Nowcasting Heavy Rainfall With Convolutional Long Short-Term Memory Networks: A Pixelwise Modeling Approach

Yi Victor Wang, *Member, IEEE*, Seung Hee Kim, *Member, IEEE*, Geunsu Lyu, Choeng-Lyong Lee, Soorok Ryu, Gyuwon Lee, Ki-Hong Min, and Menas C. Kafatos, *Life Member, IEEE*

***Abstract*—The recent decades have seen an increasing academic interest in leveraging machine learning approaches to nowcast, or forecast in a highly short-term manner, precipitation at a high resolution, given the limitations of the traditional numerical weather prediction models on this task. To capture the spatiotemporal associations of data on input variables, a deep learning (DL) architecture with the combination of a convolutional neural network and a recurrent neural network can be an ideal design for nowcasting rainfall. In this study, a long short-term memory (LSTM) modeling structure is proposed with convolutional operations on input variables. To resolve the issue of underestimation of heavy rainfall that challenges most of the DL models, a pixelwise modeling approach is adopted to facilitate a stratified sampling process in generating training data points for calibrating models to predict rain rates at locations. The proposed pixelwise convolutional LSTM (CLSTM) models are applied to data on mesoscale convective systems during the warm seasons over the Korean Peninsula. Results show a significant and consistent improvement in prediction skill scores produced by the CLSTM models than a traditional rainfall nowcasting method, the McGill algorithm for precipitation nowcasting by Lagrangian extrapolation, across all considered lead times from 10 to 60 min. Future work needs to reduce the relatively large false positive rates produced by the CLSTM models and their blurring effect in mapping spatial distributions of rain rates, in particular for longer lead times.**

***Index Terms*—Artificial neural network, convolutional neural network, deep learning (DL), dual-polarimetric weather radar, early warning, hydrometeorological hazard, long short-term memory (LSTM) network, mesoscale convective system, rainfall nowcasting, recurrent neural network (RNN), remote sensing, storm.**

Manuscript received 11 September 2023; revised 13 December 2023 and 12 February 2024; accepted 23 March 2024. Date of publication 2 April 2024; date of current version 22 April 2024. This work was supported by the Korea Meteorological Administration Research and Development Program under Grant RS-2023-00237740. *(Corresponding author: Seung Hee Kim.)*

Yi Victor Wang is with the Department of Emergency Management, Massachusetts Maritime Academy, Buzzards Bay, MA 02532 USA (e-mail: vwang@maritime.edu).

Seung Hee Kim and Menas C. Kafatos are with the Institute for Earth, Computing, Human and Observing, Chapman University, Orange, CA 92866 USA (e-mail: sekim@chapman.edu; kafatos@chapman.edu).

Geunsu Lyu, Choeng-Lyong Lee, and Soorok Ryu are with the Center for Atmospheric REmote sensing (CARE), Kyungpook National University, Daegu 41566, Republic of Korea (e-mail: geunsulyu@gmail.com; lchly777@knu.ac.kr; sryu@knu.ac.kr).

Gyuwon Lee and Ki-Hong Min are with the BK21 Weather Extremes Education & Research Team, Department of Atmospheric Sciences, Center for Atmospheric Remote sensing (CARE), Kyungpook National University, Daegu 41566, Republic of Korea (e-mail: gyuwon@knu.ac.kr; kmin@knu.ac.kr).

Digital Object Identifier 10.1109/JSTARS.2024.3383397

I. Introduction

HIGHLY short-term forecasting of up to a few hours, or nowcasting, of precipitation is essential for managing hazardous emergencies, such as flash floods, under sudden extreme meteorological conditions [1]. To provide accurate and precise short-term projections of heavy rainfall at a high spatiotemporal resolution, however, the numerical weather prediction (NWP) models that are widely used for daily weather forecasting usually lack reliability, mainly, due to the uncertainties in the derivation of the initial and boundary conditions, the exceptionally long spin-up time, and the sophisticated process of physical parameterization [1], [2], [3], [4].

To improve the performance of rainfall nowcasting, traditionally, a number of tracking and computational methods have been proposed and applied based on algorithms and predefined rules to fuse meteorological data from multiple remotely sensing and ground-based sources and extrapolate radar echoes to locate and estimate precipitations [5], [6], [7], [8], [9]. These methods include, for example, the Storm Cell Identification and Tracking system [5], the Thunderstorm Identification, Tracking and Nowcasting system [6], the Tracking Reflectivity Echoes by Correlation system [7], [8], and the Auto-Nowcast system [9]. Among these traditional rainfall nowcasting methods, the McGill algorithm for precipitation nowcasting by Lagrangian extrapolation (MAPLE) is one of the most recently developed algorithms to use a variety of meteorological equations to estimate the locations and amounts of future precipitations with prior radar observations of storm motion [10], [11], [12], [13], [14]. Based on the variational echo tracking technique, the MAPLE is capable of producing rainfall estimates at a high spatiotemporal resolution [15]. It has shown relatively high predictive skill when being applied for precipitation nowcasting, particularly, in regions with complex terrains, such as the European Alpine region [16] and the Korean Peninsula [17], [18]. Albeit with improved predictive performance compared to the NWP models, the traditional nowcasting methods, such as MAPLE, are still considered to be limited largely because their physical assumptions only partially represent the true atmospheric states [19].

Unlike the traditional methods, machine learning can be leveraged to predict rainfall with information extracted directly from

observational data without relying on physical assumptions. In particular, deep learning (DL) that is based on an artificial neural network (ANN) architecture has been gaining increasing popularity during recent decades for nowcasting rainfall [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32]. An ANN model of the generic form of the multilayer perceptron (MLP) can be trained with observational input and output data to efficiently predict rainfall of storms at specific moments [1], [20]. To capture the spatial associations of input and output data across the model area for locational nowcasting, convolutional computations can be conducted in addition to the matrix multiplications within an MLP-ANN. The convolution-based ANNs, exemplified by the conventional convolutional neural network [19], [21], [22], [23], the generative adversarial network [24], [25], and the U-net [26], [27], [28], [29], have shown good predictive performance in rainfall nowcasting. To better generalize the temporal patterns of sequential training data, modelers have also proposed to leverage a recurrent neural network (RNN) design, with memory cells and feedback connections, to augment the feedforward convolution-based ANN models [30], [31], [32]. In particular, the long short-term memory (LSTM) architecture is capable of retaining long-term temporal dependencies in the data without causing the vanishing gradient problem that is commonly associated with training RNNs [33]. Despite the unbiased good predictive performance in general, DL models, especially the convolution-based ones producing an entire image as the output, tend to significantly underestimate heavy rainfall values, while the ability to accurately identify and quantify heavy rainfall is essential to the provision of reliable early warnings against extreme pluvial conditions.

In this study, a set of convolutional LSTM (CLSTM) regression DL models was proposed with remotely sensed data to nowcast rain rates of 10-min intervals of mesoscale convective storms (MCSs) for specific locations in a pixelwise fashion. For capturing the spatial associations of input variables, a convolutional method was used in addition to the conventional LSTM modeling. The pixelwise modeling approach was adopted to resolve the issue of underestimation of heavy rainfall. Rainfall images were used as the ground truth output data for training and testing models. The CLSTM models were trained with data for the Korean Peninsula to provide predictions of rain rates with lead times of 10, 20, 30, 40, 50, and 60 min. The predictive performances of the proposed models were compared with the ones produced by the traditional rainfall nowcasting algorithm MAPLE.

