https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

MRDF-Net: A Model with Multidimensional Reconstruction Convolution and Dynamic Force Unit for Radar Nowcasting

Guangxin He[123],Wang Zhang[123],Xiaoran Zhuang[4*],Yuxuan Feng[4],Juanzhen Sun[5],Yubao Qiu[6],Lei Lei[3*],and Jingjia Luo[1]

1 State Key Laboratory of Climate System Prediction and Risk Management (CPRM) /
Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD) /
Key Laboratory of Meteorological Disaster, Ministry of Education (KLME) /
International Joint Research Laboratory on Climate and Environment Change (ILCEC),
Nanjing University of Information Science and Technology,
Nanjing 210044, China;

2 Nanchang Meteorological Bureau,Nanchang 330000, China;

3 Beijing Meteorological Observatory, Beijing 100097, China;

4 Jiangsu Meteorological Observatory, Nanjing 210008, China;

5 National Center for Atmospheric Research (NCAR), Boulder, CO 80305, USA;

6 Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China;

Correspondence to: Xiaoran Zhuang (zxrxz3212009@163.com) and Lei Lei (leilei_bjmet@126.com)

Abstract. Accurately predicting rapid weather changes is essential for meteorological services and disaster prevention, an area typically addressed by radar nowcasting. However, achieving accurate and stable predictions over extended forecasting horizons remains a challenging task due to the increasing uncertainty and error accumulation inherent in spatiotemporal sequence modeling. To address this challenge, this paper proposes MRDF-Net, a novel spatiotemporal sequence prediction model that integrates a multidimensional reconstruction convolution module with a dynamic force unit module to enhance forecasting accuracy and stability.The reconstruction convolution module adopts a dual reconstruction strategy across spatial and channel dimensions, which effectively reduces feature redundancy while preserving sensitivity to complex meteorological patterns. The dynamic force unit module, on the other hand, simplifies nonlinear operations in the self-attention mechanism to improve computational efficiency and feature representation.Experimental results demonstrate that MRDF-Net achieves state-of-the-art performance on standard short-term forecasting tasks, as measured by the Critical Success Index (CSI) and Heidke Skill Score (HSS). More notably, the model maintains its superior predictive capability in extended two-hour forecasting scenarios. MRDF-Net effectively alleviates the echo weakening commonly observed in other models by better preserving strong echo regions, resulting in more accurate predictions of detailed spatial structures. These results highlight the strong potential of MRDF-Net for operational meteorological forecasting.

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

Keywords: precipitation forecasting, radar short-range forecasting, multidimensional reconstructive convolution, dynamic force unit, long- and short-term memory networks

1 Introduction

Radar forecasting, a critical component of weather prediction, utilizes radar data for short-term (usually 2 to 12 hours) and nowcasting (usually 0 to 2 hours) applications (Li et al., 2023a;Chen et al., 2020). Traditional radar nowcasting methods encompass optical flow extrapolation, physical modeling, spatiotemporal correlation, and empirical statistics (Wang et al., 2024). Among these, the optical flow method predicts future echo changes by estimating motion vectors between successive radar images (Czibula et al., 2020a;Caseri et al., 2022;Yin et al., 2021;Zhang et al., 2023). Popular techniques like the Horn-Schunck and Lucas-Kanade approaches operate with the assumptions of image grayscale invariance and motion smoothness. Consequently, their performance is limited in scenarios involving complex or non-rigid motion. Physical model-based methods simulate precipitation processes in the atmosphere using kinetic and thermodynamic equations, as exemplified by numerical weather prediction (NWP) models (Liu et al., 2022). While relatively accurate for medium- and long-term forecasts, these methods are less effective for short-term applications due to their sensitivity to initial and boundary conditions and their high computational cost. Methods based on spatialtemporal correlation employ statistical models, such as ARMA, to predict changes in echo intensity and distribution by leveraging the inherent correlations in radar data across time and space. While high effective for very short time nowcasting, their performance degrades over longer time scales. In contrast, empirical statistical methods (e.g., the nearest neighbor point matching method) rely on historical data and statistical features. While proficient under specific conditions, their general applicability is limited (Huang et al., 2023;Mao and Sorteberg, 2020;Han et al., 2021;Socaci et al., 2020;Czibula et al., 2021;Ko et al., 2022;Cuomo and Chandrasekar, 2021).

The application of artificial intelligence algorithms, in particular Long Short-Term Memory networks (LSTM), has advanced radar short-term forecasting in recent years (Niu et al., 2024;Cao et al., 2022;Li et al., 2023b;Marrocu and Massidda, 2020;Pan et al., 2021). LSTM can effectively capture long-term dependencies within radar echo sequences, overcoming the limitations of traditional methods in modeling complex, nonlinear weather systems (Yao et al., 2022). By remembering and processing data from multiple time steps, LSTM not only predicts the spatiotemporal evolution of radar echoes more accurately, but also dynamically adjust the importance of input information. This capability enhances the accuracy and timeliness of short-term weather forecasts (Thiruvengadam et al., 2020), making this deep learning approach a key technology for producing more forward-looking and practical predictions (Nie et al., 2021;Spyrou et al., 2020).

ConvLSTM (Convolutional Long Short-Term Memory) is the first model to integrate Convolutional Neural Network (CNN) with LSTM for radar short-term forecasting (Shi et al., 2015). By effectively capturing spatiotemporal features in radar echoes, it enhances the accuracy of precipitation forecasts and pioneered the application deep learning to weather forecasting. Subsequent models have built upon this foundation. For instance, TrajGRU (Trajectory Gated Recurrent Unit) addressed limitation of fixed convolution kernel in ConvLSTM by employing dynamic convolution kernels that adapt to

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

motions of varying directions and scales (Shi et al., 2017), thus better capturing complex echo evolution. Furthermore, PredRNN (Predictive Recurrent Neural Network) introduces a Spatiotemporal Memory mechanism within its LSTM unit to model feature interaction more effectively (Wang et al., 2017). An advanced version, PredRNN++, incorporates causal LSTM units and Gradient Highway units to mitigate the gradient vanishing problem (Wang et al., 2018).

Other significant models in radar short-term forecasting are PhyDNet and Rainformer. PhyDNet's core innovation is its unique decoupling architecture featuring a physically constrained concurrent unit (PhyCell) (Guen and Thome, 2020). This approach models physical dynamics by isolating them from other unknown factors and uses convolution to simulate partial differential equations (PDEs). In contrast, Rainformer's architecture employs a simple and yet efficient architecture. It utilizes a Feature Extraction Balancing Module (FEBM) to separately capture local and global features, which are then fused using a Gated Fusion Unit (GFU) for optimal balance (Bai et al., 2022).

These methods represent classic algorithms in the field of radar nowcasting and feature open-source and reproducible implementations, making them suitable as baseline models for fundamental comparisons with the model proposed in this paper. Furthermore, several cutting-edge approaches have also been proposed in recent years, such as FourCastNet (Pathak et al., 2022), Earthformer (Gao et al., 2022), SmaAt-UNet (Trebing et al., 2021), Broad-Unet (Fernández and Mehrkanoon, 2021), ViViT-Prob (Qiu et al., 2025), MF-UFNO (Ye et al., 2025), and RadarDiT (Wan et al., 2025). In this paper, we selected the three models most relevant to our work for further comparative analysis: ViViT-Prob (Qiu et al., 2025), MF-UFNO (Ye et al., 2025), and RadarDiT (Wan et al., 2025).

These models represent video vision transformers, Fourier neural operators, and diffusion-transformer architectures, respectively, providing a diverse overview of modern approaches to radar nowcasting. ViViT-Prob adapts video vision Transformers by incorporating 3D convolution and sparse attention for effective spatiotemporal fusion (Qiu et al., 2025). MF-UFNO employs frequency-domain feature extraction and multivariate fusion to better capture mesoscale and small-scale radar echo structures (Ye et al., 2025). RadarDiT integrates diffusion models with Vision Transformers and introduces image constraints to enhance radar extrapolation (Wan et al., 2025).

Although existing models have introduced certain innovations, they remain limited in meeting the demands of radar nowcasting. To address this gap, this paper proposes MRDF-Net. Built on ST-LSTM, MRDF-Net incorporates a Multidimensional Reconstruction Convolution to extract more representative spatiotemporal features and a Dynamic Force Unit to enhance global contextual modeling. This complementary design captures both local dynamics and global patterns in radar echoes, resulting in enhanced long-term forecasting, superior handling of complex meteorological phenomena, and an overall improved feature representation. Experimental results demonstrate that MRDF-Net achieves higher accuracy in two-hour nowcasting compared to benchmark models.

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

2 Algorithm description

In this section, we illustrate the MRDF-Net model, describe in detail its overall structure and operation mechanism, and analyze its key components. First, the multidimensional reconstruction convolution module and the dynamic force unit module are explained, then, the internal units of MRDF-Net is elucidated, and finally, its overall structure is described.

2.1 Multidimensional Reconstruction Convolution

The multidimensional reconstruction convolution module consists of two units: the spatial reconstruction unit (SRU) and the channel reconstruction unit (CRU), which are placed sequentially. Specifically, for the intermediate input feature X in the bottleneck residual block, a spatial refinement feature X^{ω} is first obtained by a spatial reconstruction operation and then channel refinement feature Y is obtained using a channel reconstruction operation. By integrating the multidimensional reconstruction convolution module into the LSTM network, it is possible to reduce the redundancy between intermediate feature maps and improve the feature representation of the network model. The multidimensional reconstruction convolution module is shown in Figure 1.

