colors of the circles indicate QPE overestimation relative to gauges while warm (red) colors indicate underestimation. Scatterplots of the (c) Q3EVAP and (d) CNN QPEs vs gauges.

gauges used in the training process within the domain is shown in Fig. 6 (~1200 gauges on average for a given hour). The selected simulation days in 2019 and 2020 involved atmospheric rivers (AR) with plumes of moist air over the Pacific Ocean feeding into the region. These types of events account for the majority of annual precipitation occurring in this area (Bytheway et al. 2020), and precise precipitation quantification within these events is essential for water resource management.

For each simulation day, the CNN model was tested by generating 24-h accumulated QPE simulations and then validating against the Community Collaborative Rain Hail and Snow Network (CoCoRaHS; https://www.cocorahs.org/; Cifelli et al. 2005) gauge accumulations. As described in the model section above, the CNN creates a simulated precipitation value by taking a set of thirteen 5 km × 5 km input fields and moving through the model layers to yield an estimated hourly precipitation value at the center point of the 5 km × 5 km grid. This procedure is then repeated over each grid point of the specified domain shown in Fig. 5 to generate a 2D hourly QPE field. The simulated hourly fields are generated for 24 consecutive hours with a resultant 24-h accumulated CNN precipitation field. Statistical evaluation was performed and included a comparison of the mean bias ratio (MBR), correlation coefficient (CC), MAE, and fractional MAE (fMAE) from the CNN method and from the operational MRMS radar QPE. The definitions of the MBR, CC, MAE, and fMAE are as follows:

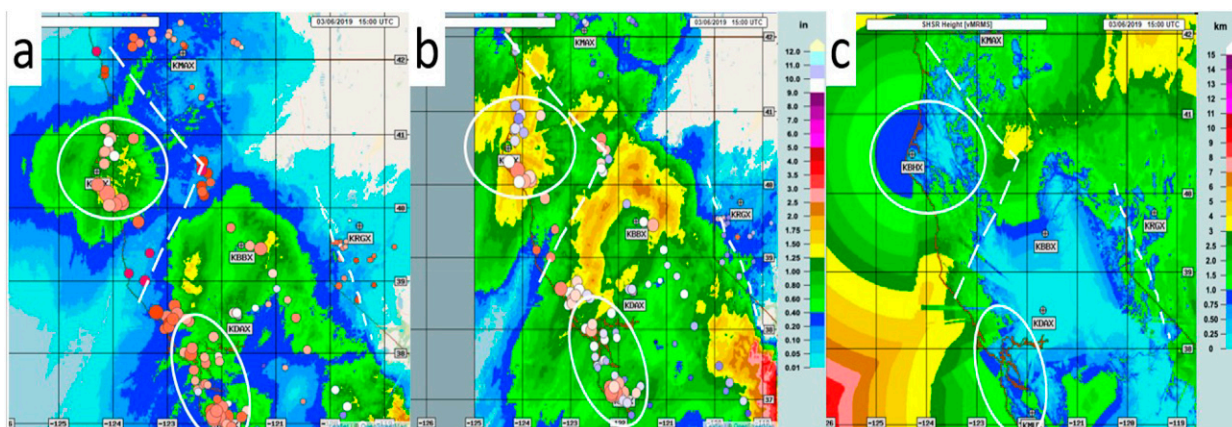$$\text{MBR} = \overline{Q}/\overline{G}, \tag{1}$$

FIG. 8. Enlarged view of the 24-h QPEs from (a) Q3EVAP and (b) CNN for the 6 Mar 2019 case over Northern California. (c) The SHSRH field valid at 0300 UTC 6 Mar 2019.

where

$$\overline{Q} = \frac{\sum_{i=1}^{N} Q_i}{N}, \tag{1a}$$

$$\overline{G} = \frac{\sum_{i=1}^{N} G_i}{N}, \tag{1b}$$

$$CC = \frac{\sum_{i=1}^{N}(Q_i - \overline{Q})(G_i - \overline{G})}{\sqrt{\sum_{i=1}^{N}(Q_i - \overline{Q})^2 \sum_{i=1}^{N}(G_i - \overline{G})^2}}, \tag{2}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|Q_i - G_i|, \quad \text{and} \tag{3}$$

$$fMAE = MAE/\overline{G}, \tag{4}$$

where N is the total number of QPE-gauge pairs for the given dataset, Q_i is the CNN (or MRMS) 24-h QPE, and G_i is the gauge-observed 24-h rainfall at the ith gauge station.

d. Masking procedure

Initial experiments showed that when certain input variables were included, the CNN model precipitation simulations would have artifacts present. The artifacts tended to be uniform, spurious light precipitation covering large portions of the domain with occasional ring-shaped features. The circular artifacts are likely associated with the arc-shaped features of the SHSRH (Fig. 1b) and RQI (Fig. 1d) fields near the edge of the radar coverage area. With the overshooting beam effects far from the radar, the model cannot rely on the radar
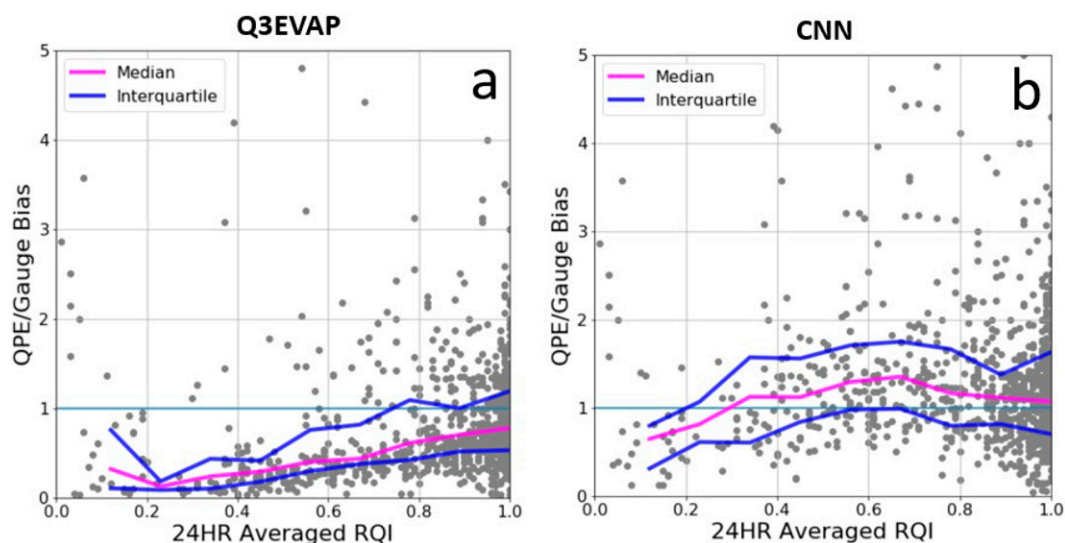

