Nowcasting Multiparameter Phased-Array Weather Radar (MP-PAWR) Echoes of Localized Heavy Precipitation Using a 3D Recurrent Neural Network Trained with an Adversarial Technique

PHILIPPE BARON, KOHEI KAWASHIMA, DONG-KYUN KIM, HIROSHI HANADO, SEIJI KAWAMURA, TAKESHI MAESAKA, KATSUHIRO NAKAGAWA, SHINSUKE SATOH, and TOMOO USHIO

National Institute of Information and Communications Technology, Koganei, Japan
Osaka University, Suita, Japan
National Research Institute for Earth Science and Disaster Resilience, Tsukuba, Japan

ABSTRACT: We present nowcasts of sudden heavy rains on meso-γ scales (2–20 km) using the high spatiotemporal resolution of a multiparameter phased-array weather radar (MP-PAWR) sensitive to rain droplets. The onset of typical storms is successfully predicted with 10-min lead time, i.e., the current predictability limit of rainfall caused by individual convective cores. A supervised recurrent neural network based on long short-term memory with 3D spatial convolutions (RN3D) is used to account for the horizontal and vertical changes of the convective cells with a time resolution of 30 s. The model uses radar reﬂectivity at horizontal polarization $Z_H$ and the differential reﬂectivity. The input parameters are deﬁned in a volume of $64 \times 64 \times 8 \text{ km}^3$ with the lowest level at 1.9 km and a resolution of $0.4 \times 0.4 \times 0.25 \text{ km}^3$. The prediction is a 10-min sequence of $Z_H$ at the lowest grid level. The model is trained with a large number of observations of summer 2020 and an adversarial technique. RN3D is tested with different types of rapidly evolving localized heavy rainfalls of summers 2018 and 2019. The model performance is compared to that of an advection model for 3D extrapolation of PAWR echoes (A3DM). RN3D better predicts the formation and dissipation of precipitation. However, RN3D tends to underestimate heavy rainfall especially when the storm is well developed. In this phase of the storm, A3DM nowcast scores are found slightly higher. The high skill of RN3D to predict the onset of sudden localized rainfall is illustrated with an example for which RN3D outperforms the operational precipitation nowcasting system of Japan Meteorological Agency (JMA).

SIGNIFICANCE STATEMENT: Temporal extrapolation of radar observations is a means of nowcasting sudden heavy rains, i.e., forecasts with a few tens of minutes and a high spatial resolution better than 500 m. They are necessary to set up warning systems to anticipate damage to infrastructure and reduce the fatalities these storms cause. It is a difficult task due to the storm suddenness, restricted area, and nonlinear behavior that are not well captured by current operational observation and numerical systems. In this study, we use a new high-resolution weather radar with polarimetric information and a 3D recurrent neural network to improve 10-min nowcasts, the current limit of operational systems. This is a first and essential step before applying such a method for increasing the prediction lead time.

KEYWORDS: Algorithms; Radars/radar observations; Nowcasting; Neural networks

1. Introduction

Localized heavy precipitations are generated by sudden convective thunderstorms developing over a small area of about $5 \times 5 \text{ km}^2$ for a few tens of minutes (Kato and Maki 2009). They are caused by individual convective cells that develop aloft in about 10 min (Shusse et al. 2015). They can be responsible for flooding causing severe infrastructure damages, often with fatalities, and their frequency is expected to increase because of climate change (Nakamura et al. 2008; Myhre et al. 2019). The prediction of such storms is essential for developing disaster prevention systems, but this cannot be realized with operational numerical weather prediction (NWP) model (Germann et al. 2006; Surcel et al. 2015; Yano et al. 2018; Kato et al. 2022). Thus, the conventional approach, hereinafter referred to as the advection nowcast model, to dealing with this problem is to extrapolate in real time the precipitation features identified in the radar observations. The extrapolations follow motion vectors, themselves extracted from the same radar measurements with optical flow techniques (Otsuka et al. 2016; Bechini and Chandrasekar 2017). However, the limit of predictability of localized heavy precipitation with advection nowcast models is less than 10 min. This is the typical lifetime of an individual convective cell, i.e., the time between the formation of the first rain droplets and the onset of heavy rain. Furthermore, current operational systems have additional limitations: 1) Nonlinear processes involved in the growth and dissipation of convective rains are in general not properly accounted for (Pulikkien et al. 2020). 2) Radar temporal resolution, typically 5 min, is not sufficient to follow the rapid evolution of the convective cell causing the rainfall (Isoda et al. 2018; Honda et al. 2022a). 3) Observations are only extrapolated from a single height although the vertical dimension is critical for the description of the...

Under the Japanese Cross Strategic Innovation Promotion Program (SIP), a series of X-band phased array weather radars (PAWR) were developed to observe the whole 3D rain field surrounding the instrument with a time resolution of 30 s (Ushio et al. 2022). The latest version of the instrument is a multiparameter PAWR (MP-PAWR) that has been operational since 2018 at Saitama University, Japan (Kikuchi et al. 2020; Asai et al. 2021; Yoshimi et al. 2021). It transmits 2 polarizations, an essential improvement for accurate precipitation analysis (P.C. et al. 2016; Augros et al. 2018; Gou et al. 2019; Tromel et al. 2021). For nowcasting of rapidly forming localized rains, the advantage of observations with high spatiotemporal resolution is highlighted in several studies. For example, an advection model designed to extrapolate PAWR echoes in 3D space has better performance than conventional 2D methods (Otsuka et al. 2016). The Riken Institute (Japan) presented an experimental real-time NWP system updated every 30 s with MP-PAWR observations that outperforms the operational system of Japan Meteorological Agency (JMA) updated every 5 min with conventional weather radar observations, and this despite the fact that the JMA system takes into account a 3D atmosphere and processes of initiation and dissipation of convective rain (Honda et al. 2022a).

Supervised neural network (DNN) has been recognized as a promising numerical method for radar echo extrapolation since it has the capability to handle the nonlinear behavior of the convective cells and can be applied in real time (Nayak et al. 2013). Typically, a DNN can extract spatiotemporal features from high-dimensional data and be trained with past observations without any description of the underlying physical processes. The interest for such a technique has strongly increased with recent improvements coming from the fields of video prediction and medical imaging (Milletari et al. 2016; Oprea et al. 2022). Most of the nowcast models use 2D or 3D convolutional neural networks (CNN) and time recurrent networks such as long short-term memory (LSTM) units (Klein et al. 2015; Akbari Asanjan et al. 2018; Han et al. 2020; Kim et al. 2021; Yao et al. 2022). Recently, 2D spatial convolutions have been included in LSTM (ConvLSTM2D) to successfully nowcast precipitation (Shi et al. 2015) and has become a popular approach (Shi et al. 2017; Tran and Song 2019; Jing et al. 2019; Kim and Ushio 2022). However, we should note that it is reported that an LSTM could be less efficient than CNN for long time series (Ayzel et al. 2020) and new techniques are already being proposed such as transformer architecture (Bertasius et al. 2021). Furthermore, it has been shown that a CNN2D-based U-Net allows for robust nowcasts of convective rain while being a much lighter model than LSTM-based ones (Han et al. 2022a).