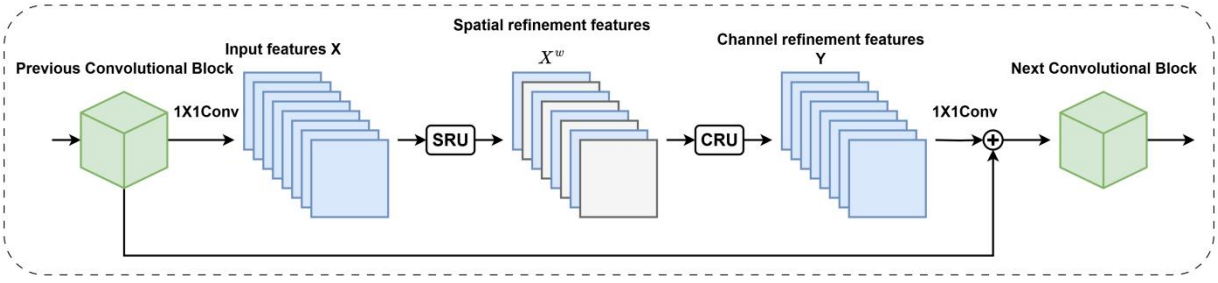

Figure 1. Multidimensional Reconstruction Convolution Module

To exploit the spatial redundancy of features, a spatial reconstruction unit (SRU) is introduced. This unit processes the features through a separation-reconstruction operation. The purpose of the separation operation is to distinguish information-rich feature maps from those with less spatial content, thereby highlighting important spatial information. Specifically, given an intermediate feature map $X \in R^{N\times C\times H\times W}$, where N denotes the batch dimension, C the channel dimension, and H and W the spatial height and width, respectively, the input feature X is first normalized by subtracting the mean value μ and then divided by the standard deviation σ. The formula is as follows:

$$X_{out} = GN(X) = \gamma \frac{X-\mu}{\sqrt{\sigma^2+\varepsilon}} + \beta$$

Here, μ and σ are the mean and standard deviation of X, ε is a small normal number added to ensure division stability, and γ and β are trainable affine transformation parameters.

The model utilizes the trainable parameters $\gamma \in R^C$ to measure of the variance of spatial pixels in each batch and channel. Richer spatial information corresponds to greater spatial pixel variance, resulting in higher γ values. The normalized weights $W_\gamma \in R^C$ are calculated by the following formula, which reflects the importance of the different feature maps:

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

$$W_\gamma = \{\omega_i\} = \frac{\gamma_i}{\sum_{j=1}^{C}\gamma_j}, i, j = 1,2,\cdots,C$$

Subsequently, the weights W_γ are mapped to the range (0,1) by an s-type function and gated by a threshold. Weights above the threshold are set to 1 to obtain the informative weights W_1, while and weights below the threshold are set to 0 to obtain the non-informative weights W_2. The entire process for obtaining W is defined as:

$$W = \mathrm{Gate}(\mathrm{Sigmoid}(W_\gamma(\mathrm{GN}(X))))$$

Multiply the input feature X with weights W_1 and W_2 respectively to obtain two weighted features: information-rich feature X_1^ω and the less information-rich feature X_2^ω. Thus, the input features are divided into two parts: X_1^ω contains informative and expressive spatial content, while X_2^ω contains little to no information and is considered a redundant component.

To reduce spatial redundancy, we have designed a new reconstruction strategy that fuses high-information features with low-information features to obtain more expressive features that occupy less space. Unlike simple stacking methods, our approach introduces a cross-reconstruction mechanism to enhance information exchange between the two feature types using weighted fusion. Finally, the features $X^{\omega 1}$ and $X^{\omega 2}$ obtained from cross-reconstruction are concatenated to generate an optimized spatial feature map X^ω.The specific formula for construction is represented as：

$$X_1^\omega = W_1 \otimes X$$

$$X_2^\omega = W_2 \otimes X$$

$$X_{11}^\omega \oplus X_{22}^\omega = X^{\omega 1}$$

$$X_{21}^\omega \oplus X_{12}^\omega = X^{\omega 2}$$

$$X^{\omega 1} \cup X^{\omega 2} = X^\omega$$

The symbol $\otimes$ represents element-wise multiplication, $\oplus$ represents element-wise addition, and $\cup$ represents element wise concatenation. The Spatially Normalized Convolution (SRU) applied to the intermediate feature X not only separates high- and low-information features but also uses this distinction to highlight key spatial components and suppress redundant ones during reconstruction.

To address channel redundancy in the features, we introduce a channel reconstruction unit (CRU) based on a segmentation-transformation-fusion strategy. Repeated standard convolution $k \times k$ is typically used to extract features, which often results in relatively redundant feature maps across the channel dimension. Let $M^k \in R^{c \times k \times k}$ represent a convolution kernel of $k \times k$, and $X, Y \in R^{c \times h \times w}$ represent the input and output features, respectively. The standard convolution can thus be defined as $Y = M^k X$. Specifically, the CRU replaces this standard convolution and is implemented through three key operations – segmentation, transformation, and fusion.

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

In short, by employing a segmentation-transformation-fusion strategy, the redundancy of spatially refined feature map X^{ω} in the channel dimension can be further reduced. The CRU extracts rich, representative features using lightweight convolutions, while managing redundant features through fast operations and feature multiplexing. The process involves three stages. First, the segmentation step separates channels into subgroups to highlight informative meteorological structures (e.g., convective cells, precipitation cores). Second, the transformation operation strengthens the discriminative capacity of these features by suppressing irrelevant or noisy channels. Finally, the fusion stage integrates the enhanced signals, retaining key meteorological patterns while discarding redundant ones. By sequentially arranging the SRU and CRU, an efficient multidimensional reconstruction convolution module is constructed.

This redundancy reduction and feature reconstruction strategy effectively enhances the model's ability to capture multi-scale, multidimensional dynamic weather features without increasing the computational costs. By embedding the multidimensional reconstruction convolution module into the radar short-range forecast model, the accuracy and efficiency of the forecasts can be improved, particularly for longer-term forecasts and high-resolution data processing, while hardly increasing the model complexity and computational overhead. Overall, the integration of this module significantly boosts the model's adaptability and generalization ability under complex weather conditions.

2.2 Dynamic Force Unit

The model structure of the dynamic force unit is shown in Figure 2. For the input $x \in R^{N\times C}$, the traditional self-attention (SA) mechanism is expressed as follows:

$$A(x) = \sigma(QK^T/\sqrt{d_k})V$$

$$Q = xW_Q,\ K = xW_k,\ V = xW_v$$

where A denotes the attention operator. If the nonlinearity of the softmax function is eliminated, $\sigma\ (QK^T)\ V$ can be decomposed into $O\ (N\times h + h \times N)$. The dynamic force unit uses XNorm instead of softmax, which allows the self-attention module to first compute K^TV. Since removing the nonlinear operation can lead to performance degradation, a simple constraint is introduced to mitigate this issue.

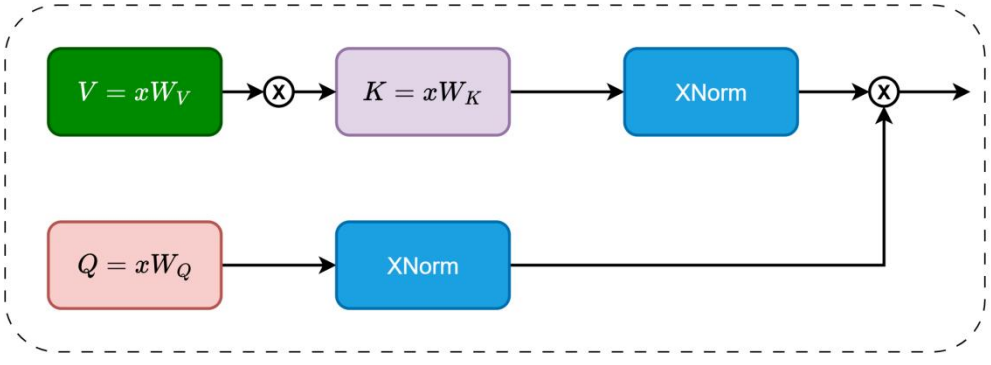

Figure 2. Dynamic Force Unit

The cross-normalization operator, XNorm, is defined as follows:

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

$$A(x) = XN_{dim=filter}(Q)(XN_{dim=space}(K^T V))$$

$$XN(a): = \frac{\gamma a}{\sqrt{\sum_{i=0}^{h} ||a||^2}}$$

Where γ is a learnable parameter and h is the number of embedding layer dimensions. This is a common L2 normalization paradigm, but it is applied across two dimensions: the spatial dimension of $K^T V$ and the channel dimension of Q. Therefore, the method is termed "cross-normalization." Using the associative law, the keys and values are first multiplied, and then the result is multiplied by the queries. As shown in Figure 2, the computation complexity of both multiplication operations is O (hNd), so the process is linear with respect to N.

To analyze the XNorm module in detail, we consider its role within a residual connections. The output of any module can be represented as:

$$x_{n+1} = x_n + f(x_n)$$

Where n and x represent the current layer index and input image, respectively. If we interpret x as the displacement of the object and n as the time parameter, the above equation can be rewritten as:

$$x_{t+1} = x_t + f(x_t)$$

$$f(x) = \frac{\Delta x}{\Delta t}$$

Since most neural networks are in discrete form, Δt can be regarded as a constant (for ease of derivation, $\Delta t=1$ is taken here). The residual part can be understood as velocity, so under the condition that the particle has a unit mass and $\Delta t=1$, this term is equivalent to the action of force.The residual part can be interpreted as the velocity that drives the evolution of the feature state. Assuming unit mass and $\Delta t=1$, this velocity term reflects the effect of the underlying driving force learned by the model.

By embedding the Dynamic Force Unit into the ST-LSTM framework, the proposed architecture achieves a significant improvement in processing large-scale, high-resolution radar data for short-range forecasting tasks, while maintaining effective feature extraction and spatial–temporal information capture. Its linear complexity mechanism also ensures fast inference on high-resolution data.