FIG. 9. The bias ratio of the 24-h (a) Q3EVAP and (b) CNN QPE against individual CoCoRaHS gauges as a function of the 24-h average RQI for the four atmospheric river precipitation events.
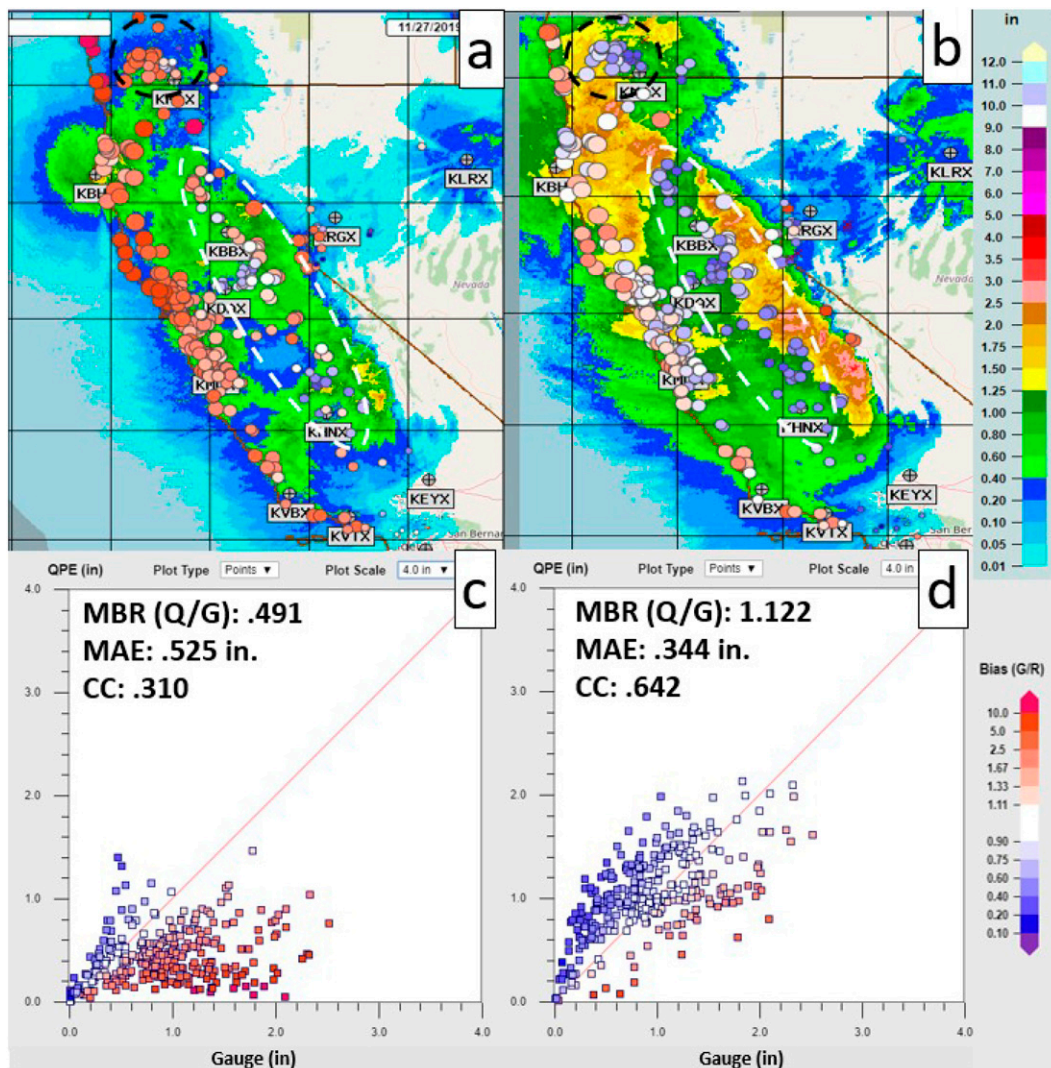
FIG. 10. As in Fig. 7, but for 1500 UTC 27 Nov 2019. The white and black dashed circles indicate areas of different Q3EVAP and CNN QPE performances, and detailed explanations can be found in the text.

reflectivity fields to accurately delineate areas where precipitation is occurring. Relying on variables with more uniform coverage, such as SHSRH and RQI, where the value of the variable is not well correlated to precipitation occurrence likely makes the model more susceptible to errors in precipitation placement. The RQI and SHSRH variables may provide useful information regarding radar sampling and related QPE biases, so to mitigate the false precipitation while retaining the benefits of these variables, a precipitation mask was applied to the CNN model output. The MRMS operational dual-pol radar synthetic hourly QPE field with evaporation correction ("Q3EVAP"; Zhang et al. 2020) was used to mask the CNN QPE predictions, such that the CNN values were set to 0 anywhere the hourly Q3EVAP field had no precipitation. Therefore, while the current CNN QPE does provide adjusted precipitation amounts relative to the Q3EVAP product, the precipitation coverage area remains unchanged from

Q3EVAP. Precipitation coverage improvements would likely come from the use of additional data sources such as satellite observations.

3. Results

After completion of the training process, the CNN model was applied to the 2019 and 2020 atmospheric river events (Table 2) and hourly QPE fields were generated over the study domain (Fig. 5). Using the hourly CNN QPE fields, 24-h accumulations were then calculated and compared to the CoCoRaHS gauge observations. The CNN model performance was also compared to the performance of Q3EVAP, which is a physically based scheme using explicit relationships between radar variables and the precipitation rate. Finally, the CNN model was tested on all precipitation days in 2021.
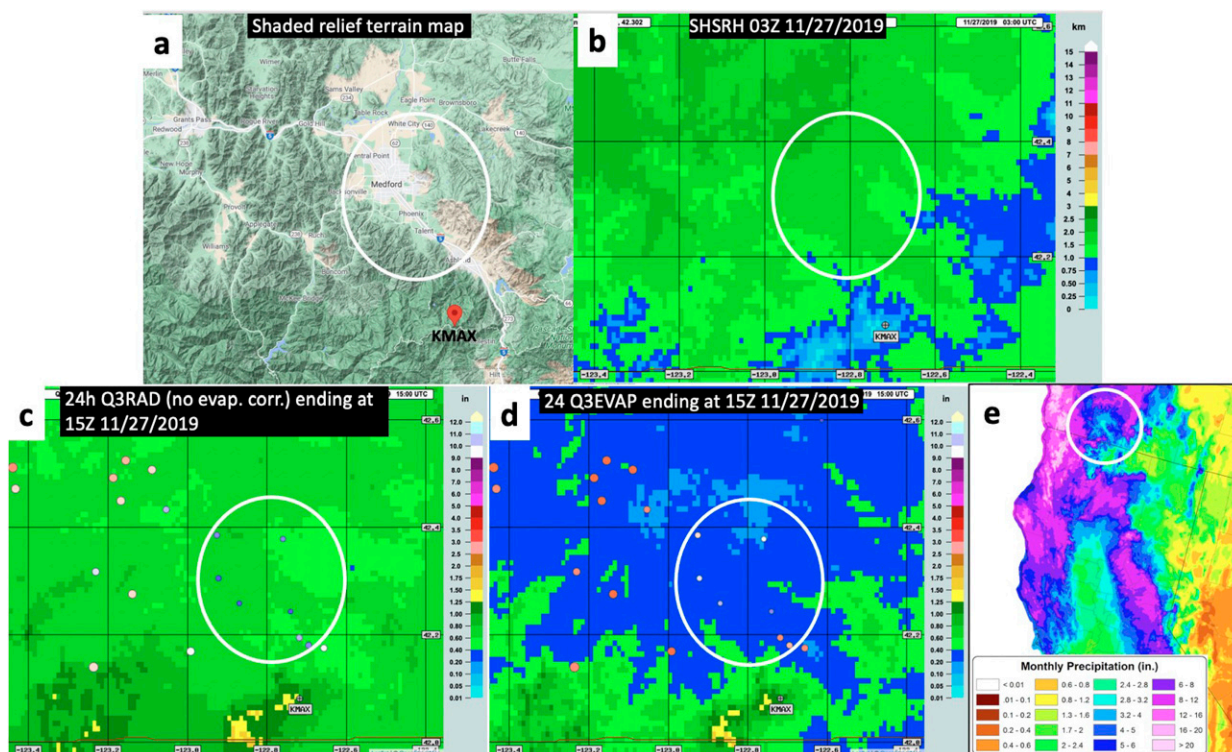
FIG. 11. (a) A shaded relief terrain map and (b) SHSRH field over the Rouge Valley of southern Oregon (circled in white). The MRMS 24-h radar-only QPE valid at 15 UTC 27 Nov 2019 (c) without ("Q3RAD") and (d) with ("Q3EVAP") the evaporation correction. The circles in (c) and (d) represent gauges as shown in Fig. 7a. (e) The PRISM 30-yr normal precipitation for November.

a. Case of 6 March 2019

A substantial atmospheric river event affected California from 5 to 7 March 2019 with moderate integrated vapor transport (IVT) values and widespread precipitation accumulations of 2–4 in. (https://cw3e.ucsd.edu/wp-content/uploads/2019/03/20190305_AR_Quicklook.pdf). There were substantial underestimations (∼48%) from the MRMS Q3EVAP product (Figs. 7a,c) with a domain MBR of 0.52. The underestimation bias from the radar-based product is common in this area due to terrain effects on radar observations (beam blockage, overshooting, etc.). Also, the variability of the precipitation processes within an orographically enhanced regime makes it difficult to apply the generalized radar-based rainfall rate equations to these scenarios. The CNN QPE for this same period shows much better agreement with gauges (Fig. 7d), reducing MAE from 0.389 to 0.276 in. (∼31% reduction), increasing CC from 0.755 to 0.794, and nearly eliminating the bias (MBR = 1.045).