The rest of this article is organized as follows. Section II introduces the collection and processing of data for the study. In Section III, the architecture of the proposed CLSTM models and the methodology for model training and testing are presented. In Section IV, the results of the study are laid out. Section V presents discussions regarding the comparisons of the predictive performances of the proposed models and the MAPLE. Finally, Section VI concludes this article and summarizes the significance and limitations of the study and points out the directions for future work.

II. Data

The MCSs were classified into three types: convective cells (CCs), mesoscale convective complexes, and squall lines. This classification was based on meteorological and remotely sensed image data collected across the Korean Peninsula during the warm seasons from June 2018 to August 2021 [1].

All the data used for training were obtained from a dual-polarimetric Doppler radar. This included horizontal reflectivity (Z_{H}), differential reflectivity (Z_{DR}), and specific differential phase (K_{DP}). Rainfall intensity (R_{I}) data, corrected from ground rain gauge readings and converted to precipitation using a *Z–R* relationship with $Z = Z_{\mathrm{H}}$, were also used as an input–output variable. Z_{H}, obtainable from single-polarized radar, is mainly used for precipitation estimation using a *Z–R* relationship. Z_{DR}, the logarithm of the ratio of Z_{H} to vertical reflectivity (Z_{V}), varies depending on the shape of the precipitation particle and is primarily used for precipitation-type classification (e.g., rain, drizzle, graupel, and hail). K_{DP} represents the difference in specific differential phase for horizontally and vertically polarized radar pulses in a given area. Precipitation estimation using K_{DP}, often combined with Z_{H}, offers several advantages over the sole use of a *Z–R* relationship for more accurate precipitation estimation. Data on these radar variables were provided in the form of plan-position indicators (PPIs) with quality control for each radar site, accessible through the application programming interface hub website operated by the Korea Meteorological Administration.[1]

To composite the PPI images from various radars, a hybrid surface rainfall radar composite algorithm [34] was employed. The training data were generated over a 2-D area extending 1200 km from north to south and 960 km from west to east.

The original data had a spatial and temporal resolutions of 1 km and 10 min, respectively. The composite R_{I} images were used as the output data and one of the major input variables. In the study, the rainfall intensity (R_{I}) in mm/h is derived by correcting the original readings using ground rain gauges. This correction is applied using the following equation:

$$R_{\mathrm{I}} = R_{\mathrm{I\ Original}} \times 2.08 + 0.48 \tag{1}$$

where the original R_{I} ($R_{\mathrm{I\ Original}}$) is computed from Z_{H} using the *Z–R* relationshop: $Z = 200R^{1.6}$ [35].

The output of a CLSTM model was the aggregated rain rate (ARR) of a coarse pixel derived from a corresponding RI image. The ARR was computed based on all 10-min rain rate values of the original data pixels with a spatial resolution of 1 km within the scope of the coarse pixel. The ARR can be the mean, median, or maximum of the rain rates of considered original data pixels. The maximum method was adopted for the ARR in this study to form datasets with sufficient numbers of data points for model training. Based on trials and errors for producing optimal results, the spatial resolution of the coarse pixel was set at 12 km.

The input of the CLSTM models was a time series of image patches of four variables, including R_{I} along with three dual-polarimetric Doppler radar variables of Z_{H}, Z_{DR}, and K_{DP},

[1] [Online]. Available: https://apihub.kma.go.kr/

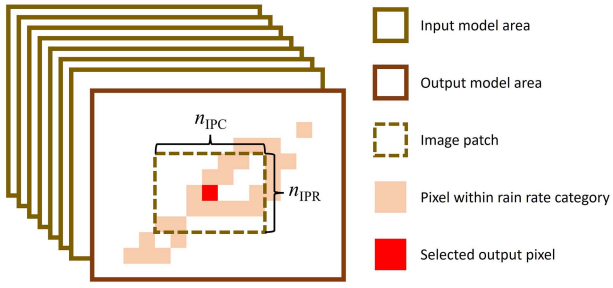


Fig. 1. Selection of a data point.

centering at the output coarse pixel. The spatial resolutions of the input variables were converted to be the same as the output resolution with the same maximum method for data aggregation as of the output data. Each time series of each input variable consisted of n_t+1 image patches of the now moment and n_t previous time stamps corresponding to n_t 10-min intervals, where n_t= 6. The determination of n_t was important, as a larger n_t would help establish models with information from input data involving more past time stamps. Meanwhile, however, a smaller n_t would be preferred to avoid an unnecessarily long training time due to too many trainable parameters of the CLSTM models. Corresponding to the six nowcasting lead times of 10, 20, 30, 40, 50, and 60 min, six datasets were created to train and test six CLSTM models, respectively.

For effectively training and testing CLSTM models, data points with paired information on input and output were generated based on their output R_I values. Adopted for rain rate output were 11 categories, including: 1) zero rain rate; 2) (0, 10]; 3) (10, 20]; 4) (20, 30]; 5) (30, 40]; 6) (40, 50]; 7) (50, 60]; 8) (60, 70]; 9) (70, 80]; 10) (80, 90]; and 11) (90, ∞) mm/h. For each output category, data points were generated separately. To facilitate training and testing models, extreme values of variables of R_I, Z_H, Z_{DR}, and K_{DP} were converted to their corresponding minimum or maximum value of the adopted ranges of [0, 100] mm/h, [–32, 70] dB, [–8, 8] dB, and [–2, 10]°/km, respectively. The variable values were then standardized to be within the range of [0, 1], accordingly. Data points with time stamps before year 2021 were used for model training, while ones of year 2021 were used for both model testing and visual presentation.

For the selection of pairs of input and output data to be included for training and testing models, a time stride n_{ts}= 4 was adopted such that the selected output time stamps were at least 10 × n_{ts} min apart from each other. For each output time stamp and each rain rate category, at most n_s= 20 output pixels were randomly selected. Fig. 1 shows an example of selection of output pixel for one rain rate category. For each selected output pixel, image patches of the size of n_{IPR}=n_{IPC}= 25 were created centering at the output pixel location for each input time stamp, where n_{IPR} and n_{IPC} were, respectively, the numbers of rows and columns of the image patches. The values of the hyperparameters of n_{ts}, n_s, n_{IPR}, and n_{IPC} were determined based on trials and errors to provide an optimal training process, regarding the maximization of information gain from data within a limited time budget. Following the data generation process, data points were prepared according to the rain rate categories and nowcasting lead times for training and testing CLSTM models, as listed in Tables I and II.