Improvements in training methods have also led to significant increases in performance. Adversarial training developed for generative adversarial network (GAN) and its conditional variant (CGAN) (Goodfellow et al. 2014; Mirza and Osindero 2014) helps to reduce the well-known blur effect inherent with pixelwise $l_p$-norm loss functions (Oprea et al. 2022). This effect smooths out spatial details and therefore reduces the spatial resolution of nowcasts. With a GAN, the model is optimized to fool a discriminator trained to recognize predicted data from actual observations. Unlike $l_p$-norm evaluation method, predictions and observations are not compared pixel by pixel, but globally as in a sort of intensity-distribution matching method. Various strategies have been proposed. The discriminator can be used as a penalty of a mean-squared error (MSE) loss (Jing et al. 2019) or in two-step procedure, where a first model is optimized with an MSE loss to generate intermediate blurred predictions, while a second model optimized with a discriminator loss refines the results (Wang et al. 2021). Ravuri et al. (2021) uses the random latent spaces of a CGAN to generate an ensemble of predictions. The training of their generator is driven by two discriminator-based losses, which analyze time and space features, respectively.

Most of the DNNs cited above consider a single altitude with meso-$\beta$-scale (scale $>100$ km) and conventional radars. They are therefore not optimal for nowcasting localized heavy rainfall. Jing et al. (2019) considered 3 altitudes between 2 and 3 km for the input and forecasts the radar echo at the middle level. The vertical dimension is treated as different channels of convLSTM2D units, similar to RGB colors in video analysis. This approach is well suited to exploit the local spatiotemporal correlations in a narrow vertical range, but it is not satisfactory for the full vertical range observed with MP-PAWR (Baron et al. 2021a). Consequently, for nowcasting heavy rains on meso-$\gamma$ scale (length $<20$ km) with PAWRs, new approaches are being studied on the basis of previous 2D models. They are designed to fully exploit the high spatiotemporal resolution of PAWRs, especially the vertical range. Kim et al. (2021) feeds a CNN3D with PAWR 4D data and nowcasts from a 3D linear advection model to extrapolate the 3D structure of the precipitation. In other studies, convLSTM2D units are organized in an architecture adapted to account for the vertical dimension (Ushio et al. 2022). Another strategy consists in replacing the 2D spatial convolutions in convLSTM2D with 3D convolutions (Otsuka et al. 2020; Baron et al. 2021a) to include the vertical range. Typically, these studies consider high horizontal resolution (better than 400 m), heights above 10 km, and lead times of 10–20 min. Overall, the DNN-based methods outperform 3D advection nowcast model but are limited for high rainfall intensity and lead time of 10 min or more.

In this study we analyze 10-min nowcasts of localized heavy rains using MP-PAWR and the latest version of the 3D recurrent neural network (RN3D) proposed by Baron et al. (2021a). Four periods with different types of localized rainfalls are considered to evaluate the model performance. Unlike other studies with PAWR, polarimetric information is considered and its influence on the precipitation nowcasts will be discussed. The model is trained using several months of observations to produce a robust model applicable to any periods while most of other DNNs for PAWR use training data within a few hours range close to the predictions. The first version of RN3D showed promising performance against convLSTM2D-based models (Baron et al. 2021a,b) and the version presented in this study is improved by using an
adversarial training technique, convLSTM3D units instead of gated recurrent units (convGRU3D), and a new training dataset.

The paper is divided as follows. Section 2 presents details of the analysis method including a description of MP-PAWR and the parameters used for the nowcasts, as well as the 3D advection model used to evaluate RN3D. Section 3 describes RN3D with a special focus on the adversarial training. Section 4, the model performance is evaluated. The benefits of the adversarial training are assessed, and the ability of the model to predict sudden and localized heavy rainfall is discussed based on 3 case studies. In section 6, we discuss the influence of radar parameters on precipitation nowcasting. Section 7 concludes by summarizing the main results and presenting future work to improve the nowcasts.

2. Method

2.a. MP-PAWR parameters

We only use data measured with MP-PAWR and provided by the National Research Institute for Earth Science and Disaster Resilience (Japan). The radar operates in the X-band frequency range (9425 MHz) and combines mechanical horizontal scanning with electronic vertical scanning. In the observation mode used in this study, the whole surrounding atmosphere is scanned every 30 s to a range of 60 km. The volume is sampled with 114 elevations from 0° to 90°, an azimuthal resolution of 1.2° and a range resolution of 75 m. For this study, the radial observations are interpolated into a regular Cartesian grid of 68 × 68 × 8 km³ (resolution 0.4 × 0.4 × 0.25 km³) with a footprint centered on MP-PAWR. The altitude of the lowest level is 1.875 km, but it will be approximated by 1.9 km in the discussion. The full observation range (60 km) is not considered to avoid problems due to spatial resolution degradation and signal attenuation with increasing range.

We use two radar parameters that we believe are relevant for the nowcasts (Table 1). The radar reflectivity $Z_H$, the power of the signal measured at the horizontal polarization, is a proxy of the precipitation intensity and, hence, the key parameter for the nowcast. It is expressed in dBZ as $10 \log_{10}(Z_H)$ and no correction for rain attenuation is applied. This is a similar approach to previous studies with PAWR (Otsuka et al. 2016; Kim et al. 2021; Yoshimi et al. 2021). The second parameter is the differential reflectivity defined as $Z_{DR} = 10 \log_{10}(Z_H/Z_V)$, with $Z_V$ being the intensity at the vertical polarization. It is sensitive to the shape of the scatterers at low elevation angles. Raindrops, which are horizontally oriented, have positive values that increase with the size of the droplet. On the other hand, ice particles are associated with negative values (Homeyer and Kumjian 2015; Mattos et al. 2016). For instance, a strong convective updraft of raindrops above the melting layer can be identified by the presence of a vertical column of large positive $Z_{DR}$ (Tamanachi and Heinselman 2016).

Two other radar parameters could have been used in this study. Doppler velocity ($V_{dpb}$) of the scatterers was originally included (Baron et al. 2021a,b), but, as we will show in section 6, it has no impact on the nowcasts. This is probably because a CNN-based model assumes translation invariant features and is therefore not suitable for analyzing $V_{dpb}$, which is defined with respect to the instrument. The other unused parameter is the specific differential phase $K_{dp}$, which is a measure of the phase difference between the two polarizations. It allows for more accurate estimation of high rainfall intensities compared to $Z_H$ (Augros et al. 2018), and its use will be further investigated in a future study.

