Deep Learning Model based on Multi-scale Feature Fusion for Precipitation Nowcasting

Jinkai Tan[1,2,*], Qiqiao Huang[1,*], and Sheng Chen[1,2,✉]

[1]Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China
[2]Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai 519080, China
[*]These authors contributed equally to this work.

Correspondence: Sheng Chen (chensheng@nieer.ac.cn)

Abstract. ~~Accurate forecast of heavy precipitation remains~~ Forecasting heavy precipitation accurately is a challenging task ~~in~~ for most deep learning (DL)-based models. ~~This study proposes~~ To address this, we present a novel DL architecture ~~named~~ ~~'~~called "Multi-scale Feature Fusion~~'~~ " (MFF) ~~for precipitation nowcasting for~~ that can forecast precipitation with a lead time of up to 3 ~~h. The basic idea is to apply~~ hours. The MFF model uses convolution kernels with ~~various sizes to achieve~~ varying sizes to create multi-scale receptive fields~~and then~~ . This helps to capture the movement features of ~~the precipitation system (e.g.~~ precipitation systems, such as their shape, movement direction, and ~~moving speed). Meanwhile~~speed. Additionally, the architecture ~~implants~~ utilizes the mechanism of discrete probability to reduce uncertainties and forecast errors, ~~so that heavy precipitations can be produced~~ enabling it to predict heavy precipitation even at longer lead ~~time. The model uses four year's~~ times. For model training, we use four years of radar echo data from 2018 to 2021~~for model training~~ , and one year's data ~~of~~ from 2022 for model testing. ~~The model is compared~~ We compare the MFF model with three existing extrapolative models: time series residual convolution (TSRC), optical flow (OF), and UNet. ~~Results~~ The results show that MFF ~~obtains relatively~~ achieves superior forecast skills with ~~a~~ high probability of detection (POD), low false alarm rate (FAR), small mean absolute error (MAE), and high structural similarity index (SSIM). ~~The most commendable result is that~~ Notably, MFF can predict high-intensity precipitation fields at 3 ~~h lead time~~hours lead time, while the other three models can not. ~~Additionally, it can be found from the results of radially averaged power spectral (RAPS) that~~ Furthermore, MFF shows improvement in the smoothing effect of the forecast field~~. Future works will pay more attention to~~ , as observed from the results of radially averaged power spectral (RAPS). Our future works will focus on incorporating multi-source meteorological variables, ~~the structural adjustments of~~ making structural adjustments to the network, and ~~the combinations~~ combining them with numerical models to further improve the forecast skills of heavy precipitations at longer lead times.

1 Introduction

Heavy precipitation ~~is a key driver of a variety of~~ can cause various natural disasters, ~~including~~ such as floods, landslides, and mud-rock flows, which ~~pose a threat to both life and property. The term 'nowcasting' refers to predicting precipitation over a certain~~ can be life-threatening and cause property damage. Nowcasting is the term used for predicting precipitation in a specific region within a short time frame (~~typically~~ usually less than 3 hours) and with a ~~fine-grained~~ high spatiotemporal resolution (Ayzel et al., 2020; Czibula et al., 2021). It ~~is an attractive~~ has become a popular research topic in ~~the field of~~ hydrometeorology. The ~~destruction of a precipitation event mainly depends on its~~ intensity, duration, and ~~falling area . Therefore~~area of precipitation determine the extent of its destruction. Consequently, accurate and timely nowcasting ~~has become an indispensable link~~ is essential for disaster early warning and emergency response (Chen et al., 2020; Ehsani et al., 2021). However, real-time, large-scale, and fine-grained precipitation nowcasting remains a challenging task due to the ~~inherent~~ complexities of atmospheric conditions (Ehsani et ~~a~~al., 2021; Kim et al., 2021).