A notable finding from this event was the ability of the CNN QPE to reduce the range-dependent bias in the radar QPE. Figure 7a shows substantial underestimation over Northern California (pink to red circles) from the MRMS Q3EVAP product. The areas of the most severe underestimation (white dashed line, Fig. 8a) line up mostly with the areas of poor low-level radar coverage seen in the SHSRH field (white dashed line, Fig. 8c), which is the bottom height of the lowest radar beam with no severe blockages (i.e., ≤50%). This underestimation is associated with the radar beam

overshooting the melting layer and therefore the radar observations are not representative of the hydrometeor phase and size distributions at the surface (Zhang et al. 2012a). The CNN product (Fig. 8b) shows better continuity across those areas (white dashed lines, Fig. 8b) with higher values that more accurately match gauge observations than the Q3EVAP radar-based algorithm. Similar improvements were observed in the other three AR cases. Figure 9 shows the bias ratios of the 24-h Q3EVAP and CNN QPEs over individual CoCoRaHS gauges as a function of the 24-h average RQI for the four cases examined in this work. The median bias of Q3EVAP (purple line, Fig. 9a) shows a persistent underestimation that worsens with decreasing RQI, generally associated with increasing distance from the radar. The median bias of CNN QPE (purple line, Fig. 9b) was closer to 1.0 than the Q3EVAP and was largely independent of the RQI until when RQI fell below 0.3, where beam overshooting of precipitation areas is more prevalent. These results demonstrated the CNN model's capability to mitigate range-dependent QPE biases, probably learned from the RQI and SHSRH fields. Further, the CNN product reduced the moderate underestimation bias near the radars (white circles, Fig. 8a) in Q3EVAP where overshooting is less of a problem.

b. Case of 27 November 2019

Another atmospheric river event impacted most of California from 26 to 27 November 2019 (https://cw3e.ucsd.edu/wp-content/uploads/2019/11/27Nov19_Outlook/27Nov19_Outlook.pdf). The
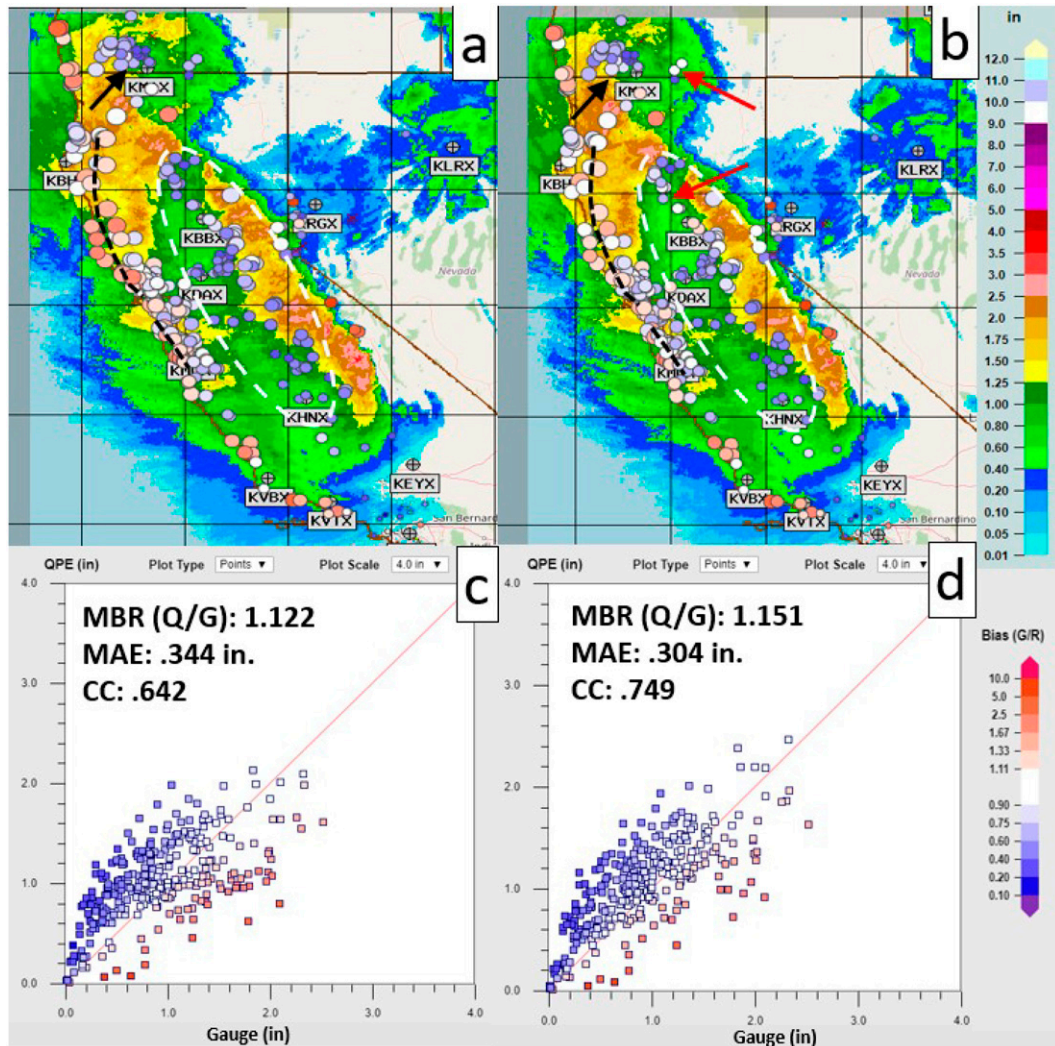
FIG. 12. As in Fig. 10, but for the (a),(c) CNN model QPE and (b),(d) CNN_PRISM model QPE. The circles and arrows indicate areas of different Q3EVAP and CNN QPE performances, and detailed explanations can be found in the text.

24-h Q3EVAP ending at 1500 UTC 27 November 2019 (Fig. 10a) showed a large underestimation bias (~51%), especially along the coastal mountain ranges in Northern California (Figs. 10a,c). The 24-h CNN QPE (Figs. 10b,d) for the same time period eliminated the severe underestimation along the coastal mountains while introducing overestimation in southern Oregon (dashed black circle, Fig. 10b) and in the Central Valley and the foothills of the Sierra Nevada range (dashed white circle, Fig. 10b). This resulted in an overall wet bias of ~12% in the CNN QPE. Nevertheless, the CNN QPE versus gauge values lined up better with the 1:1 line (Fig. 10d) and the MAE dropped from 0.525 in. for Q3EVAP to 0.344 in. for the CNN QPE, while the CC increased from 0.310 to 0.642. A closer look at the overestimation in southern Oregon indicated a possible evaporation effect that was not captured in the CNN model. Figure 11a shows a terrain map centered around the Rouge Valley, a region that is 1.5–2.0 km below the lowest tilt (Fig. 11b) of the nearest radar,