Figure 1a shows $Z_H$ for convective rains observed at 1.9-km height over a continuous period of 1.5 h [1520–1700 Japan standard time (JST)]. Occultations by buildings can be noticed in the upper-left corner (gray areas). In the first hour, several isolated convective rainfalls are visible with intensity larger than 40 dBZ (=15 mm h⁻¹). At 1550 JST, two cores occur in less than 10 min (center of the black window). They merge and grow until 1640 JST to form a band of mixed stratiform and convective precipitations that dissipates in about 30 min. The moving target indicator (MTI) procedure is used to remove unwanted clutter, although some clutter is still visible at certain periods such as that presented in these examples.

2.b. Training and validation data

Training data are gathered between May and September 2020 (Table 1), except three days in July used as validation data to select the best model state. We consider a 20-min-long sequence with 0.5-min time step. The first half is used for model input, and the second half is used as true data for nowcast error assessment.

First, periods with rainfall are manually selected except if problems are noticed with the measurements. Cases with heavy rain widely spread over the full area are also rejected because of strong attenuation of the signal for ranges longer than 5 km (typhoon events fall in this case). Observations with only small stratiform rains are partly rejected to reduce their statistical weight in the database (1 event of 4 is kept).

Then, the training database of 20-min-long sequences is built (30 s time steps) with the following rules: 1) The two radar parameters must be available. 2) No missing time steps in the sequence is accepted except if only one isolated data is missing. In that case, the gap is filled by averaging the time

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Considered range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radar reflectivity</td>
<td>$Z_H$ (dBZ)</td>
</tr>
<tr>
<td>Differential</td>
<td>$Z_{DR}$ (dB)</td>
</tr>
<tr>
<td>reflectivity</td>
<td>$Z_V$ (dB)</td>
</tr>
<tr>
<td>Cartesian grid</td>
<td>Horizontal range (km)</td>
</tr>
<tr>
<td></td>
<td>Vertical range (km)</td>
</tr>
<tr>
<td></td>
<td>$[1.875, 9.625]$ (resolution 0.25)</td>
</tr>
</tbody>
</table>

Table 1. MP-PAWR parameters used in the study. Time is given in Japan standard time (JST). The instrument location is 35.8615°N, 139.6090°E. The associated datasets are 1) the training data, which cover May–October 2020 except during periods reserved for the validation; 2) validation data, which extend from 1600 to 2000 JST 20 Jul, from 1630 to 1900 JST 22 Jul, and from 0700 JST 23 to 0000 JST 24 Jul 2020; and 3) test data, which cover July–October 2018 and April–October 2019.
steps just before and after. 3) Two consecutive sequences must be separated by at least 2 min and intermediate sequences are removed.

To increase the ratio of pixels with high $Z_H$ and to minimize the GPU memory usage during the training, the observed volume is divided into smaller ones of $16 \times 16 \times 8$ km$^3$ ($40 \times 40 \times 32$ voxels). Sequences of reduced volume without rain are rejected. Rain detection is positive if, in the middle of the sequence, at least 1.5% of the pixels at one of the heights of 1.9, 4, and 6 km have an intensity greater than 15 dB$Z$. Figure 2 shows daily distribution of selected sequences. A total of about 29,100 samples are selected within 37 days between May and September. A sample corresponds to a sequence of reduced volume.

c. Reference models

The nowcast skills of RN3D are evaluated against two other methods. The first one is a persistent nowcast model (PERS). It is the most basic prediction since the atmospheric state is considered constant after the last observation. Such a nowcast sets the minimum predictive skill that should be achieved by RN3D. For example, a high predictive score would not be so relevant if it is smaller than that of PERS.

The second model is a 3D advection nowcast model (A3DM) already described in the introduction (Otsuka et al. 2016). It is based on the 2D method named Continuity Of Tracking Radar Echoes by Correlation (COTREC) (Li et al. 1995). A3DM also includes noise reduction techniques optimized for PAWRs’ 30-s sampling. Typically, it can detect a convective cell at any altitude and extrapolate its future position following a constant motion vector, but cannot account for its initiation or dissipation. The implementation is the one used by Kim et al. (2021). The same team produced the A3DM nowcasts of this study using their own MP-PAWR dataset (24 and 29 July 2018). The A3DM data are defined in a volume of $40 \times 55 \times 10$ km$^3$ with a resolution of 100 m. The
horizontal grid extends from −15 to 25 km in the west–east direction (x axis) and from −20 to 35 km in south–north direction. For the comparison with RN3D, the high resolution of the predicted maps is degraded to match that of the RN3D predictions (400 m along x and y, and 250 m along z), and only the predictions at the altitude of 1.9 km are used. Then, A3DM scores are calculated with the same MP-PAWR data used to produce the A3DM nowcasts. This procedure ensures that A3DM is correctly manipulated and that no unnecessary errors are added in the assessment of nowcast skills. There are indeed small differences between their data and those used to train RN3D due to different gridding methods and the use of an additional Gaussian filter to further reduce clutters of A3DM input.

d. Nowcast scores

Usual quantitative scores for assessing nowcasting skills are probability of detection (POD), false alarm rate (FAR), critical success index (CSI), and Heidke skill score (HSS) (Shi et al. 2015; Otsuka et al. 2016; Kim and Ushio 2022; Yao et al. 2022). The CSI value is a pixelwise measure of the model ability to predict whether rainfall intensity will be above a threshold $Z_{TH}$. It is calculated as follows:

$$CSI(Z_{TH}) = \frac{N_{TP}}{N_{TP} + N_{FP} + N_{FN}},$$

where $N_{TP}$, $N_{FP}$, and $N_{FN}$ are the number of pixels identified as true positive, false positive, and false negative, respectively (examples will be given in section 4a). The thresholds of 10 and 37 dBZ have been chosen to be consistent with previous studies (Kim et al. 2021; Ushio et al. 2022). Data with $Z_H > 10$ dBZ ($\sim 0.05$ mm h$^{-1}$) include all rain events and $Z_H > 37$ dBZ ($\sim 10$ mm h$^{-1}$) the heavy convective ones only. The CSI combines the POD and FAR, which are defined as

$$POD(Z_{TH}) = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad \text{and}$$

$$FAR(Z_{TH}) = \frac{N_{FP}}{N_{TP} + N_{FP}}.$$

HSS is a complementary score to CSI because it is more sensitive to spatial mismatches between observations and nowcasts. It is defined as

$$HSS(Z_{TH}) = \frac{2(N_{TP}N_{TN} - N_{FN}N_{FP})}{(N_{TP} + N_{FN})(N_{FP} + N_{TN}) + (N_{TP} + N_{FP})(N_{FP} + N_{FN})}.$$

Note that, unlike other scores, HSS can be negative.