~~Conventional~~ There are two main conventional approaches for precipitation nowcasting~~mainly include~~ : numerical weather prediction (NWP)-based methods (Sun et al., 2014; Yano et al., 2018) and radar echo-based quantitative forecasts (Liguori et al., 2014). The NWP models ~~describe atmospheric phenomenons~~ predict precipitation dynamics by solving a series of differential equations ~~and thus predict precipitation dynamics~~ to describe atmospheric phenomena (Dupuy et al., 2021)~~. They represent the main tools for precipitation forecasts. However, these models~~ , but they are computationally intensive ~~,~~ and time-consuming, and ~~difficult to assimilate the local data,~~ their forecast products depend on initial/boundary conditions (Marrocu et al., 2020; Ehsani et al., 2021). ~~Besides~~Moreover, the first few hours of precipitation predictions by NWP models are invalid, so they are not commonly used in nowcasting (Han et al., 2019; Yan et al., 2020). ~~The~~ On the other hand, radar echo-based quantitative models use the ~~so-called~~ Z-R relationship ~~(radar reflectivity 'Z' and precipitation intensity 'R')~~ to drive precipitation rates and ~~further~~ estimate precipitation accumulations. ~~Especially, the optical flow (OF)~~ The optical flow model is the simplest technique in radar echo-based quantitative forecast models~~. It~~ , which consists of tracking and extrapolation~~, where~~ . In this technique, an advection field is estimated from a series of consecutive radar echo images~~. This field~~ , and it is then used to extrapolate recent radar echo images through semi-Lagrangian schemes or interpolation procedures (Ayzel et al., 2019). ~~Progress~~ Many studies have documented the progress and achievements in precipitation nowcasting with variations of the OF model ~~have also been documented in many other studies~~ (Marrocu et al., 2020; Pulkkinen et al., 2019; Ayzel et al., 2019; Prudden et al., 2020; Liu et al., 2015; Woo et al., 2017; Li et al., 2018). ~~Although~~ However, the OF model ~~and its variations achieved great advances in precipitation nowcasting, they have~~ has certain limitations due to the assumption of a constant advection field (Prudden et al., 2020; Li et al., 2021).

~~Recently~~In recent years, deep learning ~~(DL) techniques in precipitation nowcastinghave drawn much attention from numerous studies~~techniques have become increasingly popular for precipitation nowcasting, due to their superior ~~performances~~ performance in tracking and processing successive frames of radar echo video/image. For ~~example~~instance, Shi et al. (2015) treated precipitation nowcasting as a spatiotemporal sequence predictive problem ~~, and proposed~~ and proposed a convolutional long short-term memory (ConvLSTM) architecture~~which helps to capture both~~ , which captures spatial and temporal features of

radar echo sequences. This model outperformed the OF method. ~~Considering the change of radar echo over time, in~~ In their follow-up study (Shi et al., 2017), they introduced a trajectory GRU (TrajGRU) model, which used the same convolutional and recursive networks as ~~in the ConvLSTM , while~~ the ConvLSTM while excavating the spatially-variant relationship of radar echo ~~is excavated by~~ through its sub-networks. Moreover, Chen et al. (2020) built a new architecture with a transition path (star-shaped bridge, SB) based on ConvLSTM, which gleans more latent features and makes the model more robust~~, the model was used~~. The model was tested in precipitation nowcasting over the Shanghai area and achieved better ~~performances~~ performance than some conventional extrapolation methods. To improve the limitation of time-step reduction in the ConvLSTM model, Yasuno et al. (2021) proposed a rain-code approach with multi-frame fusion, ~~thus the model has~~ allowing the model to have a forecast lead time of 6 hours. Ronneberger et al. (2015) presented a ~~deep network with~~ U-shaped architecture deep network, namely U-Net, consisting of a contracting path to capture context and an expanding path that enables precise positioning. This model was ~~first~~ initially used in biomedical segmentation applications. Numerous attempts ~~to develop an~~ have been made to develop a UNet-based precipitation nowcasting model~~and obtained certain success, such as~~ , including the 'RainNet' in Germany (Ayzel et al., 2020), the 'MSDM' in ~~eastern~~ Eastern China (Li et al., 2021), the 'Convolutional Nowcasting-Net' with IMERG products (Ehsani et al., 2021), the 'SmaAt-UNet' in the Netherlands (Trebing et al., 2021), the 'FURENet' for convective precipitation nowcasting (Pan et al., 2021), and the nowcasting system with ground-based radars and geostationary satellites imagery (Lebedev et al., 2019)~~, and~~ . Additionally, Sadeghi et al. ~~,~~(2020) used a UNet convolutional neural network and geographical information ~~for improving~~ to enhance near real-time precipitation estimation.