TABLE I
NUMBERS OF GENERATED TRAINING DATA POINTS FOR PREDICTION LEAD TIMES AND RAIN RATE CATEGORIES

Rain rate (mm/h)	10-min	20-min	30-min	40-min	50-min	60-min
0	256 289	255 656	255 991	255 698	255 226	254 815
(0, 10]	263 628	262 825	262 322	262 375	261 884	261 211
(10, 20]	115 801	115 622	115 216	115 338	115 616	115 422
(20, 30]	70 306	70 239	70 191	70 104	70 216	70 150
(30, 40]	42 045	41 758	41 902	41 834	42 008	41 705
(40, 50]	24 449	24 515	24 426	24 226	24 444	24 497
(50, 60]	15 057	14 832	14 748	14 899	15 052	14 823
(60, 70]	9130	8994	9285	9077	9121	8993
(70, 80]	5886	5825	6000	5786	5878	5820
(80, 90]	3831	3884	3793	3942	3828	3881
(90, ∞)	8199	8070	8119	8141	8196	8067

TABLE II
NUMBERS OF GENERATED TESTING DATA POINTS FOR PREDICTION LEAD TIMES AND RAIN RATE CATEGORIES

Rain rate (mm/h)	10-min	20-min	30-min	40-min	50-min	60-min
0	2140	2120	2120	2120	2120	2100
(0, 10]	2140	2120	2120	2120	2120	2100
(10, 20]	2108	2077	2076	2082	2091	2059
(20, 30]	1832	1767	1782	1776	1825	1760
(30, 40]	1291	1275	1278	1242	1285	1271
(40, 50]	879	902	901	896	879	900
(50, 60]	627	605	581	617	625	605
(60, 70]	404	419	446	419	404	419
(70, 80]	310	304	281	283	310	304
(80, 90]	233	205	219	188	233	205
(90, ∞)	448	444	445	474	448	444

III. METHODOLOGY

In this study, six CLSTM regression DL-ANN models were trained and tested with the processed datasets for six lead times, i.e., 10, 20, 30, 40, 50, and 60 min, respectively, for nowcasting rainfall values. The models were trained with Python 3.7.7 and TensorFlow 2.1 on a server of graphics processing unit of NVIDIA-SMI 440.95.01 with the driver version 220.950.01 and CUDA version 10.2 [36], [37].

A. Proposed Model

The CLSTM models proposed in the study shared the same architecture, as shown in Fig. 2. For each input variable, i.e., R_I, Z_H, Z_{DR}, or K_{DP}, an image patch corresponding to a time stamp of a data point was processed with three sets of convolutional operations with batch normalization. In between the convolutions were two operations of max pooling before the resulting four feature maps of the size of 2 by 2 were flattened into a vector of 16 nodes and fed into an LSTM cell. The "Other input" in Fig. 2 refers to the flattened vectors of the other input variables following the same convolutional processes. Each LSTM cell had 32 hidden computational units. These 32 hidden units were fully connected to another two vanilla ANN layers consisting of

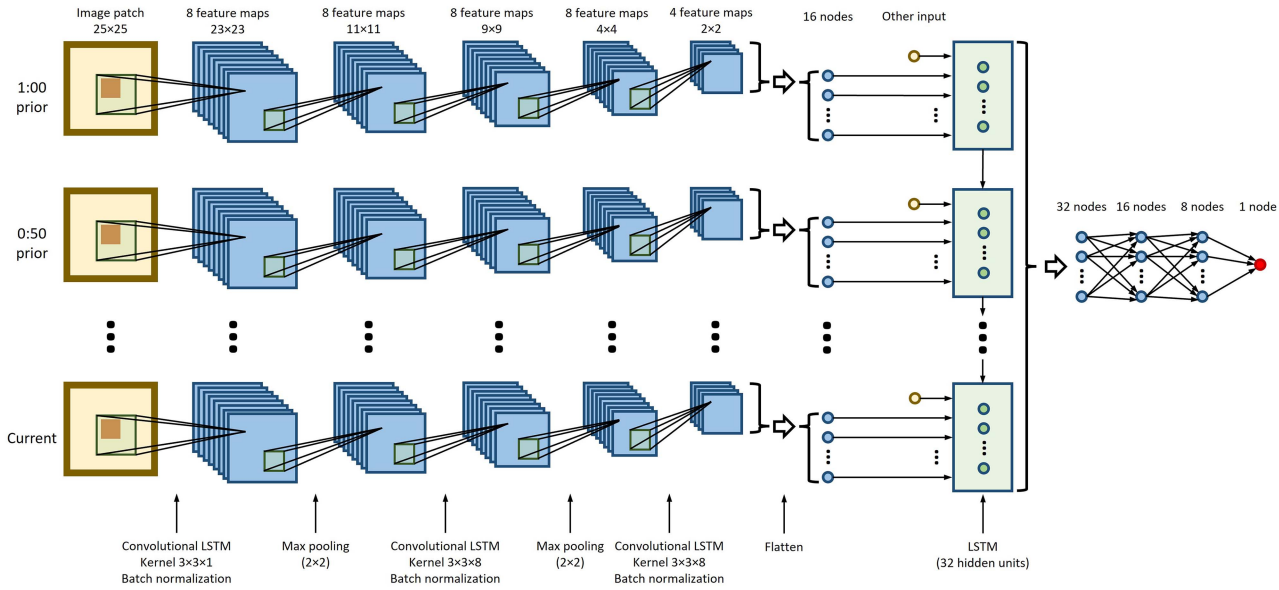


Fig. 2. Architecture of the proposed model.

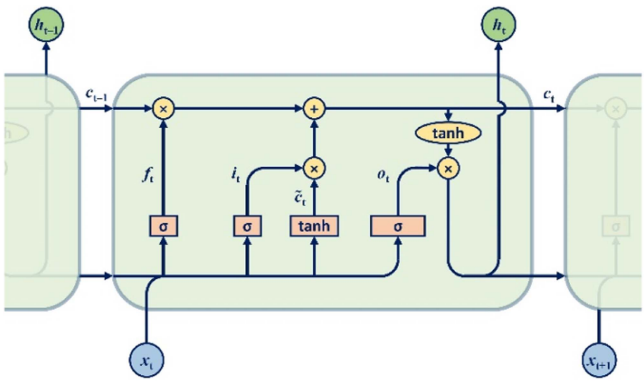


Fig. 3. LSTM cell.

16 and 8 nodes, respectively. The final output layer of the model had only one node, corresponding to the predicted rain rate value at the pixelated location at the center of the image patches of the data point. Apart from the output layer with an identity activation function, all convolutional and fully connected layers were assigned with an exponential linear unit (ELU) activation function [38]

$$A_{\mathrm{ELU}}(x) = \begin{cases} \exp(x) - 1, & x \leq 0 \\ x, & x > 0. \end{cases} \tag{2}$$

The temporal interdependencies of the input data were captured via the adoption of the LSTM cells, as detailed in Fig. 3. For each LSTM cell at time stamp t, it was connected to the cell at the previous time stamp $t-1$ through the cell state vector ($\boldsymbol{c}_{t-1}$) and the hidden state vector ($\boldsymbol{h}_{t-1}$). The LSTM cell adopted in the study consisted of three computational gates, i.e., the forget gate, input gate, and output gate. The forget gate determined whether and to what extent to keep or ignore information from previous time stamps. At the forget gate