3. RN3D model

a. Model description

RN3D is a 3D multiparameter sequence-to-sequence model (Fig. 3a). The input consists of volumetric sequences of $Z_H$ and $Z_{DR}$, and the output is the predicted sequence of $Z_H$ at 1.9-km height. The input sequence is 10-min long (0.5 min sampling, 20 time steps), a duration chosen because it is the typical time range to generate sudden localized storms (see examples in sections 4 and 5).

The model uses LSTM units in which the scalar products are replaced by convolution functions as in a CNN (Shi et al. 2015). To analyze the temporal variation of 3D spatial features, 3D convolutions are used (convLSTM3D) instead of 2D ones used in most other studies. This is a very natural approach to describe 3D spatial observations, but it considerably increases the number of operations compared to 2D convolutions. It also inherits the limitation of CNNs due to their location invariant properties (Shi et al. 2017).

The convLSTM3D units are organized in an encoder-decoder architecture (Fig. 3b). Each unit has 64 filters of size $3 \times 3 \times 3$. The encoder fuses $Z_H$ and $Z_{DR}$, and detects spatiotemporal features in the input sequence. The decoder extrapolates $Z_H$ into the 4D space–time from encoded features. The encoder–decoder has four layers with two convLSTM3D units per layer. Details are given in appendix B. The horizontal receptive field (Le and Borji 2017; Luo et al. 2017) increases from 2 km for an output pixel of the upper layer (first layer of the encoder) to 5.2, 11.6, and 24.4 km for the next three layers. Regarding the vertical dimension, the receptive fields of the output pixels are 1.25, 3.25, 7.25, and 15.25 km.

A CNN-based module (postdecoder module) refines the decoder output to make the predicted sequence at 1.9-km height (see appendix C). A shortcut (He et al. 2015) connects the last $Z_H$
a) Overview

Scan of the input sequence (at time $T$) to the output. Therefore, in order to achieve the simplest prediction of a persistent model, the main model should simply return zero. Note that the model will be adapted in the future for nowcasting the 3D structure of the convective core, by simply changing the postdecoder module.

b. Discriminator and adversarial training

The training is adapted from the adversarial method presented by Jing et al. (2019). RN3D acts as the generator of a sequence of $Z_H$ at the altitude of 1.9 km. The discriminator $D$ is a standard CNN classifier. It returns the probability that a time series of $Z_H$ images is real and not a predicted one. The sequence is composed of 20 images of size of $16 \times 16$ km$^2$ (40 $\times$ 40 pixels) as shown in Fig. 4. The model is made of four 3D CNN layers with 64, 128, 256, and 512 filters, respectively, and stride of 2 and 3 $\times$ 3 $\times$ 3 kernels. A fully connected layer feeds a sigmoid activation function to generate a single probability value between 0 and 1.

Compared to Jing et al. (2019), 3D convolutions are used instead of 2D convolutions to detect features in both temporal and spatial dimensions.

The discriminator is optimized with a binary cross entropy loss function $L_D$ defined as

$$L_D = -(l_k \log[D(x^*_{k}/y_k) + \epsilon] + (1 - l_k) \log[1 - D(x^*_{k}/y_k) + \epsilon],$$

(4)

Discriminator

Evaluates if a sequence of $Z_H$ at 1.9 km is real observation

**Real sequence (20 min, 1 min sampling)**

- Observation ($x$)
  - T -9 min
  - T +1 min
  - T +10 min

- Observation ($y$)
  - T -9 min
  - T +1 min
  - T +10 min

$D(x//y) \rightarrow 1$

**Fake sequence**

- Observation ($x$)
  - T -9 min
  - T +1 min
  - T +10 min

- Prediction ($y_p$)
  - T -9 min
  - T +1 min
  - T +10 min

$D(x//y_p) \rightarrow 0$

FIG. 4. Schematic representation of the discriminator $D$ trained to differentiate a sequence of true observations at 1.9-km height from predictions. The discriminator is a 3D convolution neural network for binary classification. The number of time steps of the input sequence is $T = 20$, and the number of pixels is $40 \times 40$. 
where // is the concatenation operator, angle brackets denote the average over a set of data \( \{ (x^*_k, y_k, l_k) \} \), for which \( x^*_k \) and \( y_k \) are ZH at 1.9 km from the model input (observations) and nowcasts, respectively. The parameter \( l_k \) is the sequence label such that \( l_k = 1 \) if \( y_k \) is actual observation or \( l_k = 0 \) otherwise. The parameter \( \epsilon \) has a value of \( 10^{-6} \) to prevent zero value in the logarithmic function.

A penalty on the nowcast model loss is added as

\[
L_p' = \text{MAE}(y_p - y) - \alpha \log \left( \frac{D[x^*_k/P(x_k)]}{\log \epsilon} \right),
\]

where the first term is the MAE assessed with \( T_o = 10 \) predicted images of size \( (X, Y) \) and the second term is the penalty computed with BCE in which the label of the sequence is set to 1 though the given sequence includes the predictions (fake sequence).

The discriminator is trained in parallel to RN3D (Fig. 5). As iterations progress, the ability of the discriminator to differentiate between small differences between the prediction and the observation increases, forcing the RN3D to produce predictions that look increasingly realistic. The model convergence is very sensitive to the value of \( \alpha \) [Eq. (5)] and the frequency of updating the discriminator trainable parameters. After trying various settings, we finally set \( \alpha = 0.002 \) and the frequency update to 1/5 (the discriminator is updated after 5 iterations).

The values of the radar parameters are clipped into the ranges given in Table 1 and normalized as

\[
p_i = \frac{(p - \min) - \max}{\max - \min},
\]

where \( p \) denotes one of the input parameters, and \( \min \) and \( \max \) are the boundary of the considered range. The trainable parameters are optimized using the stochastic gradient descent algorithm (see appendix A) with a batch of 8 samples per iteration. The model is evaluated with the validation data (Table 1) after iterating over 320 samples using an MAE loss function.

Figure 6a shows validation loss functions with respect to the number of epochs. One epoch is set to 5000 samples, i.e., the approximate number of independent data in the training database. The samples randomly change from one iteration to another to use the whole training dataset (29144 samples).

Without adversarial training, a clear negative trend is observed until epoch 250, while, with adversarial training, the loss function decreases until epochs 100–150 before rising again. The lowest MAE is found for the training without adversarial technique (0.035 vs 0.040), which is expected since the model is optimized to minimize the MAE while the adversarial method introduces a penalty that moves the convergence away from this limit. Therefore, it does not necessarily mean that the model without adversarial training outperforms the one with. The plateau of the adversarial loss is not due to overfitting because it is also noticed with training data.