~~Apart from ConvLSTM-based and UNet-based models, many~~ When it comes to radar-based nowcasting, there are several plug-and-play modules ~~for radar-based nowcasting~~ available that use different network architectures. Some models use ConvLSTM or UNet-based architectures, while others either trim deformable network architectures or implant various feature extraction operations into the network architectures. For ~~example~~instance, Ravuri et al. (2021) ~~presented~~ proposed a conditional generative model for ~~the probabilistic nowcastingwhich~~ probabilistic nowcasting. Their model produced realistic and spatiotemporally consistent predictions with a lead time of up to 90 minutes~~. This model eliminate the blurry nowcasting maps and outperformed~~ , outperforming UNet and PySTEPS (Pulkkinen et al., 2019). The Google Research group (Sønderby et al., 2020) developed "MetNet", a weather probabilistic model ~~'MetNet' which used~~ that uses axial self-attention mechanisms to unearth weather patterns from large-scale radar and satellite data~~, it~~ . The model provided probabilistic precipitation maps for up to ~~8~~ eight hours over the continental United States at a spatial resolution of 1 ~~km2 and~~ km and a temporal resolution of ~~2~~ two minutes. The Huawei Cloud group (Bi et al., ~~2022~~2023) devised a 3D ~~earth specific~~ earth-specific transformer module and developed ~~'~~"Pangu-Weather~~'~~", a high-resolution system for ~~the global weather forecast. The~~ global weather forecasting. This system showed good application prospects for its superior performance in many downstream forecast tasks such as wind, temperature, and typhoon forecasts. Researchers from DeepMind and Google (Lam et al., 2022) proposed a novel machine learning weather simulator named ~~'GraphCast', which~~ "GraphCast". It was an autoregressive model based on graph neural networks and a high-resolution multi-scale mesh representation~~. The model~~ , which produced medium-range global weather forecasting for up to 10 days. The Microsoft Research group (Tung et al., 2023) developed and demonstrated the ~~'ClimaX' modelwhich~~ "ClimaX" model. This model extended the Transformer architecture with novel encoding and aggregation blocks,

the model resulted resulting in superior performance on benchmarks for both weather forecasting and climate projections. Similarly, the author of Chen et al. (2023) presented an advanced data-driven global medium-range weather forecast system named 'FengWu', which is "FengWu". This system was equipped with model-specific encoder-decoders, cross-modal fusion Transformer, and a replay buffer mechanism, and it . It solved the medium-range forecast problem problems from a multi-modal and multi-task perspective. Marrocu et al. (2020) proposed the 'PreNet' modelwhich is "PreNet" model, which was based on a widely-used semisupervised semi-supervised and unsupervised learning DL method named '"generative adversarial network'" (GAN, Goodfellow et al., 2014). The model's performance was compared with state-of-the-art OF procedures and shown showed remarkable superiority. Zheng et al. (2022) established the 'GAN–argcPredNet' model"GAN-argcPredNet" model, which was also based on GAN architecture, and it . It can reduce the prediction loss in a small-scale space and show more detailed features among prediction maps.

However, some limitations/challenges in the above DL-based models for precipitation nowcasting are widely reported . First, because precipitation dynamics are quite complex and DL models are difficult have their limitations and challenges, as reported by various studies. Firstly, these models struggle to extrapolate short-term local convection or precipitation fields by learning the prior knowledge from due to the complex nature of precipitation dynamics, and the fact that DL models rely solely on historical radar echo data alone to learn prior knowledge (Su et al., 2020; Chen et al., 2020; Ehsani et al., 2021), it is . This makes it challenging to predict fast-moving precipitation systems or short-term local convections with characterized by rapid growth and dissipation. Second, Secondly, iterative forecasts tend to result in accumulative errors and uncertainties usually occur during iterative forecasts due to the discrepancy due to discrepancies between the model's training and testing process processes (Ayzel et al., 2020; Prudden et al., 2020; Li et al., 2020; Singh et al., 2021; Huang et al., 2023), resulting in . This can lead to low values of heavy precipitation, smoothing or blurry forecast fields. ThirdThirdly, the convolution operation used in DL models covers precipitation fields as comprehensively as possiblebut is unable to , but it cannot reveal the rapid changes in echo intensity, deformation, and movement of precipitation fields (Ehsani et al., 2021; Kim et al., 2021). ThereforeConsequently, DL models inevitably produce some undesirable forecast outputs, such as declining forecast skills with increasing lead time, smoothing and blurry precipitation fields, missing extreme precipitation events, and poor forecast skills for precipitation growth and dissipation.