KMAX (Medford, Oregon). Small precipitation particles may be partially or even completely evaporated/sublimated before reaching the ground when the environment near the surface is sufficiently dry. The physically based radar QPE without the evaporation correction (Fig. 11c) showed a similar overestimation bias (white circle, Fig. 11c) as in the CNN QPE (dashed black circle, Fig. 10b) in the area, while Q3EVAP had no wet bias (white circle, Fig. 11d). Further, the PRISM 30-yr precipitation climatology for November (Fig. 11e) shows a local minimum and indicates a persistent deficit of precipitation in the area. Toward reducing this local wet bias, the PRISM monthly precipitation climatology field was added to the input variables and a new CNN model ("CNN-PRISM") was trained and applied to this case. However, the resultant QPE (Fig. 12) did not have a significant reduction of precipitation in the Rouge Valley area (black arrows, Fig. 12), most likely due to the small area and
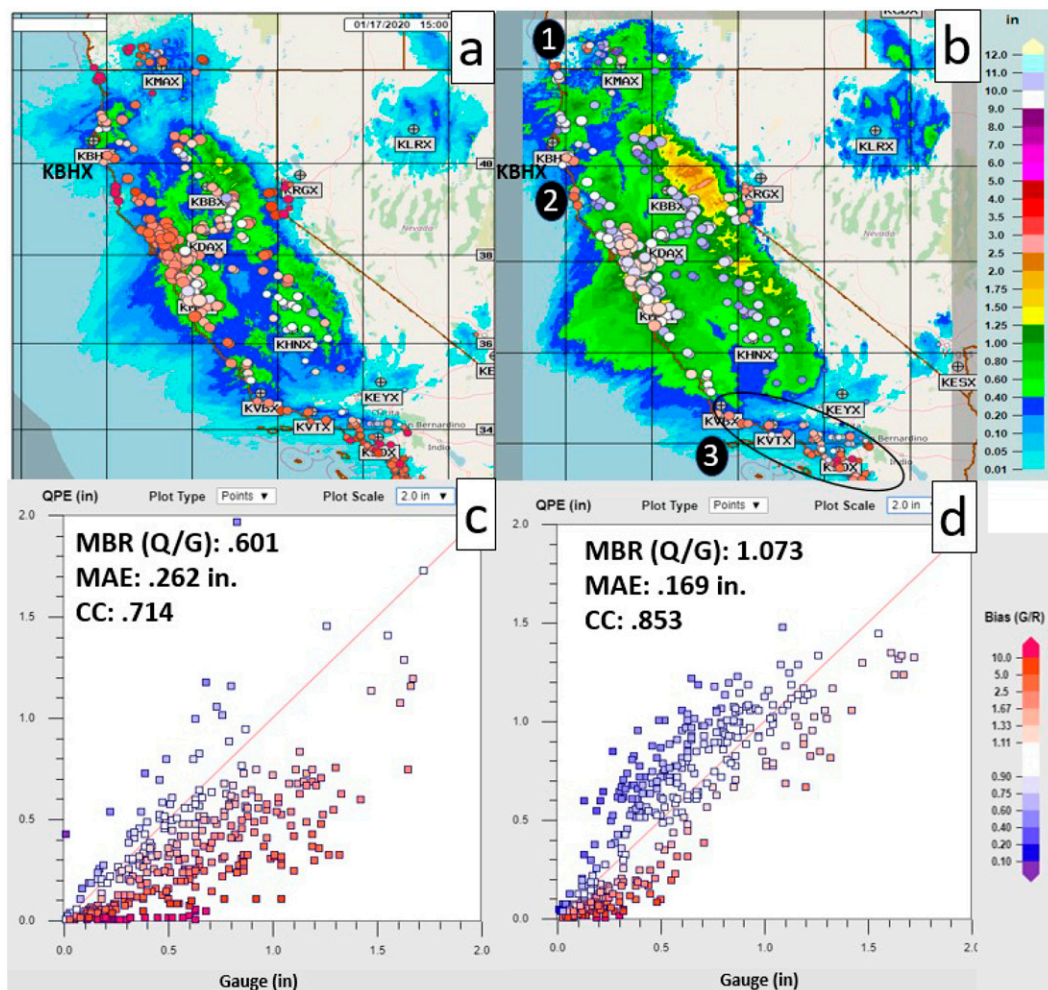
FIG. 13. As in Fig. 7, but for 1500 UTC 17 Jan 2020. The black circle and numbers indicate areas of specific Q3EVAP and CNN QPE performances, and detailed explanations can be found in the text.

limited data samples for the training. This result highlights the challenge of the machine learning technique in capturing relatively "rare" and localized events. Meanwhile, the CNN-PRISM QPE reduced the underestimation in the CNN QPE along the coastal mountain ranges (black dashed lines, Fig. 12) and the overestimation in the northern end of the Sierra Nevada range and to the east of the KMAX radar (red arrows, Fig. 12). The improvements yielded a lower MAE (0.304 vs 0.344 in.) and higher CC (0.749 vs 0.642) than the original CNN model (Fig. 12d vs Fig. 12c).

While the CNN models reduced the severe dry biases in Q3EVAP, both of them introduced overestimation in the Central Valley (white dashed circles, Figs. 12a,b). This was likely due to different drop size distribution (DSD) in the valleys and in the mountain ranges (Kingsmill et al. 2006). DSD differences near the surface can be large given the complex interactions between the atmospheric flow and the topography in the region. Such local variations may have not been fully captured in the current input variables (Table 1) and warrant future studies using additional environmental (e.g., moisture

and wind) information, geographic variables, and more data samples.

c. Case of 17 January 2020

The third case examined was the atmospheric river event occurring from 16 to 17 January 2020 (http://cw3e.ucsd.edu/wp-content/uploads/2020/01/Jan152020_Outlook/15Jan20_Outlook.pdf). Similar to the previous cases, the MRMS Q3EVAP product had a large underestimate bias (~40%) (Figs. 13a,c) in this orographically enhanced precipitation regime. The CNN model (Figs. 13b,d) nearly removed the bias and had only ~7% overestimation. It also reduced the MAE from 0.262 to 0.169 in. (a 35% improvement) and increased the CC from 0.714 to 0.853 (~19% improvement). The remaining underestimation mainly came from three areas (Fig. 13): 1) the southwestern corner of Oregon, 2) along the coastline south of Eureka, California (KBHX), and 3) Southern California.

For areas 1 and 2, the radar beam may have completely overshot the precipitation clouds and resulted in missing reflectivities. The model is heavily reliant on the reflectivity
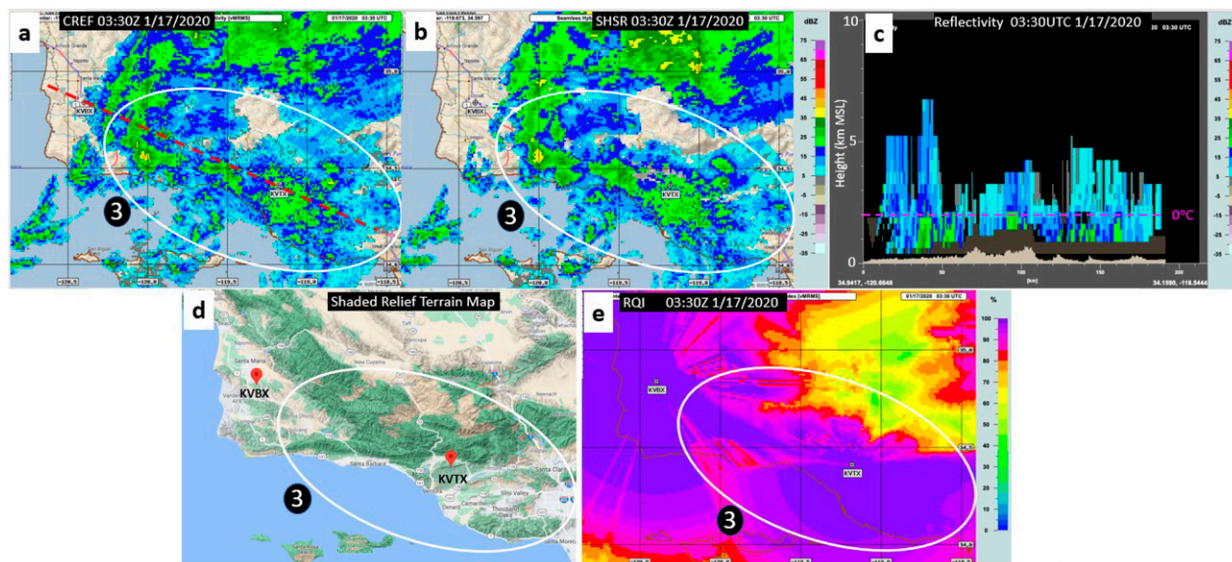
FIG. 14. (a) CREF, (b) SHSR, (c) a vertical cross section of reflectivity along the red dashed line in (a) with freezing-level height shown via the pink dashed line, (d) a shaded relief terrain map, and (e) RQI fields at 0330 UTC 17 Jan 2020. The white circles indicate the same area 3 as in Fig. 13.

variables, especially SHSR, and the performance is significantly reduced when this reflectivity information is missing. Also, the current precipitation masking based on the radar QPE results in removal of CNN QPE in areas where the radar beam is completely above the cloud top. Additional precipitation data sources, such as satellite QPEs and/or QPFs from NWP models are needed to fill in the precipitation information under these situations and the precipitation masking process will need to be refined to allow those data sources to fill in the radar gaps.