The loss minimum is located near epoch 140. Finding the minimum of the loss function is quite ambiguous since the function is noisy. Note that the noise is only due to change of the model state because the same validation samples are used for each evaluation. Therefore, the state with the highest 37-dBZ CSI [Eq. (1)] in the vicinity of the loss minimum (epochs 100–180) is selected. It is expected to be best suited for heavy rainfall nowcasting. The two best scores are almost

![Fig. 5. Flowchart of the training operations at an iteration step: (a) RN3D loss from MAE, (b) RN3D loss from the discriminator, (c) D loss, (d) update of RN3D, and (e) update of D.](image-url)
identical (Fig. 6b) and the first one (epoch 168) is chosen for this study because it has the higher 25-dBZ CSI of the two. Note that although the selection rules were defined independently of test data, the choice of the model is, however, not entirely free of subjectivity. The selected state depends on choice of metric and validation dataset. For example, the state at epoch 131 would have been selected if only the 25-dBZ CSI had been taken into account.

4. Nowcast performance

a. Adversarial training

Figure 1 shows 10-min nowcasts with RN3D trained without and with adversarial technique. The location, intensity, and behavior of short-lived isolated convective rains are well predicted by both versions. This is well illustrated by the nowcasts of two convective cores at 1600 JST that are rapidly formed in less than 20 min (middle of the black window). Similar results are also found for the gradual dissipation of the rainband after 1640 JST. However, it is clear that the fine structures observed in the rain field are smoothed out when the adversarial technique is not used. The ability to reproduce such small features with adversarial technique is clearly highlighted with the nowcast of clutters seen in the bottom-right corner of the black window while they are smoothed out with the nonadversarial method.

Figure 7 shows the pixel-to-pixel comparison of the nowcasts with the observations (ground truth) between 1500 and 1725 JST. Without adversarial training, the model performs well between 10 and 30 dBZ, but it significantly underestimates high intensities due to blurring. The model with adversarial training fits the observations better although high values of ZH are still underestimated. Regarding, for example,
true $Z_H$ of 40 dBZ, the nowcasts without adversarial training span between about 12 and 42 dBZ with a median value of 32 dBZ, but with adversarial training, the median is improved to 36 dBZ and the range is moved up to 16–48 dBZ.

Figure 8 shows the nowcast scores (section 2d) for the period of Fig. 1. The 10-min-lead-time nowcast corresponds to the last time step of the sequence predicted by RN3D. For the threshold of 10 dBZ, the CSI and HSS scores of both models are very similar and higher than the persistent model (PERS). The mean CSIs are 0.64 and 0.5 for RN3D and PERS, respectively, and 0.70 and 0.56 for HSS. RN3D with adversarial training (ADV) has slightly higher FAR than that without ADV (mean value of 0.28 and 0.19, respectively), which is compensated by a better POD for the model with ADV (0.83 versus 0.76).

For heavy rainfalls (threshold of 37 dBZ), the scores with adversarial training are significantly better than those without, which is consistent with Fig. 7. The mean CSIs of RN3D and PERS are 0.32 and 0.18, respectively, and 0.44 and 0.32 for HSS. Except in the middle of the period, only the nowcasts with ADV significantly outperform those with PERS, which is due to better POD scores than RN3D without ADV.

In conclusion, the adversarial technique helps generate more realistic images with better occurrence of high-intensity pixels. However, it provides only one state among the set of possible states while the network without adversarial training produces the average state. Both models provide complementary information and, though the model with adversarial training will be preferred in this study, the nonadversarial model could be used in the future to infer uncertainties on the predicted location of heavy rains.

b. Comparison with 3D advection model

The RN3D with adversarial training is compared with the 3D advection model (A3DM) (section 2c). Two time series of about 1 h showing different types of localized rains are considered. Figure 9 shows nowcasts for the first period on 24 July 2018. Rapidly evolving localized heavy convective rainfalls can be seen in different phases of their evolutions. The first precipitation dissipates after 30 min (2040–2110 JST). A second rainfall event appears at 2055–2100 JST and expands until the end of the time series (2135 JST). For the second event, rain intensity reaches values larger than 50 dBZ (70 mm h$^{-1}$) between 2110 and 2130 JST at $(x, y) \sim (0 \text{ km}, 5 \text{ km})$. Both models predict correctly the main characteristics of the rainfalls in terms of location, coverage, and intensity. However, the small spatial structures are not well reproduced. The quality of the RN3D nowcasts looks lower than that obtained on 1 August 2020 although similar convective rains are observed. A3DM tends to overestimate the moderate rainfall, which has also been reported in previous studies (Kim et al. 2021), and this is likely because it cannot account for rain dissipation. The formation of the precipitation at 2055–2100 JST is predicted by RN3D but missed by A3DM (upper side in the black window of Fig. 9).

The pixel-to-pixel comparison between nowcast and ground truth (Fig. 10) looks at first glance similar for both models. The median nowcasts are close to ground truth up to 30 dBZ,
but beyond that, it saturates at 30 and 32 dBZ for RN3D and A3DM, respectively. However, the standard deviation of the nowcast errors is larger for A3DM (15 and 12 dBZ for A3DM and RN3D, respectively) and the tendency of A3DM to overestimate moderate precipitation is clearly noticeable, especially between 15 and 30 dBZ.

**Figure 11** shows the nowcast scores with respect to time. Over this period A3DM has slightly higher scores for the

---

**FIG. 9.** Observations and 10-min-lead-time nowcasts of $Z_H$ at 1.9-km height for 24 Jul 2018: (a) MP-PAWR, (b) PERS, (c) RN3D, and (d) A3DM. The color levels are defined in Fig. 1.

**FIG. 10.** As in Fig. 7, but for nowcasts from (left) RN3D and (right) A3DM with data of 24 Jul 2018.
10-dBZ CSI (0.53 and 0.55 for RN3D and A3DM, respectively). These scores are significantly larger than those of PERS (0.48). Similar results are found for HSS although the difference between the RN3D and A3DM models is more pronounced (mean value of 0.60 and 0.65 for RN3D and A3DM), especially in the second half of the period. Probably, the precipitation locations predicted by RNN3D have larger errors than A3DM when storms are well developed.