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_{\boldsymbol{f}}[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t]^{\mathrm{T}} + \boldsymbol{b}_{\boldsymbol{f}}\right) \tag{3}$$

where $\boldsymbol{f}_t$ is the forget gate vector, $\sigma(\cdot)$ is the logistic sigmoid activation function, $\boldsymbol{W}_{\boldsymbol{f}}$ is the weight matrix at the forget gate, $\boldsymbol{x}_t$ is the input vector flattened from the convolutional operations, T is the transpose operator, and $\boldsymbol{b}_{\boldsymbol{f}}$ is the bias vector at the forget gate. The input gate provided updates to the cell state vector via

$$\boldsymbol{i}_t = \sigma\left(\boldsymbol{W}_{\boldsymbol{i}}[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t]^{\mathrm{T}} + \boldsymbol{b}_{\boldsymbol{i}}\right) \tag{4}$$

$$\tilde{\boldsymbol{c}}_t = \tanh\left(\boldsymbol{W}_{\tilde{\boldsymbol{c}}}[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t]^{\mathrm{T}} + \boldsymbol{b}_{\tilde{\boldsymbol{c}}}\right) \tag{5}$$

where $\boldsymbol{i}_t$ is the input gate vector, $\tilde{\boldsymbol{c}}_t$ is the cell input vector with activation, $\tanh(\cdot)$ is the hyperbolic tangent activation function, and $\boldsymbol{W}_{\boldsymbol{i}}$, $\boldsymbol{W}_{\tilde{\boldsymbol{c}}}$, $\boldsymbol{b}_{\boldsymbol{i}}$, and $\boldsymbol{b}_{\tilde{\boldsymbol{c}}}$ are, respectively, the corresponding weight matrices and bias vectors. With the forget gate and input gate, the cell state at time stamp t could be computed as

$$\boldsymbol{c}_t = \boldsymbol{f}_t \otimes \boldsymbol{c}_{t-1} + \boldsymbol{i}_t \otimes \tilde{\boldsymbol{c}}_t \tag{6}$$

where $\otimes$ denotes the Hadamard product operation. At the output gate, the value of the hidden state at time stamp t was decided via

$$\boldsymbol{o}_t = \sigma\left(\boldsymbol{W}_{\boldsymbol{o}}[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t]^{\mathrm{T}} + \boldsymbol{b}_{\boldsymbol{o}}\right) \tag{7}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \otimes \tanh(\boldsymbol{c}_t) \tag{8}$$

where $\boldsymbol{o}_t$ is the output gate vector, and $\boldsymbol{W}_{\boldsymbol{o}}$ and $\boldsymbol{b}_{\boldsymbol{o}}$ are the weight matrix and bias vector at the output gate, respectively.

B. Model Training

To establish each CLSTM model corresponding to one of the six considered lead times, a shallow training scheme was adopted with 400 iterations. During each iteration for each model, a

TABLE III
PERCENTAGES OF TRAINING DATA POINTS SELECTED FOR MODEL TRAINING FROM GENERATED DATA POINTS CORRESPONDING TO RAIN RATE CATEGORIES

Rain rate (mm/h)	Percentage
0	10%
(0, 10]	9%
(10, 20]	8.1%
(20, 30]	7.3%
(30, 40]	6.6%
(40, 50]	6%
(50, 60]	5.5%
(60, 70]	5.1%
(70, 80]	4.8%
(80, 90]	4.6%
(90, ∞)	4.5%

stratified sampling process was used to randomly select data points without replacement from each of the 11 training datasets corresponding to the rain rate categories. The percentages of selected training data points are listed in Table III. The selected data points were then pooled together. The pooled training dataset was later used to train the same CLSTM model for five epochs before the next iteration commenced. The mean squared error (MSE) was used as the loss function for model training

$$\text{MSE} = \frac{1}{n_y}\sum_{i=1}^{n_y}(y_i - \hat{y}_i)^2 \tag{9}$$

where n_y refers to the number of data points, y_i is the observed output value of the ith data point, and $\hat{y}_i$ is the model prediction. The adaptive moment estimation (Adam) algorithm was adopted for training the CLSTM models [39]. At the beginning of each training iteration, a learning rate (R_{Lg}) was updated with

$$R_{\text{Lg}} = \frac{R_{\text{LInitial}}}{1+r_{\text{LRD}}\,(g-1)} \tag{10}$$

where $R_{\text{LInitial}} = 5 \times 10^{-4}$ is the initial learning rate, $r_{\text{LRD}} = 0.1$ is the ratio of learning rate decay, and g is a dummy variable referring to the number of training iteration. For each epoch of training, the batch size was set at 128, and 10% of the training data points were used for deriving training validation MSE. The values of R_{LInitial}, r_{LRD}, batch size, and ratio of training data points for validation were determined to provide optimal training performance based on trials and errors.

C. Model Testing

To compare the predictive performances of the trained CLSTM models with the ones of the MAPLE, the testing datasets were used to derive eight testing skill metrics for two sets of classification scenarios based on 11 rain rate categories. The first set, or the within scenarios, considered whether predictions were correct regarding being within or outside a rain rate category. The within scenarios included 11 categories of rain rate, as mentioned previously. The second set, or the greater scenarios, examined whether predictions were correct in terms of being greater than a threshold of rain rate. For the greater scenarios, ten thresholds of rain rate were used, including 0, 10, 20, 30, 40, 50, 60, 70, 80, and 90 mm/h. The testing skill metrics included the false positive rate (FPR), false alarm rate (FAR), recall, precision, F1 score, Heidke skill score (HSS), critical success index (CSI), and equitable threat score (ETS)

$$\text{FPR} = \frac{\text{FP}}{\text{FP}+\text{TN}} \tag{11}$$

$$\text{FAR} = \frac{\text{FP}}{\text{TP}+\text{FP}} \tag{12}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}} \tag{13}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}} \tag{14}$$

$$\text{F1} = \frac{2\ \text{Precision}\ \cdot \text{Recall}}{\text{Precision}+\text{Recall}} \tag{15}$$

$$\text{HSS} = \frac{\text{TP}\ \cdot\ \text{TN} - \text{FN}\ \cdot\ \text{FP}}{(\text{TP}+\text{FN})\ \cdot\ (\text{FN}+\text{TN}) + (\text{TP}+\text{FP})\ \cdot\ (\text{FP}+\text{TN})} \tag{16}$$

$$\text{CSI} = \frac{\text{TP}}{\text{TP}+\text{FP}+\text{FN}} \tag{17}$$

$$\text{ETS} = \frac{\text{TP} - w_{\text{ETS}}}{\text{TP}+\text{FP}+\text{FN}-w_{\text{ETS}}} \tag{18}$$

where TP, FP, TN, and FN are the numbers of true positives, false positives, true negatives, and false negatives, respectively, and

$$w_{\text{ETS}} = \frac{(\text{TP}+\text{FP})\ \cdot\ (\text{TP}+\text{FN})}{\text{TP}+\text{FP}+\text{TN}+\text{FN}}. \tag{19}$$

IV. RESULTS

Results of this study consisted of three parts. They included the visualization of the process of training individual CLSTM models, comparative analyses of testing skill metrics between the proposed CLSTM models and the MAPLE, and the comparisons between the nowcasting maps of rain rate values generated by the proposed CLSTM models and the MAPLE.

A. Training Processes

The processes for training the proposed CLSTM models were monitored via derivation of the training and validation MSEs at each epoch. Fig. 4 displays an example of the process of training the CLSTM model with the lead time of 30 min. As shown in Fig. 4(b), the zigzag shapes of the training MSE curves correspond to the five-epoch iterations of the adopted shallow training scheme. The training processes for the CLSTM models with other lead times produced the similar curves of the training and validation MSEs as the ones shown here. The curves indicate good performance of the model training. They also imply that, with more training epochs, the training and validation MSEs may be further reduced, although such a further loss reduction may be insignificant as well as highly inefficient given the exceptionally long training time required.