Large-scale precipitation systems are affected by many factorssuch as influenced by several factors, including prevailing westerlies, trade-wind zone, mesoscale weather systems, land-sea distributions, and topography effects (Huang et al., 2023; Luo et al., 2023). Therefore, accurate and As a result, real-time precipitation nowcasting is still a very challenging issueand accurate precipitation forecasting remains a challenging task. In this study, we apply utilized large-scale radar echo data and elaborately design a DL architecture named designed a deep learning (DL) architecture called 'Multi-scale Feature Fusion' (MFF), which . The MFF model focuses on detecting radar echo multi-scale feature (e.g. features of radar echoes such as intensity, movement direction, and speed), and , which is expected to improve forecast skillsfor enhance precipitation forecasting skills, particularly in predicting precipitation growth and dissipation, fast-moving precipitation systems, and heavy precipitations. The rest of this This article is organized as follows: Section 2 presents the data materials, the detailed method, and the framework of the model.

Section 3 describes the experimental results ~~including~~ from two precipitation cases ~~, and discuss~~ and discusses the advantages and disadvantages of the four models. Finally, Section 4 ~~draws conclusions~~ concludes and explores some possible future works.

2 Materials, Methods, and Models

2.1 Radar Reflectivity Image Products

Weather radar is ~~the main monitoring instrument for~~ a crucial tool for monitoring precipitation systems and severe convective weather events such as hail, ~~gale, tornado~~gales, tornadoes, and flash ~~flood~~floods. As of November 2022, the China Meteorological Administration has ~~deployed the China Next Generation Weather Radar (CINRAD) network composed~~ installed a network of 236 C-band and S-band Doppler weather radars ~~over China(~~across China, known as the China Next Generation Weather Radar (CINRAD) network (see Fig. 1). ~~The~~ However, the CINRAD network is distributed heterogeneously across China except in complex terrain (Min et al., 2019)~~, and measures the moving speeds of the meteorological target relative to radars and further inverts~~ . The network measures the speed of meteorological targets relative to the radars and then produces various types of meteorological products. This study ~~collects and sorts out~~ focuses on collecting and organizing radar reflectivity image products ~~of~~ from five seasons (March to August) ~~from~~ between 2018 ~~to 2022, its temporal resolution is~~ and 2022. The data has a temporal resolution of 6 ~~min and its coverage area is~~ minutes and covers an area over $(73^oE - 135^oE, 10^oN - 55^oN)$. The data pre-processing steps ~~are as follows~~include the following:

(i) ~~Because radar echoes are affected by low-altitude objects (e.g. massif, building, tree, etc), so sham echoes are often produced at low-elevation areas~~Low-altitude objects such as mountains, buildings, and trees can produce sham echoes in radar images. Therefore, we ~~firstly remove the~~ remove these anomalous radar echoes and detach ~~the surplus annotations (e.g. city name, demarcation, and river)~~ unnecessary annotations like city names, demarcations, and rivers from each image. ~~Secondly, to reduce the influence~~ To reduce the impact of sham echoes on the extrapolative model, we use a local-mean filter algorithm ~~for radar image denoising, and then~~ to denoise the radar images. After this, we transform the radar reflectivities ~~are transformed~~ into precipitation values ~~based on~~ using the Z-R relationship.