For the heavy rainfall (37-dBZ threshold), the CSI and HSS scores of the 3 models are low (0.13–0.14 for the CSI and 0.19–0.21 for HSS). Three different periods can be noticed. Until 2055 JST, only a rainfall in its decreasing phase is observed (near $y = -10$ km). All models have a bad FAR close to 1. However, the number of pixels with intensity of $>37$ dBZ is few, and the relevance of the scores might be limited. The period 2055–2110 JST corresponds to the onset of heavy rainfall near $y = 0$ km. The CSI score of RN3D (0.4) is significantly larger than those of A3DM and PERS (<0.3), mostly due to better FAR. After 2110 JST, the storm is well developed and covers most of the studied area. A3DM score is better, which is due to a clear underestimation by RN3D of high intensity leading to lower POD and upper FAR.

For the second test case (Fig. 12), the precipitation characteristics are very different. The intensities are most of the time $< 40$ dBZ indicating much weaker convection than in previous cases, and a strong horizontal advection is noticeable. Two bands of precipitations cross the observed area northeastward between 0120–0145 and 0145–0210 JST, respectively. Both models capture the horizontal advection as well as the rainfall spatial range. A3DM seems to overestimate moderate intensity and underestimates heavy ones. As for the previous case, fine spatial structures are more visible in the RN3D nowcasts than in A3DM ones.

Figure 13 shows that A3DM is on average closer to ground truth than RN3D. However, the results should be taken with caution for $Z_H > 37$ dBZ because of the small number of representative pixels.

However, the statistical scores (Fig. 14) show a different picture. Those of RN3D for precipitation larger than 10 dBZ are most of the time greater than those of other models (average CSI of 0.52, 0.42, and 0.34 for RN3D, A3DM, and PERS, and average HSS of 0.56, 0.48, and 0.34). The significant discrepancy between RN3D and A3DM is primarily due to the low scores of A3DM during the dissipating phase of the first rain event between 0140 and 0155 JST, because of a combination of high FAR (>0.6 between 0140 and 0150 JST) and low POD (<0.3 between 0145 and 0155 JST). However, if we ignore these periods, A3DM predictions look better than RN3D predictions though smoother. It is expected since the changes in the precipitation pattern are mainly due to advection.

For heavy rainfalls (threshold of 37 dBZ), very low average CSIs of 0.05–0.06 are obtained for RN3D and A3DM, but greater than those of PERS (0.02). This is due to large FAR close to 1 over most of the period, indicating an overestimation of the spatial coverage of heavy rainfall predictions. Very similar results are found for HSS. The RN3D scores show noticeable improvements in comparison to both A3DM and PERS between 0155 and 0205 JST when the second rainfall event occurs within the studied area (HSS of 0.2 and 0 for AM3D and PERS). The bad scores of AM3D are probably due to its tendency to smooth the rain patterns in its
prediction leading to overestimate the spatial range of the rainfall (high FAR) and underestimate its intensity (low POD).

To summarize, the statistical scores evaluated over the full time ranges do not allow us to clearly identify whether one model is better than the other. However, the RN3D shows better nowcast skills when the rainstorm is formed or dissipated. It also better reproduces fine spatial structures.

Fig. 13. As in Fig. 10, but for data of 29 Jul 2018.
5. Sudden rainfall

a. Description of a typical case

Figure 15 shows the 3D evolutions from $T = 2047$ JST of convective cells responsible of the heavy rain observed in Fig. 9 ($y = 0$ km) at 2105 JST. Between $T$ and $T + 2$ min, three convective cells appear at the altitude of 4 km. The one at $y = 2$ km grows and rises up to 5–8 km between $T + 4$ and $T + 8$ min. Then, the core keeps on growing and descends to produce heavy rainfalls at the height of 2 km after $T + 10$ min. A new convective cell is generated at $T + 10$ min near 6 km that enhances the rainfall intensity from $T + 16$ min.

This illustrates the need of considering the 3D space with high spatiotemporal resolution as well as the growth and dissipation of rains due to thermodynamical processes. For RN3D, the nowcast of the increase in rainfall from 2058 to 2103 JST is satisfactory, with only minor mismatches in spatial coverage (Fig. 16).

On the other hand, the inability of A3DM in accounting the thermodynamic growth of rain prevents it from nowcasting the rainfall onset and enhancement. Like PERS, it cannot predict precipitation in the black window of Fig. 16a at 2058 JST (initial time 2048 JST). However, it is able to predict those at 2103 JST by extrapolating downward the rain pattern observed at nearly 5 km height in the initial state (Fig. 15, $T = 2053$ JST) but with a significantly underestimation of its intensity. Like the cases presented in the previous sections, A3DM smooths the rain pattern and the reasons for this smoothing should be examined in future studies. The sensitivity of A3DM to the initial state in nowcasting can be noticed through the significant discrepancy in intensities between nowcasts initialized at 2053 JST (Fig. 16b) and 2055 JST (Fig. 9), which are not observed with MP-PAWR.

b. Storm onset

In this section we focus on the onset of three localized sudden storms observed on 24 July 2018 (Fig. 9) and 1 (Fig. 1)
and 25 August 2019, respectively. Unlike the first two days, the latter has not been described in the previous sections but is studied in detail by Honda et al. (2022a). Figure 17 shows the observed and predicted precipitation maxima in an area of $5 \times 5$ km$^2$ around the rainfall cores. All of the events are displayed over a period of 50 min. The weather quickly changes from no or very light rain to heavy one in less than 15 min and convective processes dominate horizontal advection.

Fig. 16. Observations and nowcasts of $Z_H$ at 1.9-km height of the rains in Fig. 15 are shown in the black windows. The initial times of the nowcasts are $T = (a)$ 2048 and (b) 2053 JST, and the lead times are $+2$, $+5$, and $+10$ min. The nowcasts are computed with PERS, RN3D, and ADM3D. The color levels are defined in Fig. 1.

Fig. 17. Time variation of maximum precipitation in rapidly forming localized convective rains observed on (a),(b) 24 Jul 2018; (c),(d) 1 Aug 2019; and (e),(f) 25 Aug 2019. (bottom) The time series of the predicted maxima for lead times of 1, 5, and 10 min together with the MP-PAWR observations [see legend in (b)]. The 10-min nowcasts with A3DM are only shown for 24 Jul (data are not available for other days). The horizontal dashed lines indicate the threshold of quasi no rain and heavy rainfall (10 and 37 dB$Z$, respectively). (top) The $Z_H$ observed with MP-PAWR at 1.9-km height. The black windows indicate the $5 \times 5$ km$^2$ areas where the maximum precipitation is selected.
The precipitation reaches a plateau of high intensity of 45–55 dBZ (30–150 mm h⁻¹) over a period of more than 20 min. For all three cases, the sudden increase of rainfall and the maximum intensity are well predicted by RN3D for the lead time of 5 min. For 10-min lead time, the storms (ZH > 37 dBZ) are also well predicted but their initiation (10-dBZ level crossing) is systematically delayed by about 5 min. On 24 July, the time of the crossing of the heavy rain threshold (37 dBZ) is accurately predicted, while for the other two days a delay of about 2 min is noticeable. Between 2040 and 2045 JST 24 July, moderate rainfall is predicted (20–30 dBZ). It is related to the incorrect prediction of rainfall near the considered area (Fig. 9). Similar false prediction can be noticed on 1 August, but they could be related to lower peaks of actual rainfall.