2.3 MRDF-Net internal unit

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

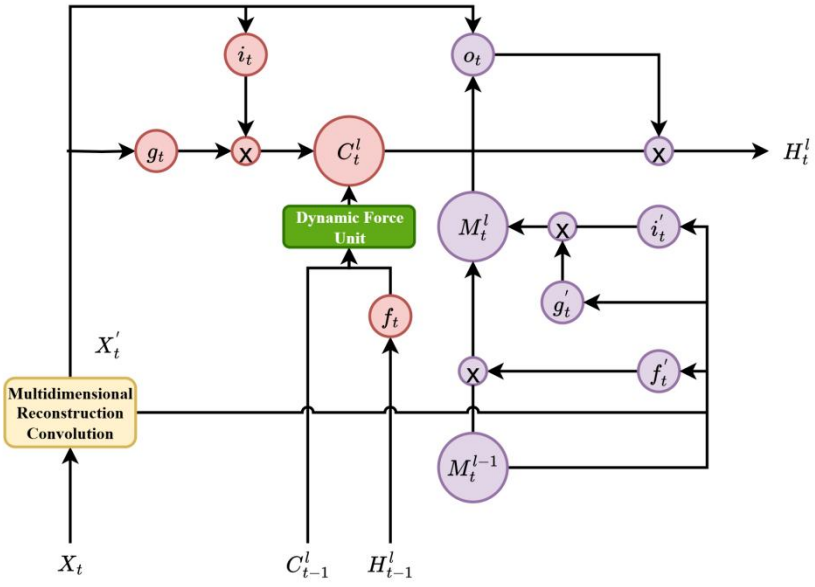

Figure 3. MRDF-Net internal unit

The internal cell structure of MRDF-Net is shown in Figure 3, where X_t is the input, H_{t-1}^l is the hidden state, C_{t-1}^l is the temporal memory unit, and M_t^{l-1} is the spatial memory unit. The gates i_t, g_t, and f_t represent the input gates, g_t input modulation gate, forget gate, respectively, corresponding to the three-gate structure of an LSTM. MRDF-Net is formulated as follows:

$$g_t = \tanh(W_{xg} * X_t' + W_{hg} * H_{t-1}^l + b_g)$$

$$i_t = \sigma(W_{xi} * X_t' + W_{hi} * H_{t-1}^l + b_i)$$

$$f_t = \sigma(W_{xf} * X_t' + W_{hf} * H_{t-1}^l + b_f)$$

$$C_t^l = CA(f_t \odot C_{t-1}^l) + i_t \odot g_t$$

$$g_t' = \tanh(W_{xg}' * X_t' + W_{mg} * M_t^{l-1} + b_g')$$

$$i_t' = \sigma(W_{xi}' * X_t' + W_{mi} * M_t^{l-1} + b_i')$$

$$f_t' = \sigma(W_{xf}' * X_t' + W_{mf} * M_t^{l-1} + b_f')$$

$$M_t^l = f_t' \odot M_t^{l-1} + i_t' \odot g_t'$$

$$o_t = \sigma(W_{xo} * X_t' + W_{ho} * H_{t-1}^l + W_{co} * C_t^l + W_{mo} * M_t^l + b_o)$$

$$H_t^l = o_t \odot \tanh(W_{1x1} * [C_t^l, M_t^l])$$

The multidimensional reconstruction convolution module is positioned between the input features X_t and the input gate of the LSTM unit, where it preprocesses the input. It reduces the spatial and channel redundancy of the input features by combining the spatial reconstruction unit and channel reconstruction unit. The resulting processed X_t' are more compact and representative, thereby improving the efficiency and accuracy of the subsequent LSTM unit processing. The dynamic force

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

unit is located between the LSTM's candidate memory unit C_t^l and its current memory unit M_t^l. Its main function is to optimize the information flow of the LSTM. It achieves this by simplifying the self-attention mechanism —for instance, by replacing softmax with L2 normalization — and introducing cross normalization (XNorm) to maintain computational efficiency. This module enhances the model's ability to capture long-term dependencies while reducing consumption of computational resource.

2.4 MRDF-Net overall structure

Figure 4 illustrates the architecture of the multi-layer MRDF-Net. Each layer of MRDF-Net units receives input from the preceding layer and passes its output to the subsequent one. The main information flow proceeds sequentially from left to right, with data being progressively processed at each stage. The connections between layers maintain the delivery logic of a traditional LSTM network, while a feature map M is introduced to enhance the information dissemination. In each layer, the hidden state H_t^l and temporal memory unit C_t^l are passed to the next layer, and the MRDF-Net units process this state information accordingly.

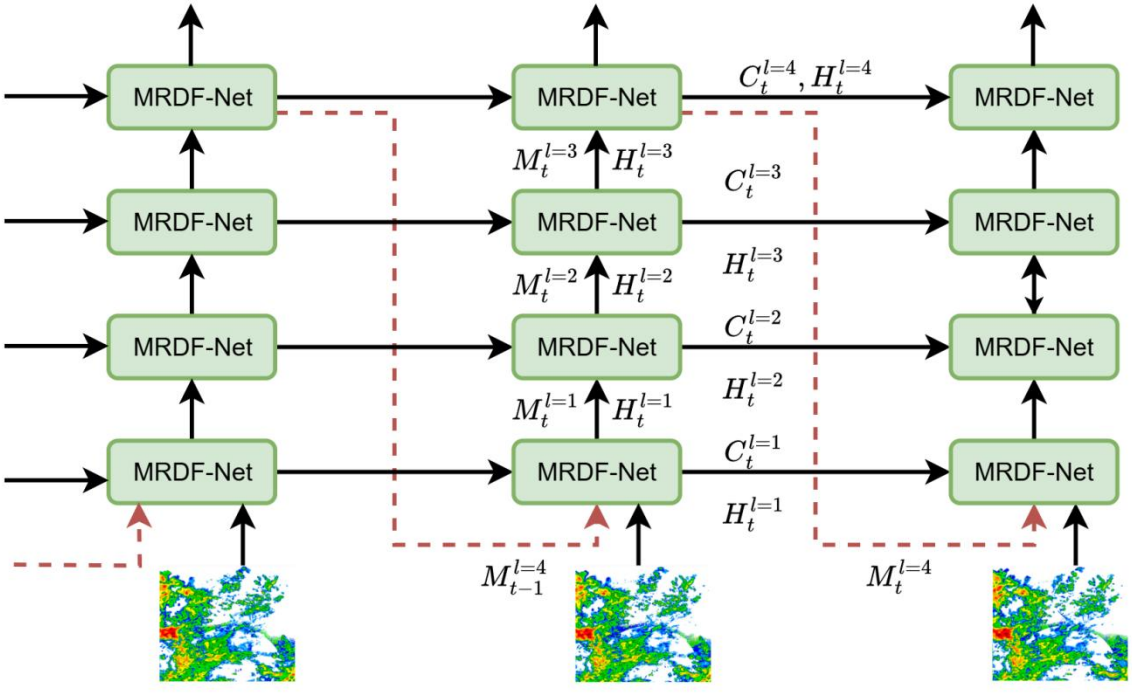

Figure 4. Overall structure of MRDF-Net

The feature map M, as an important component of the MRDF-Net, continuously updated at each time step. M_t^{l-1}, representing the feature map from the previous layer, is processed by the MRDF-Net unit to generate the feature map of the current layer M_t^l. The overall model features a multi-layer, stacked design. At each time step t, the model consists of multiple layers, each containing an MRDF-Net cell. Each MRDF-Net unit receives the hidden state H_t^l and temporal memory unit C_t^l

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

from the lower layer and passes its processed result to the MRDF-Net unit in the upper layer. At the bottom layer, the model receives the original radar echo input image. This input is then processed by the multidimensional reconstruction convolution module, which performs spatial and channel reconstruction to remove redundant features, resulting in a streamlined yet informative feature map M_t^l.

These processed feature maps are passed through each layer and used as input for each MRDF-Net unit. The states and memories output by each unit are then optimized by the dynamic force unit, which enhances information flow via a simplified self-attention mechanism. The red dashed arrows in the figure represent cross-layer connections. These skip connections allow higher-level MRDF-Net units to directly access the state information from lower layers, which enhances the model's feature extraction capability and ensures robust longer-term dependency modeling.

Through its multi-layer stacking, skip-connected architecture — combined with the use of multidimensional reconstruction convolution module and dynamic force unit—the MRDF-Net model significantly improves the accuracy and efficiency of short-range radar forecasting when processing complex time series data.

3. Data sets

3.1 Moving-MNIST dataset

The Moving-MNIST dataset(Modified National Institute of Standards and Technology database) is a standard benchmark dataset for video prediction and time-series modeling, widely used to evaluate deep learning models on complex forecasting tasks. It is generated by simulating the random movement of handwritten digits from the raw MNIST dataset across a sequence of consecutive frames, creating a dynamic scene that provides an ideal testing environment for various models.

The primary task for the Moving-MNIST dataset is to predict the content and motion trajectory of the next M frames (e.g., 10 frames) based on the first N frames of images (e.g., the first 10 frames). This task requires a model to possess strong temporal modeling capability, motion inference capability, and the ability to learn complex nonlinear relationships. Due to complexities such as random motion, occlusion, and boundary collisions, this dataset effectively simulate the challenges of real-world dynamic object prediction, making it ideal for evaluating the performance of predictive models.

For our experiments, we use the Moving-MNIST dataset as a benchmark to evaluate the prediction performance of the MRDF-Net model. Specifically, the dataset consists of 10,000 training samples and 10,000 test samples, where each sequence features digits moving along unique paths and speeds. This serves as a foundation for subsequent testing on more complex datasets and in practical applications.

3.2 Radar echo dataset

The radar echo dataset is crucial for meteorological forecasting, particularly in short-term precipitation nowcasting (e.g., at the hourly level). It comprises a sequence of two-dimensional images generated by meteorological radar scans and post-observation analysis. This images reflect the spatial and temporal distribution of atmosphere water vapor, clouds, and

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

precipitation, as well as their evolutionary trends. Analyzing this data enables the prediction of future intensity and path of precipitation.

The resolution of radar images ranges from 1 km to several kilometers, with a coverage area typically spanning from tens to hundreds of kilometers. Each sequence in the dataset captures the evolution of precipitation over time, reflecting its spatial structure and temporal dynamics.

However, these data sequences often contain significant noise and uncertainty, necessitating high-quality prediction models. Complex precipitation patterns they capture, such as convective systems, frontal precipitation, and tropical cyclones, make them ideal for testing and validating spatio-temporal prediction models.

We used a time-series of quality-controlled mosaic data from multiple S-band weather radars in Jiangsu province, China. Reflectivity factors (in dBZ) is directly associated with the density of atmospheric water droplets and ice particles; higher values indicate greater concentration of hydrometeors or more intense precipitation. The specific parameters of the dataset are provided in Table 1.