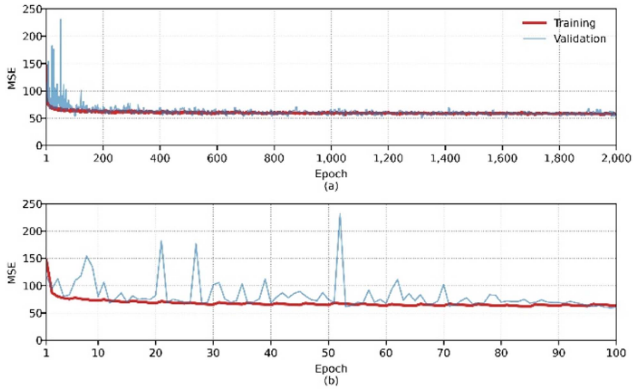


Fig. 4. Training and validation MSEs during the training process for the CLSTM model with a 30-min lead time for (a) all 2000 and (b) the first 100 training epochs.

B. Testing Skills

To compare the predictive performances of the CLSTM models and the MAPLE, eight skill metrics were computed for both the within and greater scenarios. Fig. 5 shows an example of the comparison of the skill scores across all six lead times for the within scenario, where models predicted whether rain rate would be within (30, 40] mm/h. Apart from FPR, FAR, and precision, all other five skill metrics indicate significant and consistent better performances of the CLSTM models than the MAPLE. For FAR and precision, the CLSTM models and the MAPLE produced similar results. Albeit with overall better performances, the CLSTM models resulted in much higher FPRs than the MAPLE, indicating a blurring effect in the predictions of rain rate by the CLSTM models. For the within scenarios with other rain rate categories, the differences between the prediction skill metric curves of the CLSTM models and the MAPLE were similar.

Regarding the greater scenarios, Fig. 6 shows the comparisons of skill scores of the CLSTM models and the MAPLE for whether predicted rain rates are greater than 30 mm/h. Skill scores for other rain rate thresholds presented similar comparisons between the CLSTM models and the MAPLE. For practical purposes, the rain rate threshold of 30 mm/h can be a good indicator for issuing early warning of heavy rainfall [1]. According to Fig. 6, apart from FPR, all other seven prediction skill metrics indicated significant and consistent better performances of the CLSTM models than the MAPLE. For FAR and precision, unlike in the within scenarios [see Fig. 5(b) and (c)], these two metrics in the greater scenarios [see Fig. 6(b) and (c)] also suggested the superiority of the CLSTM models over the MAPLE. The results on the FARs and precisions indicated that although the CLSTM models and MAPLE predicted rain rates with similar performances for each individual rain rate categories in terms of FAR and precision, the CLSTM models tend to better predict high rain rate values, which is preferred for provision of early warnings of heavy rainfall. In addition, the F1 scores of the CLSTM models stayed above 0.5 across all lead times considered in the study, indicating a high quality in the predictive performance.

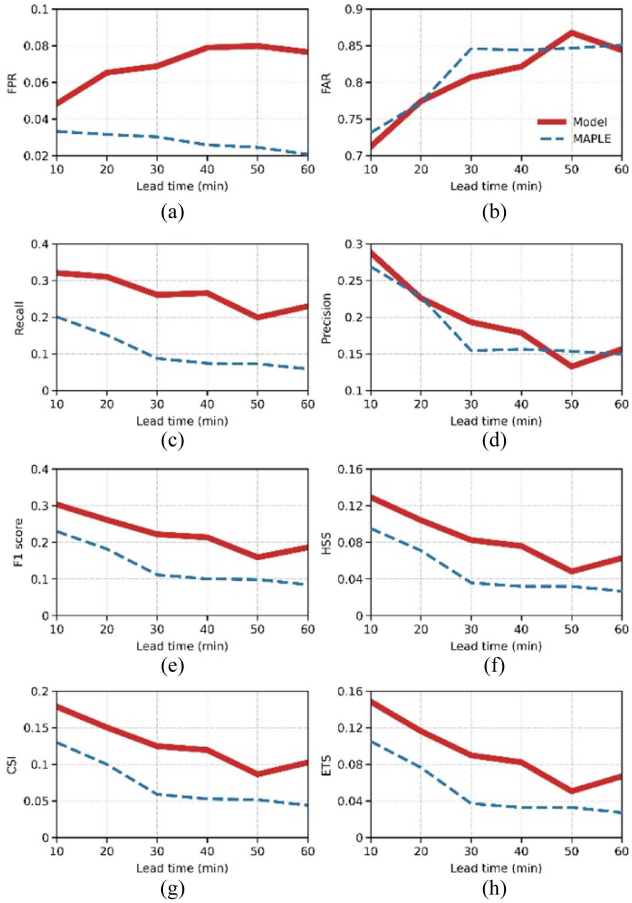


Fig. 5. Comparison of prediction skill scores derived with CLSTM models and the MAPLE for predicting rain rate within (30, 40] mm/h with lead times of 10, 20, 30, 40, 50, and 60 min, in terms of (a) FPR, (b) FAR, (c) recall, (d) precision, (e) F1 score, (f) HSS, (g) CSI, and (h) ETS. Note that a higher score indicates better performance, except for the FAR, where a perfect score is 0.

Similar to the within scenarios, the CLSTM models tend to result in worse FPRs than the MAPLE in the greater scenarios. As shown in Fig. 6(a), apart from the 10-min lead time, all the other FPRs of the CLSTM models were larger than the ones of the MAPLE for the rain rate threshold of 30 mm/h. The results on FPRs suggested that the CLSTM models, in particular with longer lead times, tend to overestimate small or 0 rain rate values and to predict a larger area with rainfall around the actual area with rainfall. This may be the cause of a blurring effect in the visual presentation of model predictions of rain rates, more so for the longer lead times.

C. Visual Presentations

The trained CLSTM models can be applied to create maps to predict the geographical distributions of rain rates with their corresponding lead times. Fig. 7 displays an example of the observed and predicted rain rates for a now moment of three MCS cases for the lead time of 60 min. Despite the overall better predictive performances in terms of the skill scores, the CLSTM models tend to produce blurred distributions of rain rates, in particular for the longer lead times (right figures in

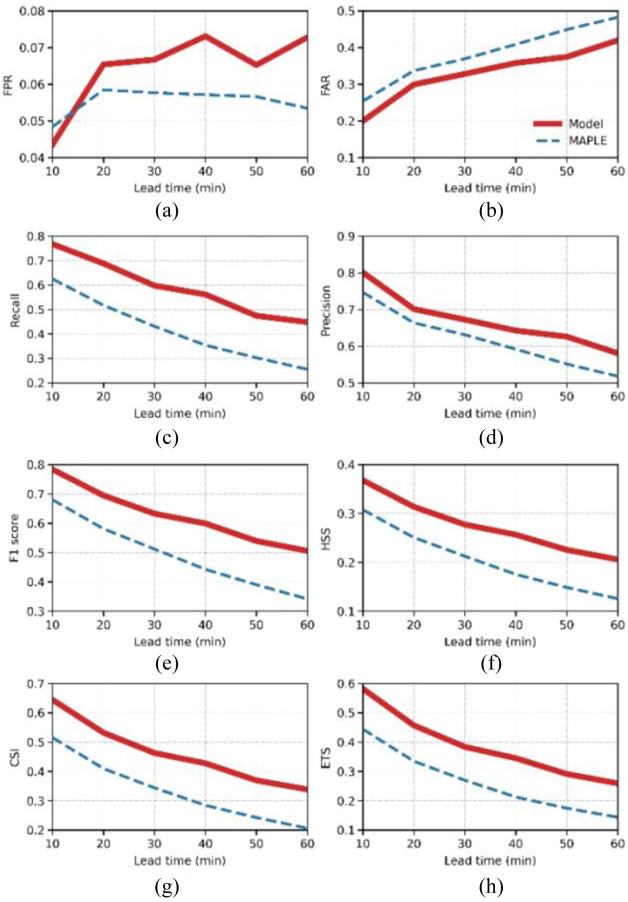


Fig. 6. Comparison of prediction skill scores derived with CLSTM models and the MAPLE for predicting rain rate greater than 30 mm/h with lead times of 10, 20, 30, 40, 50, and 60 min, in terms of (a) FPR, (b) FAR, (c) recall, (d) precision, (e) F1 score, (f) HSS, (g) CSI, and (h) ETS. Note that a higher score indicates better performance, except for the FAR, where a perfect score is 0.