A closer look at the underestimation in area 3 (Fig. 14) indicates shallow, warm rain precipitation processes occurring in the lowest levels. Both the CREF (Fig. 14a) and SHSR (Fig. 14b) at 0330 UTC 17 January 2020 showed weak to moderate reflectivities. A vertical cross section of reflectivity (Fig. 14c) from the area showed very shallow precipitation cores (e.g., reflectivity > 20 dBZ) that were largely below the freezing level (pink dashed line, Fig. 14c), indicating warm rain processes. Xu et al. (2008) showed that warm rain processes often involve coalescence growth near the surface, resulting in an increasing reflectivity with decreasing height. The complex terrain (Fig. 14d) in area 3 could also contribute to enhanced precipitation near and just above the surface. However, the lowest radar observations in the area were 0.5 to 1.0 km above the surface (Fig. 14c) and had severe blockages (see RQI, Fig. 14e). The warm/orographic rain enhancements near the surface were likely not well represented by any of the reflectivity fields and thus resulted in underestimation. The CNN-PRISM QPE (not shown) yielded a very similar result as in the original CNN model, with an MBR of 1.105 (vs 1.083), a CC of 0.853 (vs 0.853), and a MAE of 0.173 in. (vs 0.169 in.). Incorporating QPFs from NWP models along with environmental (e.g., moisture, temperature, and wind) and

orographic variables may help capture the microphysical and dynamical processes involved in these warm/orographic rainfall regimes and will be explored in a future study.

d. Case of 13 March 2020

The final case study involved a heavy rainfall event in Southern California from 12 to 13 March 2020 (http://cw3e. ucsd.edu/wp-content/uploads/2020/03/17Mar2020_Summary/ 17Mar2020_Summary.pdf). The Q3EVAP product (Figs. 15a,c) underestimated by ~41% relative to 24-h CoCoRaHS gauge accumulations. The underestimation is likely related to the known issues with radar blockages, overshooting and inadequate radar rainfall rate equations for the orographically enhanced rainfall regimes. The CNN product nearly eliminated the overall bias (~2% underestimation) and reduced the MAE by ~32% from 0.567 to 0.388 in. (Fig. 15) relative to Q3EVAP. Further, Q3EVAP had a gap near southeastern California and southwestern Arizona (area 1, Fig. 15a). The gap pattern matches an area of minimum RQI (Fig. 16) and appears to be an artifact of poor radar coverage. The CNN QPE (area 1, Fig. 15b) filled the gap and provided a more continuous precipitation distribution over the region. There is still a notable spread about the 1:1 line from the CNN product (Fig. 15d), with some overestimation introduced in the northwest portion of the precipitation area (area 2, Fig. 15b) and underestimation in the southeast part of the precipitation area (area 3, Fig. 15b). The overestimation in area 2 was probably caused by brightband contamination. Figure 17 shows the CREF and SHSR fields and a vertical cross section of reflectivity at 2330 UTC 12 March 2020. The inflated reflectivities in CREF (Fig. 17a) were clearly associated with a bright band in the vertical cross section of reflectivity (Fig. 17d). While the bright band was corrected in SHSR (Fig. 17b), the effect
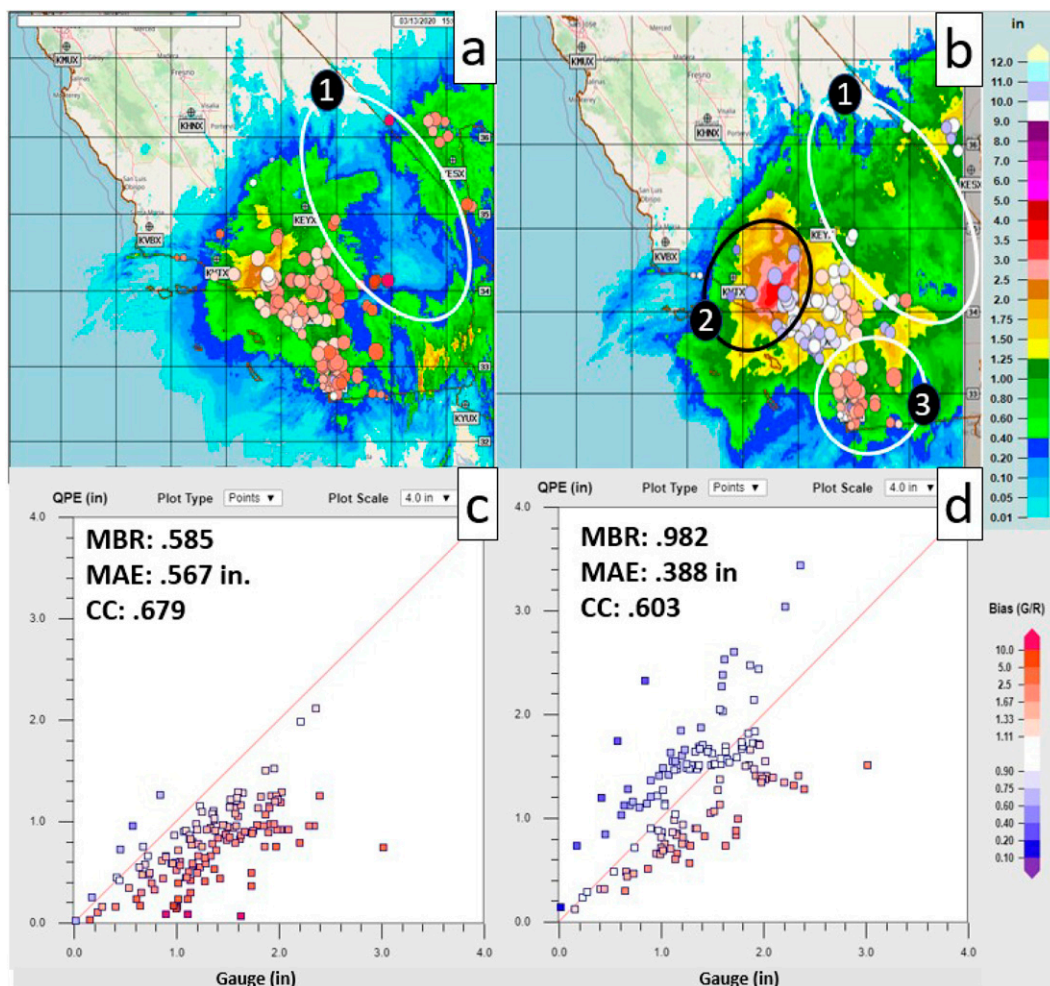
FIG. 15. As in Fig. 7, but for 1500 UTC 13 Mar 2020. The white and black circles and numbers indicate areas of specific Q3EVAP and CNN QPE performances, and detailed explanations can be found in the text.

was not corrected in other input fields such as CREF (Fig. 17a) and RALA (Fig. 17c). The CNN model may not have had sufficient gauge data samples in brightband areas to fully account for the inflated effect. More training data from brightband areas are needed to avoid such overestimation in the CNN model.

The underestimation in area 3 showed a similar warm rain situation as in area 3 of the 17 January 2020 case (Fig. 14). There was no brightband signature in the CREF (Fig. 18a) or in the vertical cross section (Fig. 18c). While the SHSR field (Fig. 18b) had similar intensities as in area 2 (Fig. 17b), the vertical cross section (Fig. 18c) showed relatively shallow precipitation cores (e.g., reflectivity > 20 dBZ) that were mostly below the freezing level. The lowest radar observations were 0.5–1.5 km above the ground (Fig. 18c) and had severe blockages (see RQI; Fig. 16) due to the complex terrain (Fig. 18d). With poor low-level coverage in this area, the warm/orographic rain enhancement is not captured by the input reflectivity fields and additional variable data would likely be needed to alleviate the underestimation bias.