(ii) The extrapolative model ~~will be~~ is difficult to converge ~~due to the great~~ because of the significant numerical differences among each echo reflectivity. ~~Therefore~~As a result, we normalize the initial radar reflectivities to ~~the $[0, 1]$ range . Then~~ a range of $[0, 1]$. Also, to assign precipitation values in areas without radar echo~~are assigned~~ , we set them as 0. Finally, we resample the precipitation values on ~~1024×880 grids~~ a 1024×880 grid for each radar image~~, while its actual spatial resolution is about~~ . The spatial resolution of one radar image, combining all grid boxes, is approximately 5 km. After completing the data preprocessing steps, we obtained $20,5848$ samples, which is a 3D matrix with ~~a size of $20,5848 \times 1024 \times 880$ is obtained~~the size of $[20,5848 \times 1024 \times 880]$.

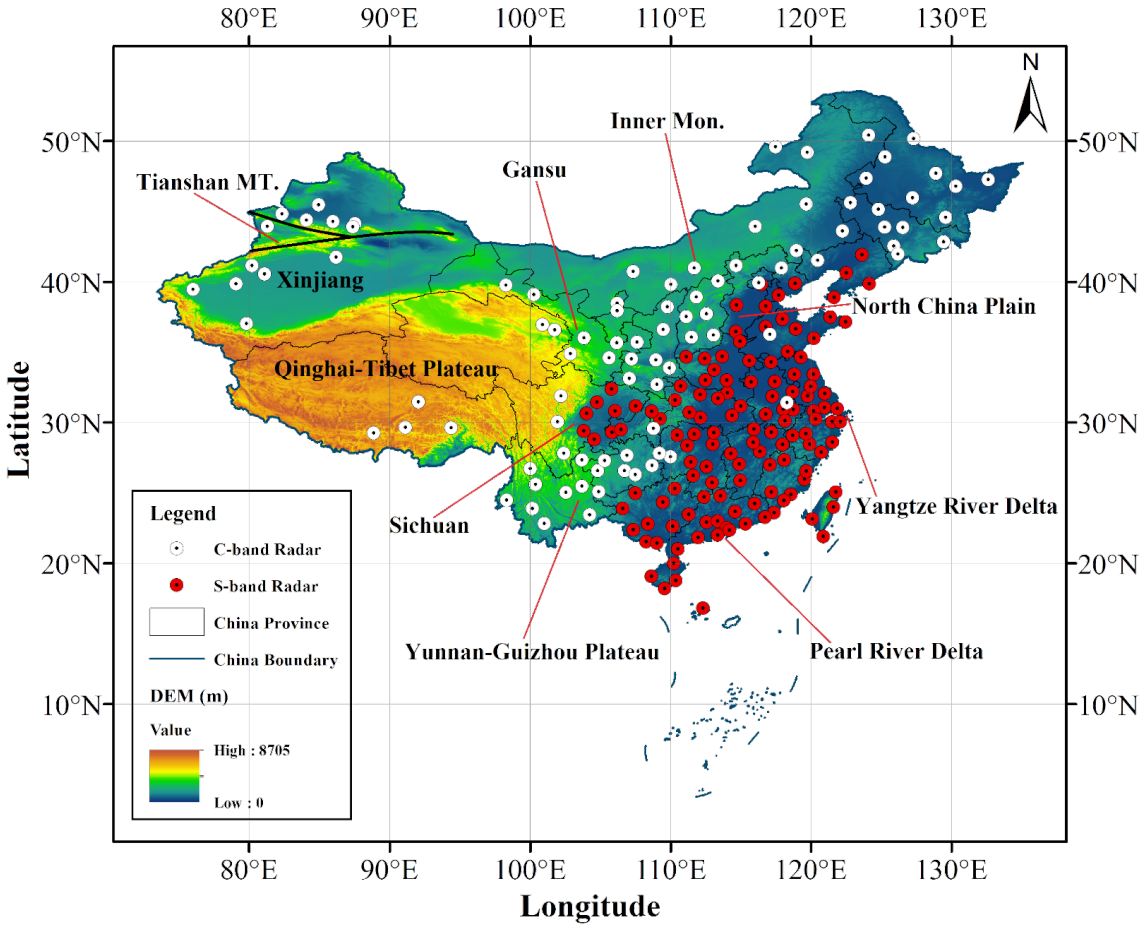

Figure 1. The distribution of the CINRAD over China and Topography (unit: m) map. White dots represent C-band radars and red dots denote S-band radars.