Table 1. Specifications of the Radar Echo Dataset

Parameter	Value / Description
Data Type	Radar echo time series
Physical Meaning	Base reflectivity at 3 km altitude
Unit	dBZ
Value Range	0 – 70
Total Images	26,378
Spatial Coverage	Entire Jiangsu Province, China
Horizontal Resolution	0.01° (approximately 1 km)
Temporal Resolution	6 minutes
Image Size	480 × 560 pixels
Radar Type	Multiple S-band weather radars
Processing Workflow	Quality control + radar mosaicking

This study uses radar data from Jiangsu Meteorological Observatory from 2019 to 2021 (April to September each year). The data is divided into a training set and a testing set. The training set contains 21,103 radar images for training the MRDF-Net model, while the testing set contains 5,275 radar images for evaluating the model's performance. Each sequence consists of 20 consecutive radar images, where the first 10 serves as model inputs and the last 10 serves as the actual outputs, representing the corresponding real observation images. The time interval between radar images is 6 minutes, therefore each 20-image sequence covers a 2-hour period. After training, the model was used to predict the next 10 images (i.e., the following hour) using the first 10 images (i.e., the previous hour) as input. This enables the model to analyze one hour of

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

radar data to forecast precipitation for the subsequent hour. Likewise, the model was also evaluated for 2-hour nowcasting; that is, it uses the first 10 images to predict the next 20 images.

4. Experimental Procedure and Results

We conducted experiments on both the Moving MNIST dataset and radar echo dataset, using the former as a benchmark test. The MRDF-Net model was implemented based on the PyTorch framework, with training and testing performed on an NVIDIA A100 GPU. For the radar echo experiments, the batch size was set to 4, the learning rate was set to 0.0001, and the Adam optimizer was used.

4.1 Moving-MNIST experiments

To verify the effectiveness of the MRDF-Net model in temporal prediction, we first conducted experiments on the Moving-MNIST dataset. The Structural Similarity Index (SSIM) and Mean Square Error (MSE) were used as evaluation metrics to measure performance in terms of spatial structure preservation and pixel-level accuracy. Lower MSE and higher SSIM values indicate better model performance.

Figure 5 shows the 1-hour prediction results using the MNIST-Net dataset performed by MRDF-Net and other mainstream prediction models including Rainformer, PhyDNet, PredRNN++, PredRNN, TrajGRU and ConvLSTM, along with the input and ground-truth.

As shown in this figure, MRDF-Net outperforms the other models on the Moving MNIST dataset. Its predictions feature clearer digital contours and better shape preservation while maintaining high temporal consistency, demonstrating strong spatiotemporal model capability. In contrast, Rainformer's results are slightly blurry and suffer from a significant loss of detail. PhyDNet performs well in shape preservation, but its boundaries are blurred, and some details are lost. PredRNN++and PredRNN generated numbers with the correct overall shape; however, the edges are blurry and some numbers are deformed, particularly in later frames. While TrajGRU performs well initially, blurriness accumulates in subsequent frames, degrading the numerical structure. Finally, ConvLSTM performs the worst, with severe distortion of digital shapes and temporal misalignment , highlighting its limitation in temporal modeling.

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

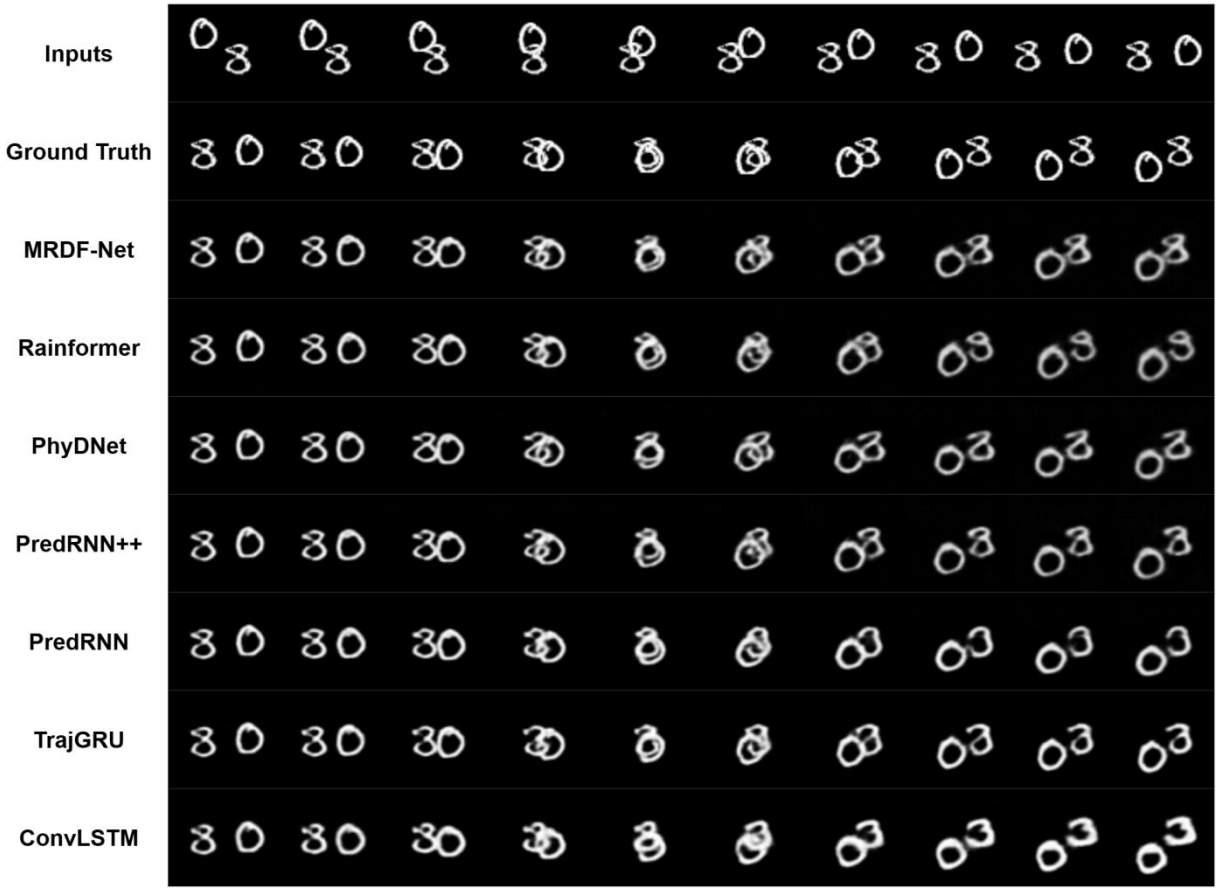

Figure 5. The input (first row), ground truth (second row), and predicted sequences from different models using the Moving-MNIST dataset. Rows from the third to the ninth represent the results of MRDF-Net, Rainformer, PhyDNet, PredRNN++, PredRNN, TrajGRU, and ConvLSTM, respectively.

As shown in Table 2, MRDF-Net achieved the best results in both SSIM and MSE metrics. This indicates that it not only better preserves the structural integrity of the images but also effectively reduces pixel-wise errors compared to the ground-truth. As shown in Figure 5 the other models produce predictions that are either more blurred or distorted, which is especially apparent for fast-moving or overlapping cells.

Table 2. Verification Results from the Moving-MNIST Experiments

Model	MSE/frame	SSIM/frame
ConvLSTM	98.7	0.745
TrajGRU	55.1	0.860
PredRNN	53.3	0.892
PredRNN++	45.8	0.901

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

PhyDNet	40.5	0.914
Rainformer	36.3	0.929
MRDF-Net	**26.4**	**0.943**

4.2 Radar echo experiments

4.2.1 Evaluation indicators

We use the two common metrics, Critical Success Index (CSI) and Heidke Skill Score (HSS), to evaluate the performance of MRDF-Net on the radar echo dataset. These metrics can effectively measure model's prediction accuracy and stability under different precipitation intensities.

The Critical Success Index (CSI), also known as the Threat Score (TS), is a classic evaluation metric used to measure a model's ability to detect target events. The CSI ranges from 0 to 1, where 1 indicates a perfect prediction and 0 indicates a complete failure. The calculation formula is:

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

Here, TP (True Positive) denotes the number of correctly predicted events (hits), FP (False Positive) denotes the number of incorrectly predicted events (false alarms), and FN (False Negative) denotes the number of actual events that are predicted negative (misses). The CSI integrates the effects of hits, misses, and misses, with a score of 1 indicating that the model detects events perfectly. It is particularly suitable for use in radar echo prediction to assess the overall detection effectiveness of the model for target weather events (e.g., rainfall, thunderstorms, etc.).

The Heidke Skill Score (HSS) measures the improvement of a classification model's prediction over random guessing, and is widely used in weather forecasting. The HSS value ranges from negative infinity to 1. A score of 1 indicates a perfect prediction, 0 indicates that the model performs no better than random guessing, and a negative value indicates that the model performs worse than random guessing. The formula for HSS is:

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

Here, TP, FP, and FN are the same as in the CSI, representing hits, false alarms, and misses, respectively, while TN (True Negative) denotes correct rejections. A higher HSS indicates a greater predictive accuracy. A key characteristic of the HSS is that it accounts for the effect of random prediction, providing a more accurate evaluation of a model's performance. Therefore, it complements the CSI, which primarily focuses on the model's absolute ability to detect target events.

4.2.2 Analysis of 1-hour forecast experiments

To evaluate the performance of MRDF-Net, it was compared with other state-of-the-art models using the average CSI and HSS over the testing period. The results, shown in Tables 3 and 4, indicate that MRDF-Net provides a significant improvement in forecast skill.