e. January–December 2021

As an extended evaluation, the CNN model was run for all the days (a total of 112) in 2021 that had at least 20 nonzero gauges in the domain. The 24-h CNN QPE performance relative to CoCoRaHS gauges was compared to the 24-h Q3EVAP product over this time. The time series plots focused on the winter months (Fig. 19) show that the CNN model outperforms Q3EVAP during these cool season months (approximately October to March), with less underestimation bias (Fig. 19a) and lower fMAE (Fig. 19b) than Q3EVAP. Many of the precipitation days during the cool season involved widespread, heavy precipitation events as indicated by the high domain average gauge values (gray "×" symbols in Fig. 19) and a large number of nonzero gauge observations (gray "−" symbols in Fig. 19). The CNN QPE showed consistent improvement over the Q3EVAP for these widespread and relatively heavy rain events, with a 20%–50% reduction in the bias and a 10%–20% reduction in the MAE for days with at least 50 nonzero gauges and average gauge amount exceeding 0.20 in. The CNN model performance
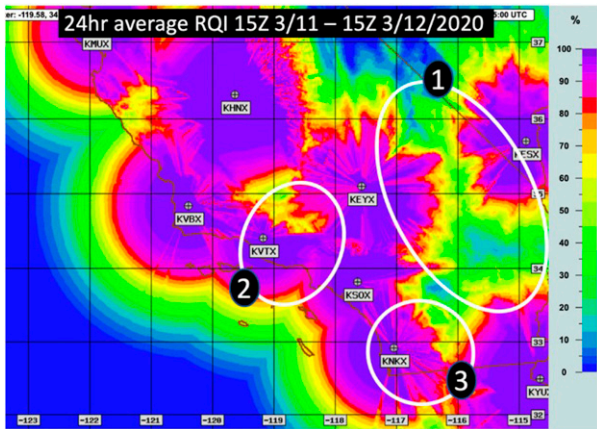
FIG. 16. The 24-h average RQI from 1500 UTC 11 Mar to 1500 UTC 12 Mar 2020. The circles and numbers indicate the same areas as in Fig. 15.

degraded CNN performance in the warm season may be largely attributed to the choice of the training dataset, which consisted of only cool season months. Further, the widespread heavy precipitation events in the training data may have dominated the CNN model formulation due to the greater contributions of these data samples to the total MAE, which is the parameter that the CNN model attempts to minimize. Future work will explore whether inclusion of warm season months in the training data and adjusting the training criteria could improve the model performance for lighter and/or sporadic precipitation regimes.

4. Summary

In this work, a CNN model was developed and applied to the precipitation estimation problem in the complex terrain of the western CONUS. The model was optimized through experimentation with different input variable combinations, preprocessing techniques, and hyperparameter settings. The model was evaluated for several heavy rainfall events associated with landfalling atmospheric rivers in the western CONUS. Results from these case studies showed that 24-h precipitation estimates from the machine learning model statistically outperformed physically based radar QPE. The CNN model consistently corrected for the large underestimates observed in the radar QPE,

drops off considerably during the summer months when the events involve more scattered, convective precipitation (Fig. 20). The CNN shows a higher overestimation bias (Fig. 20a) and increased fMAE (Fig. 20b) compared to the Q3EVAP product for many of these lighter or sporadic precipitation events. The
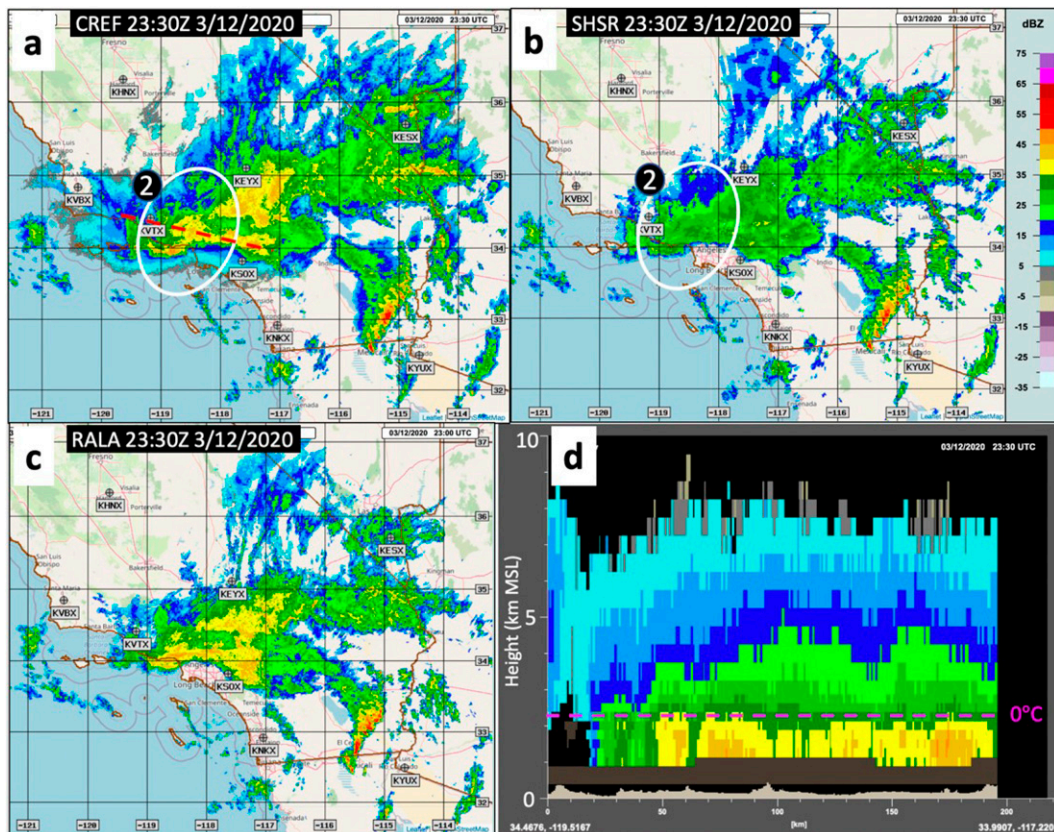

FIG. 17. (a) CREF, (b) SHSR, and (c) RALA fields at 2330 UTC 12 Mar 2020. (d) A vertical cross section of the reflectivity at the same time, which was taken along the red dashed line in (a). The pink dashed line represents the freezing level.
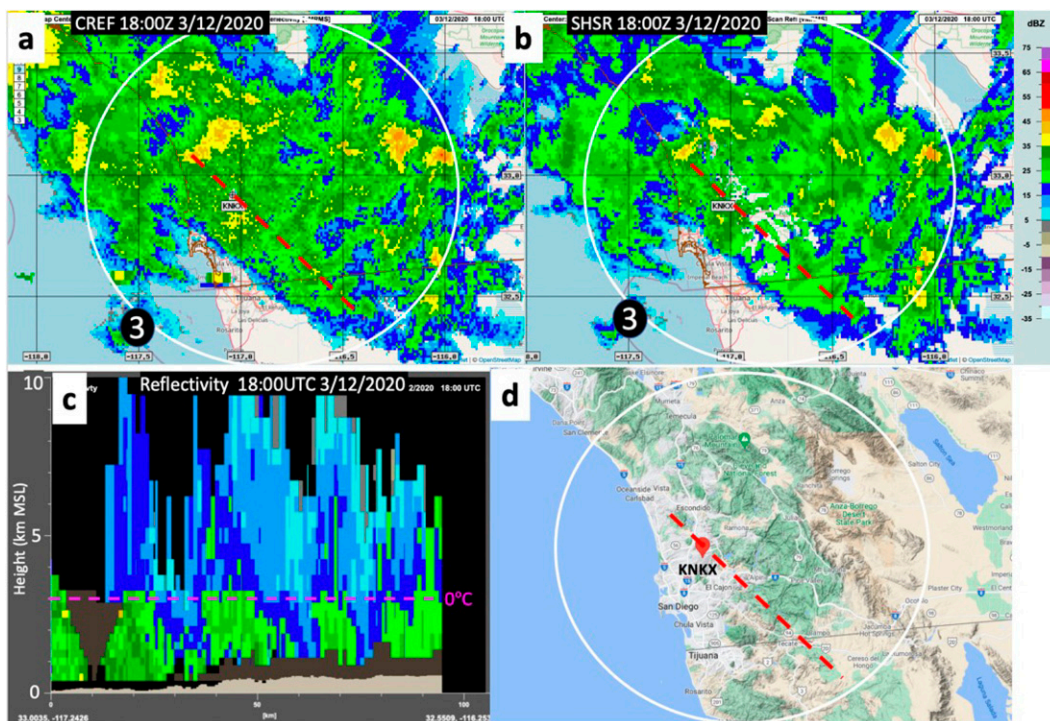
FIG. 18. (a) CREF and (b) SHSR fields at 1800 UTC 12 Mar 2020. (c) A vertical cross section of the reflectivity at the same time, which was taken along the red dashed line in (a). (d) A terrain map in the same area.

reducing the domainwide MAE in each case. Further, the CNN model consistently corrected for the range-dependent biases seen in the physically based radar QPE, with performance only dropping off substantially at the farthest ranges from the radar where the beam overshoots the cloud top.