2.2 Multi-scale Feature Fusion (MFF)

~~The extrapolative technique of radar echo~~ Radar echo extrapolation is an important ~~vehicle for precipitation nowcasting with investigating several~~ technique for predicting precipitation by analyzing key variables such as ~~the~~ convective cloud intensity, shape, movement direction, and ~~moving speedof convective cloud~~speed. However, ~~there are different targets (e.g.~~ echo images may have different targets such as light rain, moderate rain, and heavy rain ~~) in an echo image or significant differences in the size of the same target collected at various resolutions. Meanwhile~~or the same target may vary in size at different resolutions. Additionally, in a ~~certain region of interest of~~ specific area of interest in an echo map, there may be ~~situations of tight arrangement and disorderly distributions of multiple targets (not least the local strong convection) which inevitably induce background noises~~multiple targets arranged in a tight or disorderly manner, which can cause background noise, particularly due to strong local convection. Therefore, using a single ~~/unique feature (refer to as 'convolution kernel ' in a DL architecture~~

~~) will lead to low forecast skills~~ feature or convolution kernel in a Deep Learning architecture can lead to lower forecasting accuracy due to the ~~relatively small receptive fields~~limited receptive field.

This study ~~proposes~~ introduces two modules for feature fusion: a 3D Multi-scale Feature Fusion (3DMFF, Fig. 2a) module and a 2D Multi-scale Feature Fusion (2DMFF, Fig. 2b) module. ~~An important part of the~~ The 3DMFF ~~is to apply convolution kernels with various sizes to gain~~ module uses convolution kernels of different sizes to capture information from different receptive fields. ~~Given~~ Assuming that the average moving speed of ~~the~~ a convective cloud is 36 km/h, the largest convolution kernel with the size of $4 \times 5 \times 5$ ~~could~~ can capture the traceability information of the convective cloud under this moving speed. Conversely, the smallest convolution kernel with the size of $4 \times 2 \times 2$ is geared toward the slow-moving clouds. ~~Besides, a key convolution kernel with the size of~~ Additionally, a $4 \times 1 \times 1$ ~~is also used which is instrumental in~~ kernel is used for dimensionality adjustments and information ~~interactions~~ interaction among channels. The ~~above multi-scale feature are then concatenated so that the module could~~ outputs of these different scale features are concatenated to store more information from the previous echo maps. Similarly, the 2DMFF ~~uses various convolution kernels with~~ module uses convolution kernels of sizes ranging from 1×1 to 4×4~~. Furthermore, we introduce the '~~, and employs the 'Channel-Shuffle~~' technique (Zhang et al., 2018)~~ ' technique to randomly shuffle the concatenated feature maps along the channel dimensions~~which~~ . This enhances the feature interaction ability between channels and ~~further~~ improves the generalization ability of the module. ~~The~~ Both the 3DMFF and 2DMFF ~~module both apply the 'Relu'~~ modules use the 'ReLU' activation function for nonlinear mapping~~which thins~~ , which helps to thin the network and ease the over-fitting problem to a certain extent.