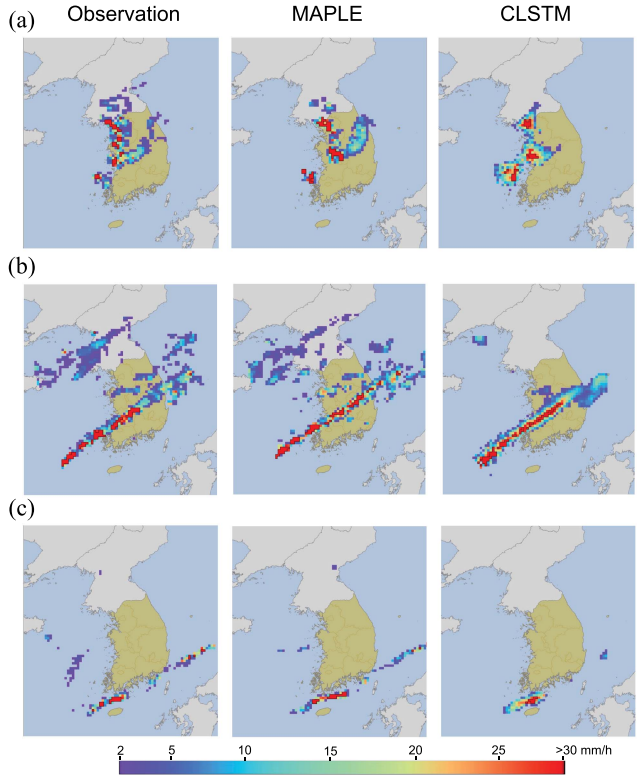


Fig. 7. Visual presentations of observed and predicted geographical distributions of three MCS cases on (a) 2021-5-30 22:00, (b) 2021-08-29 09:00, and (c) 2021-08-28 03:50. Left figures show observations, middle figures are MAPLE predictions, and left figures are CLSTM predictions. Note that solid black lines demarcate the areas with observed or predicted heavy rainfall beyond the rain rate threshold of 30 mm/h.

Fig. 7). This blurring issue is common for rainfall nowcasting with machine learning methods, as it showcases the degrees of uncertainty associated with the predictions. This issue is also consistent with the previous discussion on the higher FPRs of the CLSTM models than of the MAPLE. Despite the blurring effect, the CLSTM models tend to hit the targets when locating the areas with heavy rainfall. Contrarily, although the MAPLE did not produce blurred predictions, the shapes of the distributions of rain rates predicted by the MAPLE looked almost exactly the same across the lead times (middle figures in Fig. 7). They looked almost identical to the observed rain rate distribution at the now moment, albeit with increasing locational differences with longer lead times. Such locational differences led to the poorer predictive performance of the MAPLE than of the CLSTM models in terms of the skill scores.

In addition to the cases of MCSs, as exemplified in Fig. 7, the results of visual presentations for other storm types such as the CC show a similar blurring effect in the rain rate predictions of the CLSTM models. To counter this blurring effect, two filters, i.e., the subtraction filter and the multiplication filter, were attempted to be applied to the nowcasting maps. The subtraction filter subtracted the predicted rain rates with a defined value. The multiplication filter multiplied the rain rate predictions with a specific number between 0 and 1. Neither of these filters produced desirable results, as the subtraction or multiplication operations altered the predicted values, leading to inaccurate predictions. Future effort needs to be made to explore how to reduce the blurring effect within the rain rate predictions of the CLSTM models without sacrificing the accuracy of the model predictions.

V. Discussion

Despite the overall better performance of the proposed CLSTM regression approach than the traditional rainfall nowcasting algorithm MAPLE, there are three issues worth highlighting. First, the proposed approach was validated only on radar products of rainfall estimations, instead of the precipitation records from gauge stations on the ground. For practical purposes, future work needs to collect and compile datasets on rainfall from gauge stations that are compatible with the spatiotemporal resolutions of the presented study to test the predictive performances of the CLSTM models.

The second issue is the need for sensitivity analyses in future works. For example, further studies should explore how

prediction results may be affected by different levels of errors associated with the input variables, such as the errors caused by the retrieval algorithm for rainfall. In addition, future inquiries need to address how the predictive performances of the CLSTM models react to the changes in the values of hyperparameters of the models as well as the size of the training datasets. To take into consideration the input data without proper corresponding output rainfall values, future work also needs to explore the integration of semisupervised machine learning methods into the CLSTM modeling framework [40], [41]. Besides that, the effects of inclusion and exclusion of other input variables, such as satellite radiances from satellite data, on the model performances should also be examined through sensitivity analyses.

Third, like other machine learning methods [43], the DL approach adopted in this study tends to provide good predictive performance regarding interpolation for data points within the general range of the input data. However, its predictive results based on extrapolation for data points outside this general range, in particular the outliers, may be far from being reliable. Along this line, there may also be latent variables, such as topographical and environmental variables, that are uniquely associated with the study area. Cautions and supervisions are needed, therefore, before directly applying the calibrated models to nowcast heavy rainfall for areas with different portfolios of latent variable values in an extrapolative manner. In addition to the latent variables, the presented study was focused on MCSs for training models to nowcast heavy rainfall values, without the consideration of other forms of heavy rain fall scenarios such as tropical cyclones. This focus on MCSs can be justified for the study area around the Korean Peninsula, as MCSs in the warm seasons are the dominating sources of heavy rainfall locally. Because of the sole focus on MCSs in the study, however, additional cautions are needed if the trained models are to be applied to other regions also significantly affected by other forms of heavy rainfall.

VI. Conclusion

The presented study proposed a pixelwise modeling approach to establish CLSTM regression DL-ANNs to nowcast rain rates for lead times from 10 to 60 min, with applications to the MCSs during the warm seasons across the Korean Peninsula. The adopted approach of pixelwise modeling facilitated the stratified sampling processes that enabled the trained CLSTMs to effectively nowcast heavy rainfall without significant underestimations of large rain rate values. The results of the study indicated significantly and consistently better performances of the CLSTM models than the traditional rainfall nowcasting algorithm MAPLE in terms of prediction skill scores across all the considered lead times. These results suggested that the proposed approach to establishing CLSTM models was appropriate for issuing short-term early warnings of heavy rainfall to enhance pre-event preparedness. In addition, the architecture of the CLSTM models was designed to be flexible so that when data on other pertinent input variables became available, the new data could be included for model training. Given the ongoing generation of weather observational data, the CLSTM model developed in this study can be periodically retrained with newly acquired datasets to enhance its performance. In the interest of efficiency and resource conservation, fine-tuning the existing model has gained popularity as a more practical approach compared to training from scratch. However, the inclusion of more input variables may increase the numbers of trainable parameters of the CLSTM models, resulting in significantly longer training time than the already long training time needed for calibrating the CLSTM models separately for different lead times. To shorten the training time of the proposed CLSTM models, future work needs to explore how to improve the design of the model architecture and how to leverage the tools for parallel computing. Besides the long training time, the proposed CLSTM models produced relatively high FPRs and a blurring effect in showing the geographical distributions of rain rates. These issues should also be resolved in future work.