In the heavy rain phase, nowcast intensity is underestimated by less than 5 dBZ, except for 25 August, for which the underestimation is almost systematically between 5 and 12 dBZ. A3DM data are only available for 24 July. The model does not predict the storm onset with 10-min lead time. At 2055 JST, A3DM predicts no rain while a moderate precipitation is observed (25 dBZ) and at 2100 JST, only moderate rainfall of 25 dBZ is forecast while heavy precipitation of 50 dBZ is observed. After 2110 JST, that is, 10 min after the beginning of the storm, A3DM predictions become closer to the observations than RN3D.

The onset of the storm of 25 August appears to be a challenging case for the operational nowcast of JMA, which uses conventional radars. Honda et al. (2022a) show that it does not predict rainfall at 0035 JST with 10-min lead time and that with 5-min lead time it predicts rainfall intensity of only 10 mm h⁻¹ (37 dBZ) for an observation of 95 mm h⁻¹ (~52 dBZ). Note that small differences in the observed precipitation values are expected with our study due to the different setting of the two analyses. For instance, the area in which the maximum prediction is taken is not the same. Despite these differences, the results clearly show the better nowcast skill of RN3D in this case, which may be due to the high spatiotemporal resolution of MP-PAWR or the use of AI technique. Further investigation will be carried out in the future using a larger dataset and direct comparisons. The same study shows nowcasts using the MP-PAWR-based Riken’s experimental real-time assimilation system (Honda et al. 2022b). The results are very close to that of RN3D. Riken’s 10-min nowcasts are 20 mm h⁻¹ (~42 dBZ) and 80–84 mm h⁻¹ (51 dBZ) at 0035 and 0040 JST, respectively. These are very promising results for RN3D since Riken’s nowcast system is the most advanced one. The main advantage of the AI approach is that it is less computationally demanding than Riken’s system, which runs on a supercomputer.

To summarize, the consistency of the results between the different events separated by 1 to 2 years of the training data shows the robustness of RN3D to predict the onset of sudden storms. The high nowcast skills of RN3D is well illustrated by the comparison with other nowcast approaches.

6. Influence of input parameters

The influence of the input parameters on the 10-min-lead-time nowcasts of 24 and 29 July 2018 and 1 August 2019 are investigated using a perturbation method. The sensitivity of the nowcast to a given input parameter is inferred by the change of the CSI when the value of the considered parameter is set to zero: δCSI = CSI₁ − CSI, where CSI and CSI₁ are the CSIs for the nonperturbed and perturbed cases, respectively. Figure 18 shows the CSI changes for thresholds of 10 and 37 dBZ. Only results for ZDR are shown since ZDR is obviously the main source of information and no good nowcasts can be performed without it.

Overall, negative changes are calculated for most nowcasts indicating deterioration of prediction skill if ZDR is set to zero. The largest differences are observed on 1 August and the lowest change are found for 29 July. On 29 July, the largest change for the threshold of 10 dBZ are found at 0155 JST (~0.15) when most of the rainfall dissipated. Otherwise, the changes remain relatively low elsewhere (absolute values < 0.05). On 24 July, the changes are smaller than on 1 August but still significant (>0.05). These results could indicate that ZDR has a larger influence when strong convection occurs. However, a more appropriate evaluation considering a larger number of observations should be performed.

As discussed in section 2b, Doppler velocity (V_dop) was initially considered as an additional input for RN3D but was dropped due to its incompatibility with a CNN-based model. Figure 19 shows the changes in CSI for such a model when ZDR and V_dop are set to zero. Note that these results were obtained with an earlier version of the model (Baron et al. 2021b). However, the CSI changes for ZDR remain compatible with those described in Fig. 18. More important here are the changes due to the Doppler velocity, which are significantly lower. This confirms that V_dop does not help RN3D for nowcasting precipitations and also illustrates that the perturbation method is able to correctly detect if a parameter is not contributing to nowcasts.

7. Conclusions

Nowcasting of heavy localized precipitation on meso-γ scale have been presented for 10-min lead time. The nowcasts are based on the extrapolation of new MP-PAWR observations
with a supervised recurrent neural network (RN3D) that account for the vertical dimension of the measurement up to 10 km height. The main input of the model is the horizontal polarization radar reflectivity, which is used as a proxy of the rainfall. The model presents key originalities compared to other PAWR nowcasting studies: 1) LSTM units with 3D convolutions are used to account for the vertical range, 2) the model uses the differential reflectivity, 3) the training is carried out with rainfall observed throughout a whole summer to produce a robust model, and 4) an adversarial training method improves the representation of fine spatial structures and improve the prediction of heavy rainfalls \((ZH > 37 \text{ dBZ})\). The predictions have been evaluated with observations of 2018 and 2019 representing different types of localized rainfalls. We show that \(Z_{DR}\) has an impact on nowcasting convective rain, unlike the Doppler velocity. RN3D nowcast skills are very high to predict the onset of sudden localized rainfall. It outperforms a 3D advection model (A3DM) as well as the JMA operational nowcasts for the analyzed cases. However, RN3D tends to underestimate heavy rainfall especially when the storm is well developed. In this phase of the storm, A3DM nowcast scores are found slightly higher. The good results obtained with data separated by more than 2 years show the great capacity for generalization of the model though a relatively small validation set is used. In the future a validation set covering a greater summer period will be used to consolidate the model.

In future work, studies will be carried out to provide a confident level value or an uncertainty to nowcasts. Further comparisons with JMA and Riken systems should be realized and techniques proposed in other studies should be tested such as ensemble prediction (Ravuri et al. 2021) or different loss functions (Tran and Song 2019; Wang et al. 2021; Oprea et al. 2022). An optical flow technique as in A3DM should be used to include motion information instead of the Doppler velocity. This should improve the predictions under strong horizontal advection. The TrajGRU architecture (Shi et al. 2017) will also be tested for this purpose. The complementarity between RN3D and A3DM should be further exploited by combining the two approaches as in Kim et al. (2021). The close performance obtained with different neural network architectures (Baron et al. 2021b) is probably a sign that key improvements could only be obtained by adding information that cannot be measured with MP-PAWR even if their resolution is lower. For example, improvements could be made using brightness temperatures measured with the Himawari geostellite (Bessho et al. 2016), which should provide constraints on cloud coverage, temperature, and humidity. Increasing the nowcast lead time will be another challenge for future work.