Table 3. CSI at different thresholds for 1-hour forecasts from different models

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

Model	10dBZ	20dBZ	40dBZ	Average
ConvLSTM	0.7283	0.4622	0.1627	0.4511
TrajGRU	0.7353	0.4761	0.1857	0.4657
PredRNN	0.7428	0.4783	0.1882	0.4698
PredRNN++	0.7497	0.4897	0.1938	0.4777
PhyDNet	0.7537	0.4923	0.2013	0.4824
Rainformer	0.7555	0.5168	0.2042	0.4922
MRDF-Net	**0.7725**	**0.5284**	**0.2282**	**0.5097**

The CSI results show that MRDF-Net achieved the best performance across all thresholds. It performs particularly well for heavy precipitation (40 dBZ), with a score of 0.2282, which is significantly higher than that of other models. MRDF-Net also achieved the highest average score (0.5097), demonstrating its superior forecasting capability, especially for intense precipitation. In contrast, Rainformer and PhyDNet show fair performance across all thresholds. PredRNN++ and PredRNN, however, show a noticeable decline in CSI at 40 dBZ, with scores of only 0.1938 and 0.1882, respectively, indicating their limitations in predicting heavy precipitation. TrajGRU and ConvLSTM have the lowest performance; ConvLSTM, in particular, has the lowest average CSI (0.4511), reflecting its weaker capability in precipitation forecasting. Overall, MRDF-Net demonstrates stable and superior predictive capability across different precipitation intensities, confirming its effectiveness in enhancing short-term precipitation forecasting accuracy.

Table 4. HSS at different thresholds for 1-hour forecasts from different models

Model	10dBZ	20dBZ	40dBZ	Average
ConvLSTM	0.7017	0.5334	0.1984	0.4778
TrajGRU	0.7135	0.5396	0.2058	0.4863
PredRNN	0.7184	0.5425	0.2149	0.4919
PredRNN++	0.7269	0.5486	0.2368	0.5041
PhyDNet	0.7283	0.5492	0.2517	0.5097
Rainformer	0.7297	0.5558	0.2731	0.5195
MRDF-Net	**0.7563**	**0.5726**	**0.2993**	**0.5427**

The HSS results in Table 4 also shows that MRDF-Net achieved the best results across all thresholds, Similar to the CSI trend, it attained a significantly higher HSS (0.2993) at the 40 dBZ threshold than other models, demonstrating its predictive capability for heavy rainfall. Additionally, MRDF-Net achieved the highest average HSS (0.5427) among all models. Rainformer and PhyDNet also performed well, with average HSSs of 0.5195 and 0.5097, respectively, indicating stable performance across different rainfall intensities. In contrast, PredRNN++ and PredRNN show lower HSSs, especially at the

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

40 dBZ threshold, with values of only 0.2368 and 0.2149, indicating their limitations in predicting heavy rainfall. TrajGRU and ConvLSTM show relatively weaker overall performance, with noticeable declines in HSS at higher thresholds, resulting in average scores of 0.4863 and 0.4778.

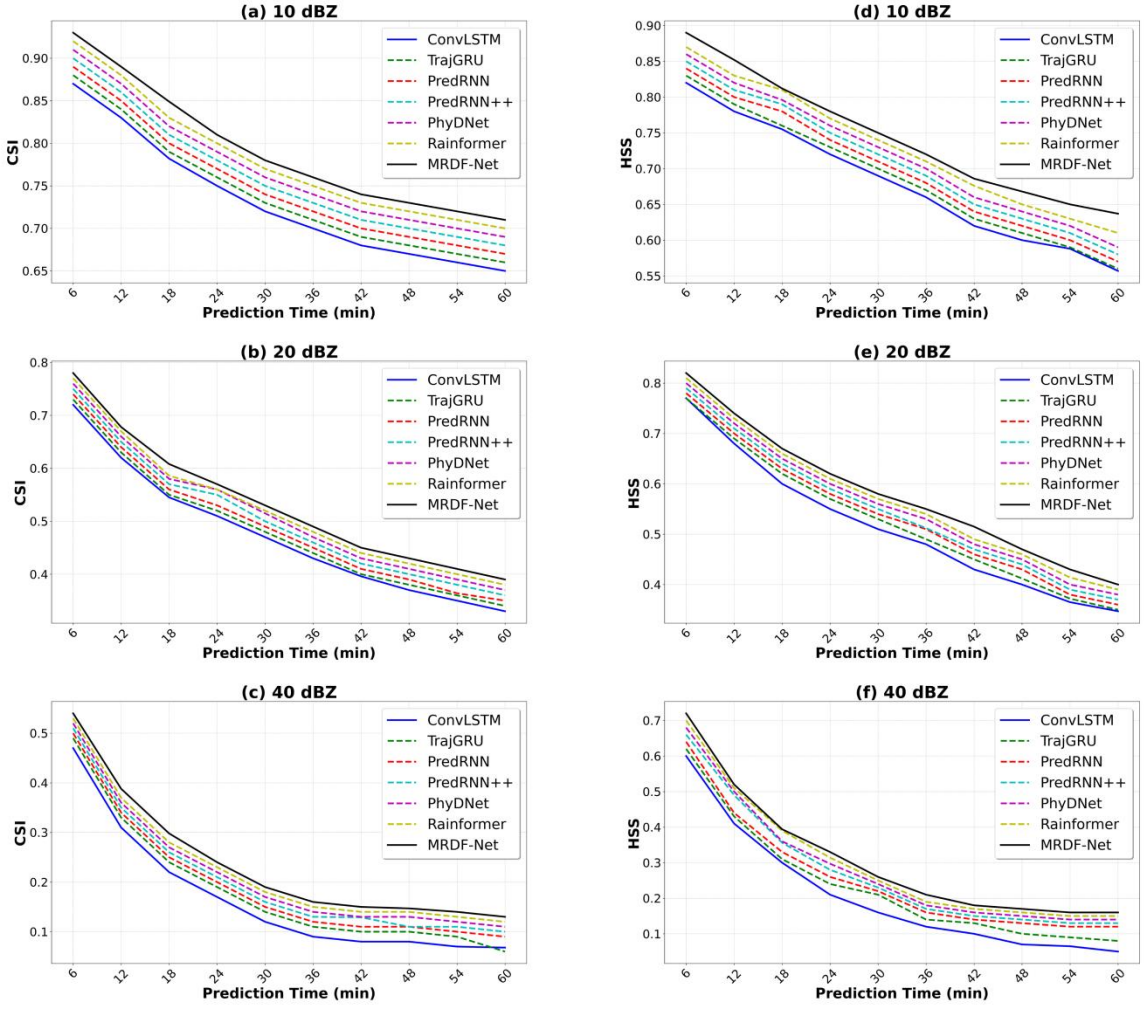

Figure 6. The CSI (left column) and HSS (right column) versus forecast lead time for different thresholds: top row — 10 dBZ, middle row — 20 dBZ, and bottom row — 40 dBZ.

Figure 6 compares the performance of different prediction models for reflectivity thresholds of 10 dBZ, 20 dBZ, and 40 dBZ. Overall, the CSI and HSS of all models decrease with forecast lead time; however, MRDF-Net exhibits the smallest decline, particularly at 40 dBZ. This demonstrates its superior stability in heavy precipitation forecasting up to one hour. This advantage is attributed to the Multidimensional Reconstruction Convolution Module, which suppresses redundant features, and the efficient modeling of longer-term dependencies in high-resolution meteorological data by the linear self-attention mechanism.

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

Across different reflectivity thresholds, the performance gap between models widens as the threshold increases. This threshold dependency highlights the inability of traditional models, such as ConvLSTM and TrajGRU, to capture the localized and sudden nature of heavy precipitation. In contrast, MRDF-Net employs a Spatial Reconstruction Unit with a dynamic gating mechanism, which enables it to more accurately locate the core regions of strong echoes. Additionally, at the moderate threshold of 20 dBZ, the CSI curve of MRDF-Net remains consistently higher than those of other models, verifying the effectiveness of the Channel Reconstruction Unit in selecting key channel features through lightweight convolution.

The figure also shows that traditional sequential models (e.g., PredRNN++) exhibit only a small performance gap compared to MRDF-Net in the initial forecast stage (<20 minutes), but their performance deteriorates more rapidly thereafter. This finding confirms that the multi-layer skip connection proposed in this study effectively mitigates the vanishing gradient problem.

Notably, in the HSS subplot for 10 dBZ, the curve for MRDF-Net remains stable before gradually declining, reaching 0.64 at 60 minutes. This is 14.3% higher than ConvLSTM's value of 0.56, indicating superior robustness in the continuous forecasting of larger-scale weak precipitation. This advantage stems from the capability of the Dynamic Force Unit module's ability to integrate global features, whereas traditional convolutional models, with their limited receptive fields, struggle to maintain spatiotemporal consistency in weak precipitation areas.