Acknowledgment

The authors would like to appreciate the anonymous reviewers who provided many insightful comments.

References

[1] Y. V. Wang et al., "Relative importance of radar variables for nowcasting heavy rainfall: A machine learning approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 4100314.

[2] Z. Sokol and P. Pešice, "Comparing nowcastings of three severe convective events by statistical and NWP models," *Atmos. Res.*, vol. 93, pp. 397–407, Jul. 2009.

[3] C. J. Short and J. Petch, "Reducing the spin-up of a regional NWP system without data assimilation," *Quart. J. Roy. Meteorol. Soc.*, vol. 148, pp. 1623–1643, Apr. 2022.

[4] R. McTaggart-Cowan et al., "Modernization of atmospheric physics parameterization in Canadian NWP," *J. Adv. Model. Earth Syst.*, vol. 11, pp. 3593–3635, Nov. 2019.

[5] J. T. Johnson et al., "The storm cell identification and tracking algorithm: An enhanced WSR-88D algorithm," *Weather Forecasting*, vol. 13, pp. 263–276, Jun. 1998.

[6] M. Dixon and G. Wiener, "TITAN: Thunderstorm identification, tracking, analysis, and nowcasting—A radar-based methodology," *J. Atmos. Ocean. Technol.*, vol. 10, pp. 785–797, Dec. 1993.

[7] R. E. Rinehart and E. T. Garvey, "Three-dimensional storm motion detection by conventional weather radar," *Nature*, vol. 273, pp. 287–289, May 1978.

[8] J. D. Tuttle and G. B. Foote, "Determination of the boundary layer airflow from a single Doppler radar," *J. Atmos. Ocean. Technol.*, vol. 7, pp. 218–232, Apr. 1990.

[9] C. Mueller et al., "NCAR auto-nowcast system," *Weather Forecasting*, vol. 18, pp. 545–561, Aug. 2003.

[10] U. Germann and I. Zawadzki, "Scale-dependence of the predictability of precipitation from continental radar images. Part I: Description of the methodology," *Monthly Weather Rev.*, vol. 130, pp. 2859–2873, Dec. 2002.

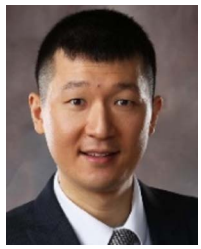
[11] U. Germann and I. Zawadzki, "Scale-dependence of the predictability of precipitation from continental radar images. Part II: Probability forecasts," *J. Appl. Meteorol. Climatol.*, vol. 43, pp. 74–89, Jan. 2004.

[12] B. J. Turner, I. Zawadzki, and U. Germann, "Predictability of precipitation from continental radar images. Part III: Operational nowcasting implementation (MAPLE)," *J. Appl. Meteorol. Climatol.*, vol. 43, pp. 231–248, Feb. 2004.

[13] U. Germann, I. Zawadzki, and B. Turner, "Predictability of precipitation from continental radar images. Part IV: Limits to prediction," *J. Atmos. Sci.*, vol. 63, pp. 2092–2108, Aug. 2006.

[14] B. Radhakrishna, I. Zawadzki, and F. Fabry, "Predictability of precipitation from continental radar images. Part V: Growth and decay," *J. Atmos. Sci.*, vol. 69, pp. 3336–3349, Nov. 2012.

[15] S. Laroche and I. Zawadzki, "Retrievals of horizontal winds from single-Doppler clear-air data by methods of cross correlation and variational analysis," *J. Atmos. Ocean. Technol.*, vol. 12, pp. 721–738, Aug. 1995.
[16] P. V. Mandapaka, U. Germann, L. Panziera, and A. Hering, "Can Lagrangian extrapolation of radar fields be used for precipitation nowcasting over complex Alpine orography?," *Weather Forecasting*, vol. 27, pp. 28–49, Feb. 2012.
[17] A. Bellon, I. Zawadzki, A. Kilambi, H. C. Lee, Y. H. Lee, and G. Lee, "McGill algorithm for precipitation nowcasting by Lagrangian extrapolation (MAPLE) applied to the South Korean radar network. Part I: Sensitivity studies of the variational echo tracking (VET) technique," *Asia-Pacific J. Atmos. Sci.*, vol. 46, pp. 369–381, Aug. 2010.
[18] H. C. Lee et al., "McGill algorithm for precipitation nowcasting by Lagrangian extrapolation (MAPLE) applied to the South Korean radar network. Part II: Real-time verification for the summer season," *Asia-Pacific J. Atmos. Sci.*, vol. 46, pp. 383–391, Aug. 2010.
[19] L. Han, J. Sun, and W. Zhang, "Convolutional neural network for convective storm nowcasting using 3-D Doppler weather radar data," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, pp. 1487–1495, Feb. 2020.
[20] C.-C. Wei, "Soft computing techniques in ensemble precipitation nowcast," *Appl. Soft. Comput.*, vol. 13, pp. 793–805, Feb. 2013.
[21] G. Ayzel, M. Heistermann, A. Sorokin, O. Nikitin, and O. Lukyanova, "All convolutional neural networks for radar-based precipitation nowcasting," *Procedia Comput. Sci.*, vol. 150, pp. 186–192, 2019.
[22] S. Samsi, C. J. Mattioli, and M. S. Veillette, "Distributed deep learning for precipitation nowcasting," in *Proc. IEEE High Perform. Extreme Comput. Conf.*, Waltham, MA, USA, 2019, pp. 186–192.
[23] Q. Yan, F. Ji, K. Miao, Q. Wu, Y. Xia, and T. Li, "Convolutional residual-attention: A deep learning approach for precipitation nowcasting," *Adv. Meteorol.*, vol. 2020, pp. 6484812, Feb. 2020.
[24] S. Ravuri et al., "Skilful precipitation nowcasting using deep generative models of radar," *Nature*, vol. 597, pp. 672–677, Sep. 2021.
[25] S. Choi and Y. Kim, "Rad-cGAN v1.0: Radar-based precipitation nowcasting model with conditional generative adversarial networks for multiple dam domains," *Geosci. Model Develop.*, vol. 15, pp. 5967–5985, Aug. 2022.
[26] G. Ayzel, T. Scheffer, and M. Heistermann, "RainNet v1.0: A convolutional neural network for radar-based precipitation nowcasting," *Geosci. Model Develop.*, vol. 13, pp. 2631–2644, Jun. 2020.
[27] K. Trebing, T. Stańczyk, and S. Mehrkanoon, "SmaAt-UNet: Precipitation nowcasting using a small attention-UNet architecture," *Pattern Recognit. Lett.*, vol. 145, pp. 178–186, May 2021.
[28] L. Han, H. Liang, H. Chen, W. Zhang, and Y. Ge, "Convective precipitation nowcasting using U-Net model," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4103508.
[29] J. Ko, K. Lee, H. Hwang, S.-G. Oh, S.-W. Son, and K. Shin, "Effective training strategies for deep-learning-based precipitation nowcasting and estimation," *Comput. Geosci.*, vol. 161, Apr. 2022, Art. no. 105072.
[30] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2015, pp. 802–810.
[31] X. Shi et al., "Deep learning for precipitation nowcasting: A benchmark and a new model," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2017, pp. 5622–5632.
[32] D. Niu, J. Huang, Z. Zang, L. Xu, H. Che, and Y. Tang, "Two-stage spatiotemporal context refinement network for precipitation nowcasting," *Remote Sens.*, vol. 13, Oct. 2021, Art. no. 4285.
[33] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, pp. 1735–1780, Nov. 1997.
[34] S. Kwon, S.-H. Jung, and G. Lee, "Inter-comparison of radar rainfall rate using constant altitude plan position indicator and hybrid surface rainfall maps," *J. Hydrol.*, vol. 531, pp. 234–247, Dec. 2015.
[35] J. S. Marshall and W. M. Palmer, "The distribution of raindrops with size," *J. Meteorol.*, vol. 5, pp. 165–166, 1948.
[36] *Python Software Foundation*, 2023. [Online]. Available: https://www.python.org/psf/
[37] TensorFlow, "Create production-grade machine learning models with TensorFlow," 2023. [Online]. Available: https://www.tensorflow.org/
[38] D.-A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (ELUs)," 2015, *arXiv:1511.07289*.
[39] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
[40] C. Baur, S. Albarqouni, and N. Navab, "Semi-supervised deep learning for fully convolutional networks," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervention*, 2017, pp. 311–319.
[41] N. Doulamis and A. Doulamis, "Semi-supervised deep learning for object tracking and classification," in *Proc. IEEE Int. Conf. Image Process.*, 2014, pp. 848–852.
[42] A. Tarvainen and H. Valpola, "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2017, pp. 1195–1204.
[43] Y. V. Wang, S. H. Kim, and M. C. Kafatos, "Verifying empirical predictive modeling of societal vulnerability to hazardous events: A Monte Carlo experimental approach," *Rel. Eng. Syst. Saf.*, vol. 240, Dec. 2023, Art. no. 109593.