While the CNN model showed promise over the physically based method, there were certain regions and precipitation regimes for which more work is required. For example, the CNN model tended to consistently overestimate precipitation relative to gauges in certain valleys in the domain. These areas may lack the orographic enhancements of precipitation that are more prominent in the windward side of mountain ranges and thus have a lower precipitation efficiency than in the latter areas. Further, precipitation in these areas may be susceptible to evaporative effects below the radar beam and the current iteration of the CNN model cannot account for
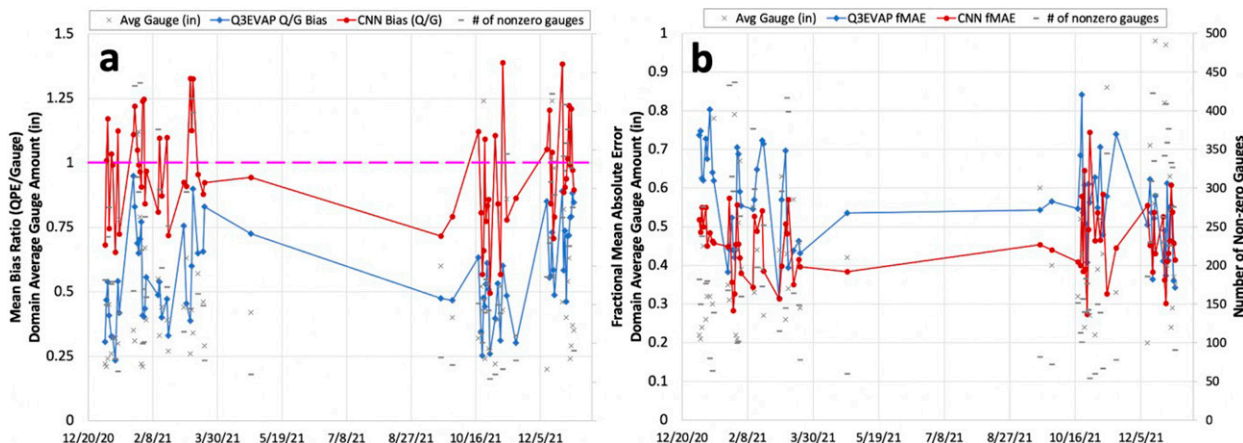

FIG. 19. (a) MBR and (b) fractional MAE of the daily Q3EVAP (blue line) and CNN (red line) QPEs for all days in 2021 with at least 50 gauges reporting nonzero values in the domain and where the domain average 24-h gauge amount exceeded 0.2 in. The gray "−" symbols indicate the number of nonzero gauges and gray "×" the domain average 24-h gauge amounts. The pink dashed line indicates no bias (MBR = 1.0).
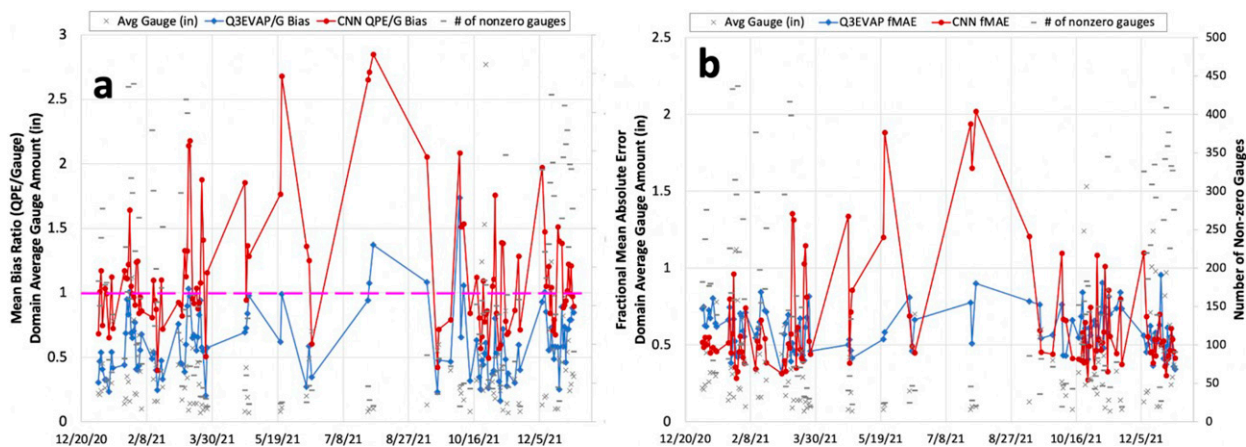
Fig. 20. As in Fig. 19, but for all days in 2021 with at least 20 nonzero gauges and domain average 24-h gauge value > 0 in.

this effect. Introducing new input variables directly related to atmospheric column moisture and wind–terrain interactions may help the model better capture these effects. The model also likely has issues at times accounting for brightband effects on the input fields, leading to overestimations such as that seen in the case of 13 March 2020. A removal of fields with uncorrected brightband contamination (e.g., CREF and RALA) might mitigate this latter issue and warrants further investigation. Another limitation seen in the CNN model results was the underestimation bias in areas where the radar beam is completely overshooting the cloud top and thus unable to give the model enough information for an accurate prediction. This can likely be improved through the use of nonradar input variables such as NWP model QPF and satellite QPE and will be explored in future work.

The long-term statistical analysis from all precipitation days in 2021 showed a strong seasonal pattern to the model's performance. The model generally shows good performance and improvement relative to the physically based QPE during the cool season when there are predominantly widespread, stratiform precipitation events. These types of events were well represented in the training dataset, and the CNN model provided a 20%–40% reduction in bias and a 10%–20% reduction in MAE over the current MRMS radar QPE. The CNN model statistics are less impressive in the warm season, with the CNN QPE often being outperformed by the physically based radar QPE. The model tends to strongly overestimate precipitation during the more scattered, convective warm season events. This difference in performance is likely a result of the training dataset exclusively containing cool season events and warrants further studies using additional training data from warm season scattered precipitation events.

This initial work is encouraging and points to the potential of deep learning approaches to supplement the MRMS precipitation products in the challenging, unique microphysical scenarios seen in complex terrain. To expand the applicability of the CNN model, additional testing and refinement of the input variables is being explored along with expanding to a more diverse training dataset. Further model interpretation

methods will be applied to better understand how the model responds to the selection of input variables, data preprocessing choices, and hyperparameter settings, all of which will help inform future model development. The CNN model will be evaluated over more cases across time and space to increase its technical readiness level for operational deployment.

Data availability statement. The data used in the current study are available from the NOAA National Severe Storms Lab Multi-Radar Multi-Sensor datasets (https://www.nssl.noaa.gov/projects/mrms/).

APPENDIX

Acronyms

AR	Atmospheric river
CC	Correlation coefficient
CNN	Convolutional neural network
CoCoRaHS	Community Collaborative Rain Hail and Snow Network
CREF	Composite reflectivity
DEM	Digital elevation model
DSD	Drop size distribution
fMAE	Fractional mean absolute error
IVT	Integrated vapor transport
LSTM	Long short-term memory model
MADIS	Meteorological Assimilation Data Ingest System
MAE	Mean absolute error
MBR	Mean bias ratio
MRMS	Multi-Radar Multi-Sensor
NWP	Numerical weather prediction

PRISM	Parameter-Elevation Regressions on Independent Slopes Model
Q3EVAP	MRMS dual-pol radar-based QPE with evaporation correction
Q3MS	MRMS multisensor QPE
Q3RAD	MRMS single-pol radar-based QPE
QPE	Quantitative precipitation estimation
QPF	Quantitative precipitation forecast
RALA	Reflectivity at lowest altitude
ReLU	Rectified linear unit
RQI	Radar quality index
SHSR	Seamless hybrid scan reflectivity
SHSRH	Seamless hybrid scan reflectivity height
VIL	Vertically integrated liquid

REFERENCES

Bytheway, J. L., M. Hughes, K. Mahoney, and R. Cifelli, 2020: On the uncertainty of high-resolution hourly quantitative precipitation estimates in California. *J. Hydrometeor.*, **21**, 865–879, https://doi.org/10.1175/JHM-D-19-0160.1.

Chen, H., R. Cifelli, and A. White, 2020: Improving operational radar rainfall estimates using profiler observations over complex terrain in northern California. *IEEE Trans. Geosci. Remote Sens.*, **58**, 1821–1832, https://doi.org/10.1109/TGRS.2019.2949214.