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

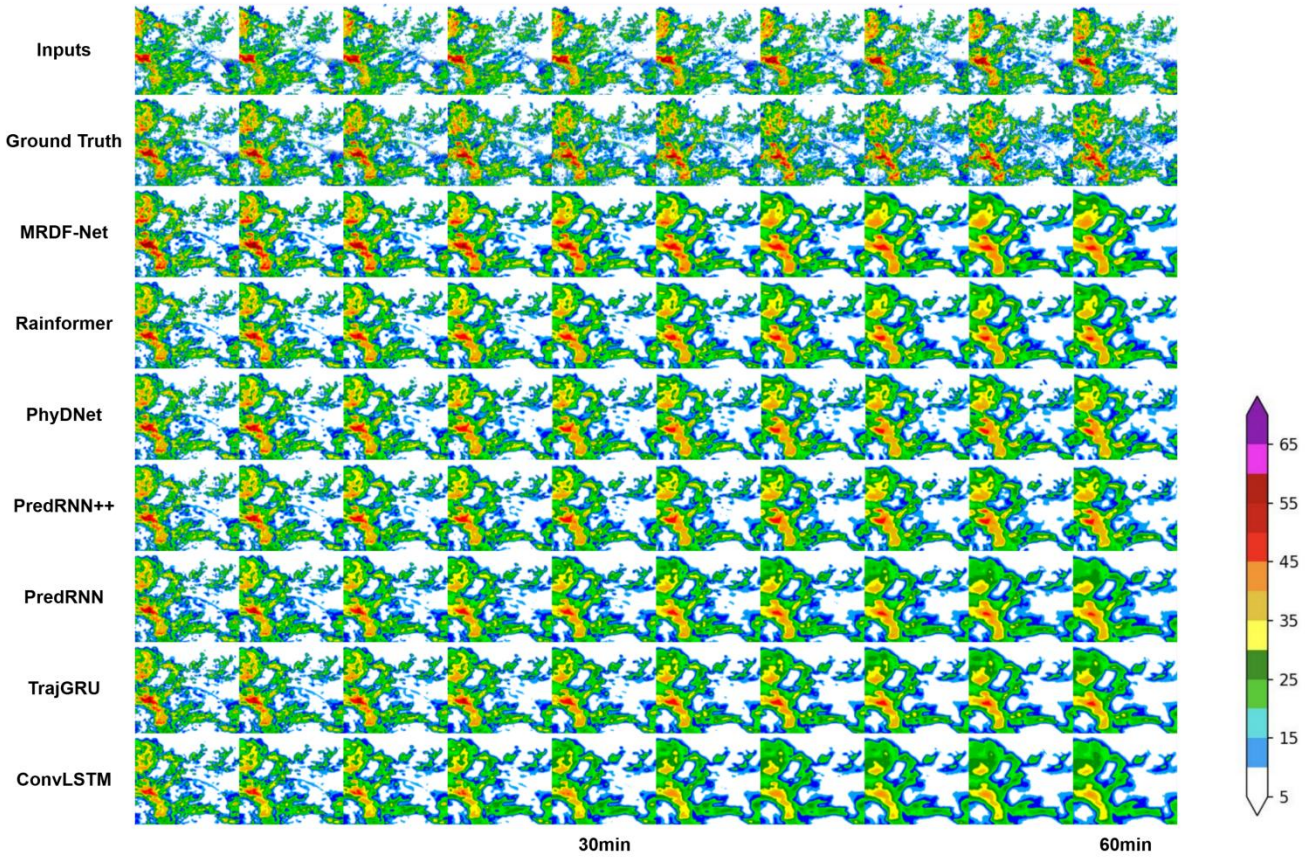

Figure 7. A radar extrapolation example from each model, shown at 6-minute intervals up to a 1-hour forecast lead time.

As shown in the extrapolated reflectivity results in Figure 7, MRDF-Net most closely matches the ground truth at all forecast lead times. At 30 minutes, the position and shape of its red core area in the main reflectivity band almost coincide with the ground truth, and the gradient transition of the surrounding weaker blue echoes is natural. At 60 minutes, although all models show varying degrees of blurriness, MRDF-Net maintains the main structure and movement direction of strong echoes, showing only slight smoothing effects, even for the small high-echo area near the northern boundary. In contrast, in Rainformer's 60-minute prediction, the predicted red area has begun to deviate from its true position, while PhyDNet shows significant attenuation in echo intensity, reducing the red area by approximately 50%.

The performance of the traditional sequence models, such as ConvLSTM and TrajGRU, is significantly weaker than that of the models with new architectures. For example, at 30 minute, the red core area in the main band in ConvLSTM has been already significantly weakened, losing the hook structure shown in the ground truth. By the-60 minute mark, the strong echo areas are further substantially weakened, indicating that the model fails to capture the physical laws of echo motion. The PredRNN series models (PredRNN, PredRNN++) maintain the echo shape quite well in 30-min predictions, but their magnitude deteriorates rapidly over time thereafter.

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

4.2.3 Analysis of 2-hour forecast experiments

To further validate the effectiveness of the proposed MRDF-Net model, we extended the forecast duration to two hours. We increased the reflectivity thresholds to 20, 35, and 45 dBZ to focus more on heavier precipitation events. This series of experiments was designed to reassessed MRDF-Net's longer-term prediction robustness in handling complex meteorological scenarios.

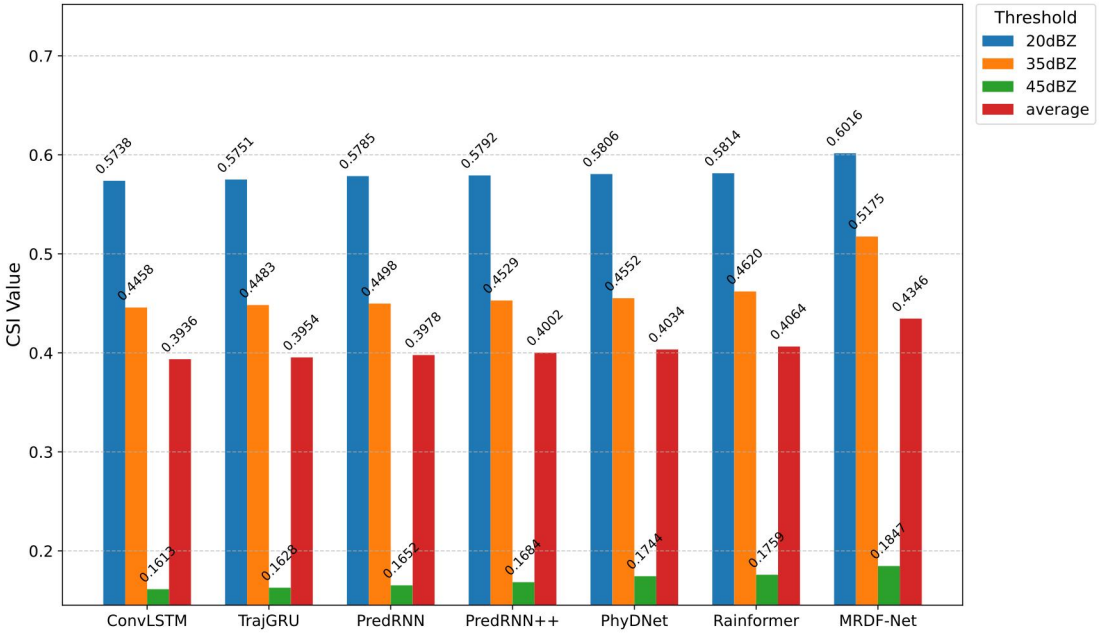

Figure 8. CSI at different thresholds for 2-hour forecasts.

The overall performance of 2-hour forecast, as shown in Figure 8, demonstrates that the MRDF-Net model achieves the highest CSI scores at the three thresholds of 20 dBZ, 35 dBZ, and 45 dBZ. With an average CSI of 0.4346, it outperforms the second-best model, Rainformer (0.4064), by 6.9%, highlighting the superiority of its architecture design. The improvement is particularly significant for moderate intensity precipitation (35 dBZ), with the CSI value of MRDF-Net (0.5175) being 12.0% higher than Rainformer (0.4620). This is likely attributable to the multidimensional reconstruction convolution mechanism. Its spatial reconstruction branch captures echo structures of different intensities through multi-scale convolution, while the temporal evolution branch maintains motion trajectory continuity using an improved attention mechanism. The dynamic interaction between these two branches effectively suppresses error accumulation in longer-term prediction.

The sensitivity of each model to threshold changes varies significantly. Although all models performed poorly at the 45 dBZ threshold, MRDF-Net still led with a slight advantage (CSI: 0.1847), thanks to its dynamic weight adjustment module, which provides targeted optimization for the strong echo core. It is worth noting that PhyDNet, which introduces physical constraints, outperforms some purely data-driven models, for example, its score (0.1744) is higher than PredRNN++'s (0.1684). This suggests the potential value of using simplified physical equations in extreme weather modeling.

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

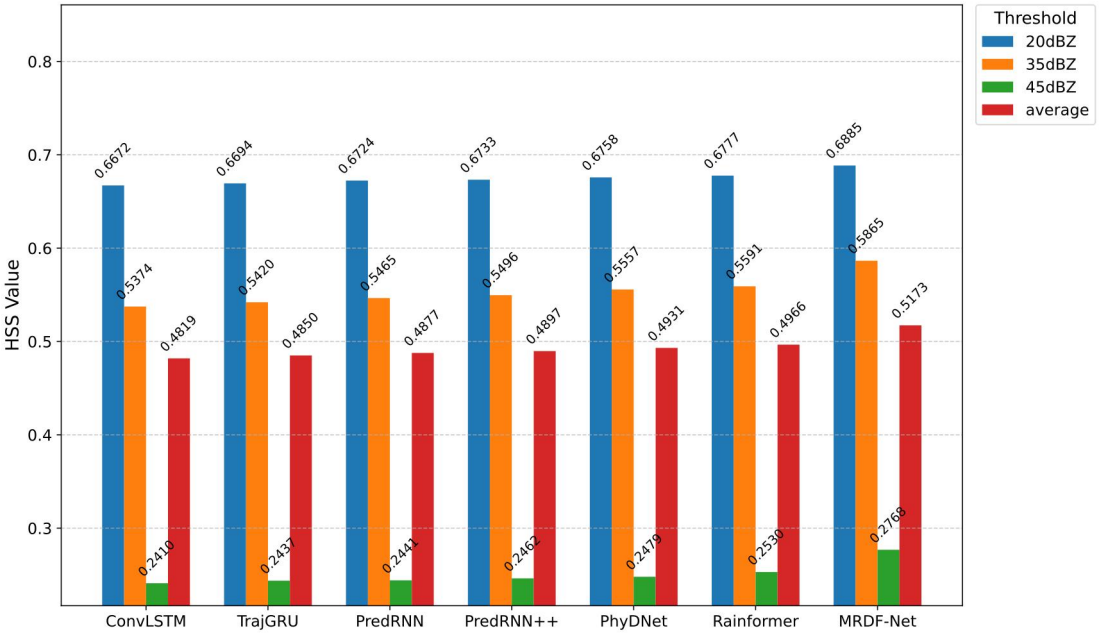

Figure 9. HSS at different thresholds for 2-hour forecasts.

According to Figure 9, MRDF-Net achieved the highest average HSS (0.5173), which is 4.2% higher than the second-best performer, Rainformer (0.4966). The improvement is more pronounced for moderate-intensity precipitation (35 dBZ), where MRDF-Net's value of 0.5865is 4.9% higher than Rainformer's (0.5591). In terms of average HSS, PredRNN++ (0.4897) shows only a 0.4% over its predecessor, PredRNN (0.4877). In contrast, the innovative cross-model architecture enabled a 6.07% increase over PredRNN. This significant improvement in HSS demonstrates that the multidimensional reconstruction convolution and dynamic force unit effectively enhance prediction stability.