Yi Victor Wang (Member, IEEE) received the Ph.D. degree in civil engineering with the direction of societal risk management from the University of Illinois at Urbana–Champaign, Champaign, IL, USA, in 2018.

He is currently an Assistant Professor with the Department of Emergency Management, Massachusetts Maritime Academy, Buzzards Bay, MA, USA. His research interests include hazard modeling, machine learning, disaster vulnerability, risk analysis, and media coverage of crises.

Dr. Wang is a Member of the American Geophysical Union, the American Society of Civil Engineers, the Earthquake Engineering Research Institute, and the European Geosciences Union.

Seung Hee Kim (Member, IEEE) received the Ph.D. degree in atmospheric and oceanic sciences from the University of California at Los Angeles, Los Angeles, CA, USA, in 2010.

He is currently with the Institute for Earth, Computing, Human and Observing, Chapman University, Orange, CA. His research interests include dynamics and physics of severe weather and mesoscale convective systems and numerical weather prediction.

Dr. Kim is a Member of the American Geophysical Union and the American Meteorological Society.

Geunsu Lyu received the master's degree in astronomy and atmospheric sciences from Kyungpook National University, Daegu, South Korea, in 2015.

He has conducted research on storm identifying and tracking, precipitation estimation, and forecasting using dual-polarization weather radars. Since 2012, he has been a Researcher with the Center for Atmospheric Remote Sensing, Kyungpook National University. His research interests include radar meteorology and hydrology and radar signal processing.

Choeng-Lyong Lee received the master's degree in radar meteorology from the Department of Astronomy and Atmospheric Sciences, Kyungpook National University, Daegu, South Korea, in 2015.

He is currently a Researcher with the Center for Atmospheric Remote Sensing, Kyungpook National University. His research interests include radar meteorology, microphysics, and dynamical analysis of cloud systems using radar and satellite observations.

Mr. Lee is a Member of the Korean Meteorological Society.

Soorok Ryu received the Ph.D. degree in mathematics from Kyungpook National University (KNU), Daegu, South Korea, in 2010, with a study on the numerical solution of partial differential equations using finite-element method and minimization technique.

She is currently a Researcher with the Center for Atmospheric Remote Sensing, KNU, where she researches the production of high-resolution ground observation data. Her research interests include severe weather and weather prediction using machine learning.

Gyuwon Lee received the Ph.D. degree in atmospheric sciences with focus on radar meteorology and precipitation physics from McGill University, Montreal, QC, Canada, in 2003.

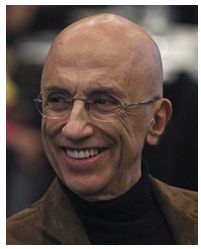
He is currently a Professor of Atmospheric Sciences with the School of Earth System Sciences and the Director of the Center for Atmospheric Remote Sensing, Kyungpook National University, Daegu, South Korea. He was a Research Scientist with the National Center for Atmospheric Research, Boulder, CO, USA, and a Radar Operator with the J. S. Marshall Radar Observatory, McGill University. His research interests include quantitative precipitation estimation, quantitative precipitation forecasting, and precipitation microphysics.

Dr. Lee is the recipient of the Presidential Award from the Korean government and the Mooksan Award from the Korean Meteorological Society for his service to the community. He is a Member of the World Meteorological Organization (WMO) Working Group on Nowcasting and Mesoscale Research and the International Organization for Standardization/WMO Committee on Weather Radar and Wind Profiler.

Ki-Hong Min received the Ph.D. degree in atmospheric sciences with focus on mesoscale and regional climate modeling from Purdue University, West Lafayette, IN, USA, in 2005.

He is currently a Professor of Meteorology with the School of Earth System Sciences, Kyungpook National University, Daegu, South Korea. He is currently interested in high-resolution modeling study of convective storms with radar data assimilation. His research interests include mesoscale modeling, remote sensing, and weather analysis of the atmosphere.

Dr. Min is a Member of the American Geophysical Union, the American Meteorological Society, the Asia-Oceania Geoscience Society, and the Korean Meteorological Society, where he is a Member of the board.



Menas C. Kafatos (Life Member, IEEE) received the Ph.D. degree in physics from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1972.

He is currently the Fletcher Jones Endowed Professor of Computational Physics with Chapman University Orange, CA, USA, where he is also the Director of the Institute for Earth, Computing, Human and Observing. From 2009 to 2012, he was the Founding Dean of the Schmid College of Science and Technology, Chapman University. He is an author, physicist, and philosopher. He works on the environment, climate change and its effects, and natural hazards, including wildfires, pollution, storms, and droughts. He is expert on quantum mechanics, cosmology, measurement theory, and the role of the mind. He is one of few geoscientists who have worked on both earth and space/astronomy, as well as data information systems for both earth and space science.

Dr. Kafatos is a Member of the American Geophysical Union, the International Astronomical Union, the American Association for the Advancement of Science, and the American Physical Society. He is an Honorary Member of the Romanian Academy of Sciences elected in 2000 and a Foreign Member of the Korean Academy of Science and Technology elected in 2018.