Chollet, F., 2018: *Deep Learning with Python*. Manning Publications, 361 pp.

——, and Coauthors, 2020: Keras. GitHub, https://github.com/fchollet/keras.

Cifelli, R., N. Doesken, P. Kennedy, L. D. Carey, S. A. Rutledge, C. Gimmestad, and T. Depue, 2005: The Community Collaborative Rain, Hail, and Snow Network: Informal education for scientists and citizens. *Bull. Amer. Meteor. Soc.*, **86**, 1069–1078, https://doi.org/10.1175/BAMS-86-8-1069.

Cocks, S. B., and Coauthors, 2019: A prototype quantitative precipitation estimation algorithm for operational S-band polarimetric radar utilizing specific attenuation and specific differential phase. Part II: Performance verification and case study analysis. *J. Hydrometeor.*, **20**, 999–1014, https://doi.org/10.1175/JHM-D-18-0070.1.

Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical–topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor. Climatol.*, **33**, 140–158, https://doi.org/10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2.

——, M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris, 2008: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int. J. Climatol.*, **28**, 2031–2064, https://doi.org/10.1002/joc.1688.

Germann, U., G. Galli, M. Boscacci, and M. Bolliger, 2006: Radar precipitation measurement in a mountainous region. *Quart. J. Roy. Meteor. Soc.*, **132**, 1669–1692, https://doi.org/10.1256/qj.05.190.

Hinton, G. E., N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, 2012: Improving neural networks by preventing co-adaptation of feature detectors. arXiv, 1207.0580v1, https://arxiv.org/abs/1207.0580.

Hochreiter, S., and J. Schmidhuber, 1997: Long short-term memory. *Neural Comput.*, **9**, 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735.

Kingsmill, D. E., P. J. Neiman, F. M. Ralph, and A. B. White, 2006: Synoptic and topographic variability of Northern California precipitation characteristics in landfalling winter storms observed during CALJET. *Mon. Wea. Rev.*, **134**, 2072–2094, https://doi.org/10.1175/MWR3166.1.

Lagerquist, R., A. McGovern, C. R. Homeyer, D. J. Gagne II, and T. Smith, 2020: Deep learning on three-dimensional multiscale data for next-hour tornado prediction. *Mon. Wea. Rev.*, **148**, 2837–2861, https://doi.org/10.1175/MWR-D-19-0372.1.

LeCun, Y., L. Bottou, Y. Bengio, and P. Haffner, 1998: Gradient-based learning applied to document recognition. *Proc. IEEE*, **86**, 2278–2324, https://doi.org/10.1109/5.726791.

Maas, A. L., A. Y. Hannun, and A. Y. Ng, 2013: Rectifier nonlinearities improve neural network acoustic models. *Proc. 30th Int. Conf. on Machine Learning*, Atlanta, GA, International Machine Learning Society, 3, http://robotics.stanford.edu/~amaas/papers/relu_hybrid_icml2013_final.pdf.

Maddox, R. A., J. Zhang, J. J. Gourley, and K. W. Howard, 2002: Weather radar coverage over the contiguous United States. *Wea. Forecasting*, **17**, 927–934, https://doi.org/10.1175/1520-0434(2002)017<0927:WRCOTC>2.0.CO;2.

Martinaitis, S. M., and Coauthors, 2020: A physically based multisensor quantitative precipitation estimation approach for gap-filling radar coverage. *J. Hydrometeor.*, **21**, 1485–1511, https://doi.org/10.1175/JHM-D-19-0264.1.

——, S. B. Cocks, M. J. Simpson, A. P. Osborne, S. S. Harkema, H. M. Grams, J. Zhang, and K. W. Howard, 2021: Advancements and characteristics of gauge ingest and quality control within the Multi-Radar Multi-Sensor system. *J. Hydrometeor.*, **22**, 2455–2474, https://doi.org/10.1175/JHM-D-20-0234.1.

Miao, Q., B. Pan, H. Wang, K. Hsu, and S. Sorooshian, 2019: Improving monsoon precipitation prediction using combined convolutional and long short term memory neural network. *Water*, **11**, 977, https://doi.org/10.3390/w11050977.

Neiman, P. J., F. M. Ralph, G. A. Wick, J. D. Lundquist, and M. D. Dettinger, 2008: Meteorological characteristics and overland precipitation impacts of atmospheric rivers affecting the west coast of North America based on eight years of SSM/I satellite observations. *J. Hydrometeor.*, **9**, 22–47, https://doi.org/10.1175/2007JHM855.1.

Pan, B., K. Hsu, A. AghaKouchak, and S. Sorooshian, 2019: Improving precipitation estimation using convolutional neural network. *Water Resour. Res.*, **55**, 2301–2321, https://doi.org/10.1029/2018WR024090.

Qi, Y., S. Martinaitis, J. Zhang, and S. Cocks, 2016: A real-time automated quality control of hourly rain gauge data based on multiple sensors in MRMS system. *J. Hydrometeor.*, **17**, 1675–1691, https://doi.org/10.1175/JHM-D-15-0188.1.

Sadeghi, M., A. A. Asanjan, M. Faridzad, P. Nguyen, K. Hsu, S. Sorooshian, and D. Braithwaite, 2019: PERSIANN-CNN: Precipitation estimation from remotely sensed information using artificial neural networks–convolutional neural networks. *J. Hydrometeor.*, **20**, 2273–2289, https://doi.org/10.1175/JHM-D-19-0110.1.

Sha, Y., D. J. Gagne II, G. West, and R. Stull, 2021: Deep-learning-based precipitation observation quality control. *J. Atmos. Oceanic Technol.*, **38**, 1075–1091, https://doi.org/10.1175/JTECH-D-20-0081.1.

Wang, Y., S. Cocks, L. Tang, A. Ryzhkov, P. Zhang, J. Zhang, and K. Howard, 2019: A prototype quantitative precipitation estimation algorithm for operational S-band polarimetric radar utilizing specific attenuation and specific differential phase. Part I: Algorithm description. *J. Hydrometeor.*, **20**, 985–997, https://doi.org/10.1175/JHM-D-18-0071.1.

Xu, X., K. Howard, and J. Zhang, 2008: An automated radar technique for the identification of tropical precipitation. *J. Hydrometeor.*, **9**, 885–902, https://doi.org/10.1175/2007JHM954.1.

Yang, Q., C.-Y. Lee, and M. K. Tippett, 2020: A long short-term memory model for global rapid intensification prediction. *Wea. Forecasting*, **35**, 1203–1220, https://doi.org/10.1175/WAF-D-19-0199.1.

Yo, T.-S., S.-H. Su, J.-L. Chu, C.-W. Chang, and H.-C. Kuo, 2021: A deep learning approach to radar-based QPE. *Earth Space Sci.*, **8**, e2020EA001340, https://doi.org/10.1029/2020EA001340.

Zhang, J., C. Langston, and K. Howard, 2008: Brightband identification based on vertical profiles of reflectivity from the WSR-88D. *J. Atmos. Oceanic Technol.*, **25**, 1859–1872, https://doi.org/10.1175/2008JTECHA1039.1.

——, Y. Qi, D. Kingsmill, and K. Howard, 2012a: Radar-based quantitative precipitation estimation for the cool season in complex terrain: Case studies from the NOAA hydrometeorology testbed. *J. Hydrometeor.*, **13**, 1836–1854, https://doi.org/10.1175/JHM-D-11-0145.1.

——, ——, K. Howard, C. Langston, and B. Kaney, 2012b: Radar quality index (RQI)—A combined measure of beam blockage and VPR effects in a national network. *Wea. Radar Hydrol.*, **351**, 388–393.

——, and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 621–638, https://doi.org/10.1175/BAMS-D-14-00174.1.

——, L. Tang, S. Cocks, P. Zhang, A. Ryzhkov, K. Howard, C. Langston, and B. Kaney, 2020: A dual-polarization radar synthetic QPE for operations. *J. Hydrometeor.*, **21**, 2507–2521, https://doi.org/10.1175/JHM-D-19-0194.1.

Zhu, Y., and R. E. Newell, 1998: A proposed algorithm for moisture fluxes from atmospheric rivers. *Mon. Wea. Rev.*, **126**, 725–735, https://doi.org/10.1175/1520-0493(1998)126<0725:APAFMF>2.0.CO;2.