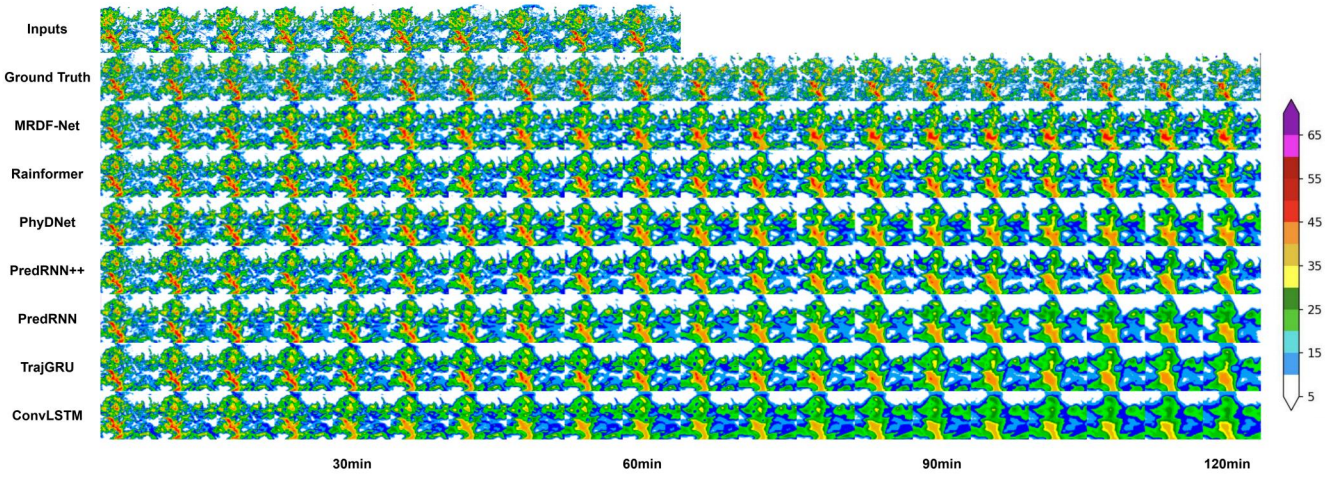

Figure 10. A 2-hour radar extrapolation example from each model, shown at 6-minute intervals up to a 2-hour lead time.

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

Figure 10 shows an example of the radar echo extrapolation results up to two hours using different methods. MRDF-Net exhibits better predictive ability compared to other methods. Throughout the entire prediction period, the radar echo field generated by MRDF-Net well successfully maintains the spatial structures and details, closely matching the ground truth. In contrast, other methods exhibit varying degrees of blurriness or structural distortion during longer-term predictions, Especially at the 90 and 120 min lead times, MRDF-Net still retains clear precipitation echo structures, while the echoes generated by ConvLSTM and TrajGRU have become blurred, and the echo intensity has significantly decreased. In addition, while PredRNN and PredRNN++ maintained the echo shape to some extent, but their predictions of heavy precipitation areas are more blurred compared those of MRDF-Net.

MRDF-Net demonstrates significant advantages for the prediction of heavy precipitation. In high echo areas above 55 dBZ, the predictions from MRDF-Net are closer to the ground truth, while Rainformer and PhyDNet tend to show premature dissipation or shape distortion in areas of heavy precipitation for longer lead times. This is likely because MRDF-Net effectively models spatiotemporal features, enabling it to maintain the morphology and intensity of heavy precipitation areas. In addition, compared to PredRNN and PredRNN++, MRDF-Net performs better at retaining the integrity of precipitation structure, thereby avoiding instability and unreasonable changes in echo shape.

MRDF-Net ' s stability in longer-term predictions is best illustrated by its coherent and reasonable precipitation distributions at the 120 minute lead time. In contrast, the predictions by Rainformer, PhyDNet, and ConvLSTM all exhibit significant dissipation. This indicates that MRDF-Net can more effectively model spatiotemporal dependencies and maintain temporal consistency of prediction, leading to greater stability in longer-term extrapolation.

In summary, MRDF-Net demonstrated superior detail preservation, improved accuracy in predicting heavy precipitation, and enhanced longer-term prediction stability, demonstrating stronger potential for application in short-term precipitation forecasting.

4.2.4 Comparison with Recent State-of-the-Art Models

We further compared MRDF-Net with three recently proposed radar echo extrapolation methods: ViViT-Prob, MF-UFNO, and RadarDiT for 1-hour forecast. These models represent the cutting-edge progress in incorporating video vision transformers, Fourier neural operators, and diffusion-transformer architectures into radar nowcasting. We conduct a quantitative evaluation under different reflectivity thresholds using CSI and HSS to compare their performances. The results are presented in Table 5.

Table 5. CSI and HSS at different thresholds for 1-hour forecasts from different models

Metrics	CSI				HSS			
Threshold	10dBZ	20dBZ	40dBZ	Avg	10dBZ	20dBZ	40dBZ	Avg
ViViT-Prob	0.7334	0.4973	0.1853	0.4725	0.7216	0.5422	0.2575	0.5071
MF-UFNO	0.7587	0.5036	0.1894	0.4839	0.7270	0.5481	0.2653	0.5135
RadarDiT	0.7857	0.5324	0.2050	0.5077	0.7835	0.5739	0.2548	0.5374

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

Metrics	CSI				HSS			
Threshold	10dBZ	20dBZ	40dBZ	Avg	10dBZ	20dBZ	40dBZ	Avg
MRDF-Net	0.7725	0.5284	0.2282	0.5097	0.7563	0.5726	0.2993	0.5427

The comparison in Table 5 reveals distinct performance characteristics among the models. At lower thresholds (10 dBZ and 20 dBZ), RadarDiT achieves slightly higher CSI values than MRDF-Net, indicating a superior ability to capture widespread weak echoes. This is consistent with the global dependency modeling capability of the diffusion-transformer architecture, which excels at better preserving large-scale echo patterns. However, at the higher threshold of 40 dBZ, MRDF-Net surpasses RadarDiT and all other models in both CSI and HSS, demonstrating its enhanced effectiveness in predicting intense convective cores and maintaining echo integrity under severe precipitation conditions. When averaged across thresholds, MRDF-Net achieves the highest scores (CSI = 0.5097, HSS = 0.5427), underscoring its overall superior performance.

5. Conclusions and discussion

In this paper, a new spatiotemporal sequence prediction method, MRDF-Net, is proposed and its performance in radar short-range forecasting is comprehensively evaluated. By introducing a Multidimensional Reconstruction Convolution module and a Dynamic Force Unit module, MRDF-Net is capable of effectively fusing multi-scale spatiotemporal features, which significantly improves the prediction performance of the model.

Experimental results on the Moving-MNIST dataset and the radar echo dataset show that MRDF-Net outperforms the existing mainstream methods in accurately capturing complex spatiotemporal dependencies, detail reproduction, and multi-scale information interaction. This can be attributed to the integration of the Multidimensional Reconstruction Convolution module and the Dynamic Force Unit into the ST-LSTM network. The multidimensional reconstruction convolution enables the model to extract more representative spatiotemporal features, while the dynamic force unit enhances its global contextual modeling capability. This complementary structure allows MRDF-Net to capture both local dynamics and global patterns in radar echoes.

Compared to conventional models, such as ConvLSTM and PredRNN, MRDF-Net introduces structural innovations that lead to superior forecasting accuracy. Specifically, the Multidimensional Reconstruction Convolution captures complex spatial correlations across multiple dimensions, whereas the dynamic force unit compensates for the inherent limitations of purely convolutional architectures by modeling long-range dependencies. The experimental evaluations on the Moving MNIST and real-world radar datasets show that MRDF-Net consistently outperforms baseline models across multiple evaluation metrics.

In addition to traditional baselines, MRDF-Net was evaluated against several recently proposed state-of-the-art models, including ViViT-Prob, MF-UFNO, and RadarDiT. These models represent cutting-edge approaches based on video vision

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

transformers, Fourier neural operators, and diffusion-transformer architectures, respectively. The results indicate that MRDF-Net achieves superior performance at higher thresholds, enabling it to better capture and maintain strong convective cores. Furthermore, it attains the highest average score across all thresholds.

The MRDF-Net model has strong promise for practical application in short-term precipitation nowcasting. Its superior prediction accuracy for radar echoes is particularly beneficial for issuing early warnings for severe convective weather. By enhancing the spatiotemporal feature representation and contextual modeling, the model enables more accurate and stable precipitation forecasts, which are crucial for disaster prevention and mitigation.

Despite its improved performance, the MRDF-Net model has some limitations. First, the model is currently trained and tested exclusively on radar reflectivity data, without incorporating other meteorological variables such as temperature, wind, or humidity. Integrating these additional variables could potentially further enhance its predictive performance. Second, the model does not explicitly quantify uncertainty, a crucial aspect in operational weather forecasting. Addressing these issues will be a key focus of future research. Moreover, we will investigate the fusion of MRDF-Net with multi-source data (e.g., satellite and ground observations) to construct a more comprehensive and fine-grained forecasting model. Additionally, research is needed to extend the model's predictive horizon beyond two hours and to evaluate its applicability across diverse regions and climatic conditions.

Code and data availability

The source code related to this study is maintained in a GitHub repository (https://github.com/zhangwang0214-arch/Radar-Nowcasting-MRDF-Net).The datasets used in this study were provided by meteorological agencies and are subject to data sharing restrictions. Therefore, the data are not publicly available, but can be accessed from the corresponding author upon reasonable request.

Author contributions

Conceptualization, G.H. and W.Z.; methodology, L.L. and J.L.; data curation and management, L.L., J.L., X.Z., and Y.F.; software development, G.H.; validation, W.Z.; formal analysis, W.Z.; investigation, J.S. and Y.Q.; writing—original draft preparation, G.H. and W.Z.; writing—review and editing, G.H., W.Z., J.S., and X.Z.; supervision, L.L. and J.L.; funding acquisition, L.L. All authors have read and approved the final manuscript.

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

Competing interests

The authors declare that they have no conflict of interest.