A direct application of the method presented in this paper to predict 20-min-long sequences did not yield satisfactory results (Baron et al. 2021a) and a better strategy has to be found. This will require the use of information about the atmosphere before the raindrops form (Kato et al. 2022). Finally, note that new MP-PAWRs are planned to be deployed in Japan in the near future and the possibility of applying the model to other sites will be studied (Han et al. 2022b).

**Acknowledgments.** The development of MP-PAWR was supported by Council for Science, Technology and Innovation (CSTI), Cross-ministerial Strategic Innovation Promotion Program (SIP), Enhancement of National Resilience against Natural Disasters (funding agency: National Research Institute for Earth Science and Disaster Resilience). Author Baron thanks Drs. S. Otsuka and M. Mdini (Riken, Japan) for valuable discussions that helped to improve this study, as well as the editor and the reviewers for their constructive comments.

**Data availability statement.** Codes are available on request. MP-PAWR and PAWR data can be requested online (https://pawr.nict.go.jp/index_en.html).

### APPENDIX A

#### Optimizer Parameters

The usual Adam stochastic gradient descent method is used to optimize the parameters of RN3D and discriminator. The keras version in tensorflow 2.5 is used. For the prediction model, the optimizer parameters are learning rate \(= 10^{-3}\).
$\epsilon = 0.1, \beta_1 = 0.9,$ and $\beta_2 = 0.999$. The same values are used for training the discriminator except that the learning rate $= 10^{-4}$.

APPENDIX B

Encoder/Decoder Module

Each of the four layers of the encoder/decoder (EC/DC) module (Fig. 3) is composed of two sublayers, each with a convLSTM3D unit (Shi et al. 2015). In the encoding branch, the operations in layer $l$ ($l = 1, \ldots, 4$) are as follows:

\[
y_{2l-1} = \text{ConvLSTM3D}(x_{2l-2}, F_{2l-1} = 64, \text{ stride} = 1)
\]
\[
x_{2l-1} = x_{2l-2} / y_{2l-1}
\]
\[
y_{2l} = \text{ConvLSTM3D}(x_{2l-1}, F_{2l} = 64, \text{ stride} = 1)
\]
\[
x_{2l} = \text{DWSC–BN}(y_{2l}), \tag{B1}
\]

where $x_{2l-2}$ is the input of the layer $l$, $x_k$ is the output of the $k$th sublayer ($k = 1, \ldots, 8$), $F_k$ is the number of spatial filters (size of $3 \times 3 \times 3$) and DWSC and BN are the downscale function to reduce the size of each spatial dimension by 2 and the batch normalization function, respectively. The setting of other convLSTM3D parameters follows conventional usage (Shi et al. 2017) (e.g., tanh and sigmoid activation functions are used for the output and the gates, respectively). Two sublayers are equivalent to a single layer with a filter size of 5 pixels. A skip connection connects directly the layer input $x_{2l-2}$ to the first sublayer output ($y_{2l-1}$) to improve the gradient flow through the EC/DC layers (He et al. 2015). The connection is realized by the concatenation operation ($\|$).

The decoder is mirror image of the encoder, except that the sublayer indexes are reversed and the downscaling operations are replaced by upscaling ones. The operations for layer $l$ are as follow:

\[
\hat{y}_{2l} = \text{ConvLSTM3D}[^{\hat{y}_{2l+1}}, F_{2l}, h_0 = y_{2l}(T), c_0 = c_{2l}(T)]
\]
\[
\hat{x}_{2l} = \hat{x}_{2l+1} / \hat{y}_{2l}
\]
\[
\hat{y}_{2l-1} = \text{ConvLSTM3D}[\hat{x}_{2l}, F_{2l-1}, h_0 = y_{2l-1}(T), c_0 = c_{2l-1}(T)]
\]
\[
\hat{x}_{2l-1} = \text{UPSC–BN}(\hat{y}_{2l-1}). \tag{B2}
\]

where $\hat{x}_{2l+1}$ is the output of the layer $l + 1$ (underneath layer $l$ in Fig. 3), $h_0$ and $c_0$ are the initial values of the hidden states of the convLSTM unit; they are equal to the hidden states at the last time step $T$ of the mirror units in the encoder branch. The upscaling function simply duplicate each voxel to increase the size of the spatial dimensions by a factor of 2. The input of the decoder branch ($\hat{x}_0$) is zero and the size of decoder output $\hat{x}_1$ is $T_o \times X \times Y \times Z \times F$ ($T_o$ is the number of time steps).

APPENDIX C

Postdecoder Module

The aim of the postdecoder module is to refine the 5D decoder output $\hat{x}_1$ [Eq. (7)] to make the nowcast at the target altitude of 1.9 km. First, the decoder features at altitude of 1.9 km are pulled out, that is a tensor $x$ with the size of $T_o \times X \times Y \times F$, where $F = 64$ the number of features (filters), $X, Y$ are the spatial dimensions, and $T_o = 10$ is the number of time steps (1-min sampling). The number of features is decreased from 64 to 1 progressively through a series of CNN2D layers as follow:

\[
y(T + 1, ..., T + 10) = \text{TD–Conv2D}(x, F = 64, K = [3, 3])
\]
\[
y = \text{LeakyRelu}(y, \text{ slope} = 0.3)
\]
\[
y = x / y
\]
\[
y = \text{TD–Conv2D}(y, F = 64, K = [3, 3]), \tag{C1}
\]
\[
y_1(T + 1, ..., T + 10) = \text{LeakyRelu}(y, \text{ slope} = 0.3)
\]
\[
y = \text{TD–Conv2D}(y_1, F = 32, K = [3, 3])
\]
\[
y = \text{LeakyRelu}(y, \text{ slope} = 0.3)
\]
\[
y = y_1 / y
\]
\[
y = \text{TD–Conv2D}(y, F = 1, K = [3, 3]), \tag{C2}
\]
\[
y_1(T + 1, ..., T + 10) = y + Z_{H}(T), \quad \text{and}
\]
\[
Z_{H}(T + 1, ..., T + 10) = \text{Relu}(y_z),
\]

where $Z_{H}(T)$ is the radar reflectivity at 1.9 km height observed at time $T$ (last time step of the input) and TD-Conv2D is a time-distributed 2D convolution function with $F$ filters of size $K = 3 \times 3$ (time-distributed means that the same function is applied to all time steps). The function LeakyRelu($x$, slope = $a$) is a nonlinear activation function that returns $x$ if $x > 0$ and $ax$ otherwise. The shortcut connection shown in Fig. 3b corresponds to the equation with $Z_{H}(T)$ before the last equation. In the last equation, the Relu function is applied to put the negative values of the output to zero.

REFERENCES