Financial support

Supported by the State Key Laboratory of Climate System Prediction and Risk Management (CPRM) initiative project (Grant No. CPRM-2025-NUIST-012); by China Meteorological Administration Capability Enhancement Joint Research Program（24NLTSQ015）; by the China Meteorological Administration Youth Innovation Team (CMA2023QN10); by the Sichuan Science and Technology Program (No. 2025YFNH0006); by China Meteorological Administration Innovation and Development Program（CXFZ2023J008）; by Jiangsu Provincial Key Research and Development Program under Grant BE2023766; by Scientific Research Project of Jiangsu Meteorological Service (KM202520); by the Open Fund Project for Heavy Rain (BYKJ2024Q23); and by the Bohai Rim Regional Meteorological Science and Technology Collaborative Innovation Fund project (QYXM202409).

Acknowledgments

We acknowledge the High Performance Computing Center of Nanjing University of Information Science & Technology for their support of this work.We also thank Jiangsu Meteorological Observatory for providing the radar echo dataset.

References

[1]Li, Y., Liu, Y., Sun, R., Guo, F., Xu, X., & Xu, H. (2023). Convective storm VIL and lightning nowcasting using satellite and weather radar measurements based on multi-task learning models. Advances in Atmospheric Sciences, 40(5), 887-899.

[2]Chen, L., Cao, Y., Ma, L., & Zhang, J. (2020). A deep learning-based methodology for precipitation nowcasting with radar. Earth and Space Science, 7(2), e2019EA000812.

[3]Wang, X., Zhao, H., Zhang, G., Guan, Q., & Zhu, Y. (2024). Spatiotemporal predictive learning for radar-based precipitation nowcasting. Atmosphere, 15(8), 914.

[4]Czibula, G., Mihai, A., & Mihuleţ, E. (2020). Nowdeepn: An ensemble of deep learning models for weather nowcasting based on radar products' values prediction. Applied Sciences, 11(1), 125.

[5]Caseri, A. N., Santos, L. B. L., & Stephany, S. (2022). A convolutional recurrent neural network for strong convective rainfall nowcasting using weather radar data in Southeastern Brazil. Artificial Intelligence in Geosciences, 3, 8-13.

[6]Yin, J., Gao, Z., & Han, W. (2021). Application of a radar echo extrapolation-based deep learning method in strong convection nowcasting. Earth and Space Science, 8(8), e2020EA001621.

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

[7]Zhang, Y., Long, M., Chen, K., Xing, L., Jin, R., Jordan, M. I., & Wang, J. (2023). Skilful nowcasting of extreme precipitation with NowcastNet. Nature, 619(7970), 526-532.

[8]Liu, W., Wang, Y., Zhong, D., Xie, S., & Xu, J. (2022). ConvLSTM network-based rainfall nowcasting method with combined reflectance and radar-retrieved wind field as inputs. Atmosphere, 13(3), 411.

[9]Huang, H., Tao, R., Zhao, K., Wen, L., & Chu, Z. (2023). Potential of snowfall nowcasting using polarimetric radar data and its link to ice microphysics: Study of two snowstorms in East China. Journal of Geophysical Research: Atmospheres, 128(15), e2022JD037654.

[10]Mao, Y., & Sorteberg, A. (2020). Improving radar-based precipitation nowcasts with machine learning using an approach based on random forest. Weather and Forecasting, 35(6), 2461-2478.

[11]Han, L., Liang, H., Chen, H., Zhang, W., & Ge, Y. (2021). Convective precipitation nowcasting using U-Net model. IEEE Transactions on Geoscience and Remote Sensing, 60, 1-8.

[12]Socaci, I. A., Czibula, G., Ionescu, V. S., & Mihai, A. (2020, May). XNow: A deep learning technique for nowcasting based on radar products ' values prediction. In 2020 IEEE 14th International Symposium on Applied Computational Intelligence and Informatics (SACI) (pp. 000117-000122). IEEE.

[13]Czibula, G., Mihai, A., Albu, A. I., Czibula, I. G., Burcea, S., & Mezghani, A. (2021). Autonowp: an approach using deep autoencoders for precipitation nowcasting based on weather radar reflectivity prediction. Mathematics, 9(14), 1653.

[14]Ko, J., Lee, K., Hwang, H., & Shin, K. (2022, November). Deep-learning-based precipitation nowcasting with ground weather station data and radar data. In 2022 IEEE International Conference on Data Mining Workshops (ICDMW) (pp. 1063-1070). IEEE.

[15]Cuomo, J., & Chandrasekar, V. (2021). Use of deep learning for weather radar nowcasting. Journal of Atmospheric and Oceanic Technology, 38(9), 1641-1656.

[16]Niu, D., Li, Y., Wang, H., Zang, Z., Jiang, M., Chen, X., & Huang, Q. (2024). FsrGAN: A Satellite and Radar-Based Fusion Prediction Network for Precipitation Nowcasting. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

[17]Cao, Y., Chen, L., Zhang, D., Ma, L., & Shan, H. (2022, May). Hybrid weighting loss for precipitation nowcasting from radar images. In ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 3738-3742). IEEE.

[18]LI, J., LI, L., FENG, P., & TANG, R. (2023). Radar rainfall nowcasting and flood forecasting based on deep learning. Advances in Water Science, 34(5), 673-684.

[19]Marrocu, M., & Massidda, L. (2020). Performance comparison between deep learning and optical flow-based techniques for nowcast precipitation from radar images. Forecasting, 2(2), 194-210.

[20]Pan, X., Lu, Y., Zhao, K., Huang, H., Wang, M., & Chen, H. (2021). Improving nowcasting of convective development by incorporating polarimetric radar variables into a deep - learning model. Geophysical Research Letters, 48(21), e2021GL095302.

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

[21]Yao, S., Chen, H., Thompson, E. J., & Cifelli, R. (2022). An improved deep learning model for high-impact weather nowcasting. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15, 7400-7413.

[22]Thiruvengadam, P., Indu, J., & Ghosh, S. (2020). Significance of 4DVAR radar data assimilation in weather research and forecast model-based nowcasting system. Journal of Geophysical Research: Atmospheres, 125(11), e2019JD031369.

[23]Nie, T., Deng, K., Shao, C., Zhao, C., Ren, K., & Song, J. (2021, October). Self-attention UNet model for radar based precipitation nowcasting. In 2021 IEEE sixth international conference on data science in cyberspace (DSC) (pp. 493-499). IEEE.

[24]Spyrou, C., Varlas, G., Pappa, A., Mentzafou, A., Katsafados, P., Papadopoulos, A., ... & Kalogiros, J. (2020). Implementation of a nowcasting hydrometeorological system for studying flash flood events: The Case of Mandra, Greece. Remote Sensing, 12(17), 2784.

[25]Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., & Woo, W. C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. Advances in neural information processing systems, 28.

[26]Shi, X., Gao, Z., Lausen, L., Wang, H., Yeung, D. Y., Wong, W. K., & Woo, W. C. (2017). Deep learning for precipitation nowcasting: A benchmark and a new model. Advances in neural information processing systems, 30.

[27]Wang, Y., Long, M., Wang, J., Gao, Z., & Yu, P. S. (2017). Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms. Advances in neural information processing systems, 30.

[28]Wang, Y., Gao, Z., Long, M., Wang, J., & Philip, S. Y. (2018, July). Predrnn++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning. In International conference on machine learning (pp. 5123-5132). PMLR.

[29]Guen, V. L., & Thome, N. (2020). Disentangling physical dynamics from unknown factors for unsupervised video prediction. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 11474-11484).

[30]Bai, C., Sun, F., Zhang, J., Song, Y., & Chen, S. (2022). Rainformer: Features extraction balanced network for radar-based precipitation nowcasting. IEEE Geoscience and Remote Sensing Letters, 19, 1-5.

[31]Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mardani, M., ... & Anandkumar, A. (2022). Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators. arXiv preprint arXiv:2202.11214.

[32]Gao, Z., Shi, X., Wang, H., Zhu, Y., Wang, Y. B., Li, M., & Yeung, D. Y. (2022). Earthformer: Exploring space-time transformers for earth system forecasting. Advances in Neural Information Processing Systems, 35, 25390-25403.

[33]Trebing, K., Stańczyk, T., & Mehrkanoon, S. (2021). SmaAt-UNet: Precipitation nowcasting using a small attention-UNet architecture. Pattern Recognition Letters, 145, 178-186.

[34]Fernández, J. G., & Mehrkanoon, S. (2021). Broad-UNet: Multi-scale feature learning for nowcasting tasks. Neural Networks, 144, 419-427.

[35]Qiu, Yunan, et al. "ViViT-Prob: A Radar Echo Extrapolation Model Based on Video Vision Transformer and Spatiotemporal Sparse Attention." Remote Sensing 17.12 (2025): 1966.

https://doi.org/10.5194/egusphere-2025-6488
Preprint. Discussion started: 23 February 2026

[36]Ye, Tingting, et al. "MF-UFNO: A Fourier Neural Operators-Based Model for Radar Echo Extrapolation." Earth and Space Science 12.7 (2025): e2024EA003740.

[37]Wan, Jiaquan, et al. "RadarDiT: An advanced radar echo extrapolation model for three gorges reservoir area via diffusion transformer." Journal of Hydrology: Regional Studies 61 (2025): 102703.

[38]Reinoso-Rondinel, R., Troemel, S., Simmer, C., & Rempel, M. (2021). Radar-based ensemble nowcasting: predictability analysis. Earth and Space Science Open Archive ESSOAr.

[39]Klocek, S., Dong, H., Dixon, M., Kanengoni, P., Kazmi, N., Luferenko, P., ... & Xiang, S. (2021). MS-nowcasting: Operational precipitation nowcasting with convolutional LSTMs at Microsoft weather. arXiv preprint arXiv:2111.09954.

[40]Shrestha, Y., Zhang, Y., Doviak, R., & Chan, P. W. (2021). Lightning flash rate nowcasting based on polarimetric radar data and machine learning. International Journal of Remote Sensing, 42(17), 6762-6780.

[41]Czibula, G., Mihai, A., & Czibula, I. G. (2020). RadRAR: A relational association rule mining approach for nowcasting based on predicting radar products' values. Procedia Computer Science, 176, 300-309.