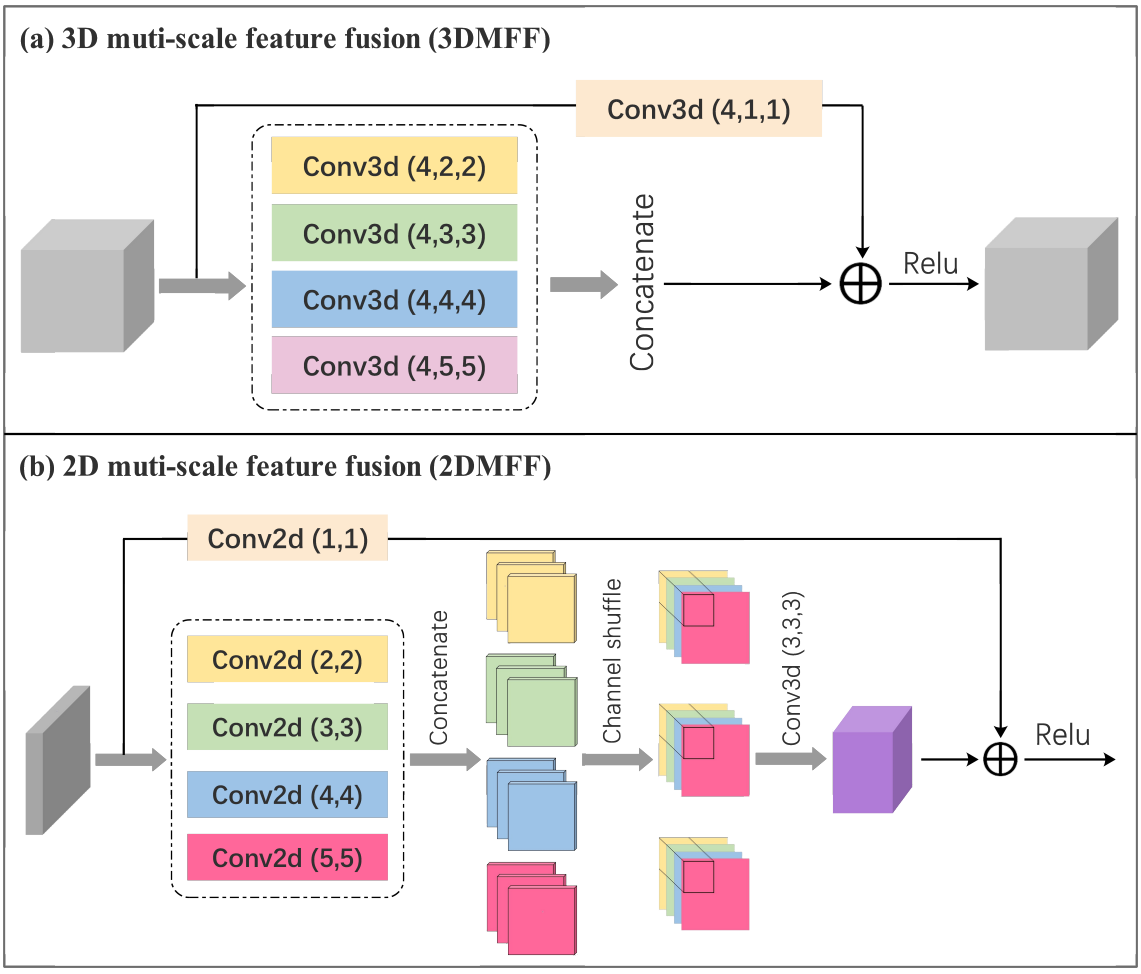

Figure 2. (a) The 3D Multi-scale Feature Fusion (3DMFF) and (b) the 2D Multi-scale Feature Fusion (2DMFF).

Consequently, as compared to the conventional ~~single feature~~ single-feature module, the ~~MFF modules make full use of the information of~~ Multi-Feature Fusion (MFF) modules use different receptive fields ~~and enhances the feature interaction ability by increasing~~ to enhance feature interaction and increase the number of network routes. ~~That is precisely the conventional single feature module is unable~~ This enables the MFF modules to fully extract feature information ~~due to less network routes(refer to as the issue of information loss). In addition~~that was previously lost due to fewer network routes. Additionally, the MFF modules introduce channel sorting and spatial-temporal convolutions to address the issue of information redundancy.

2.3 The Framework of the Nowcasting Model based on MFF

~~We present the complete framework~~ Here is a detailed description of the precipitation nowcasting model framework (Fig. 3) ~~. Overall, the model uses 60-min~~ that we developed. The model is trained using 60-minute radar echo maps~~for training with the~~, with an input size of $(1, 10, 880, 1024)$, and ~~generates the 180-min nowcasting outputs with the~~ produces nowcasting outputs of

180 minutes, with a size of $(1, 30, 880, 1024)$. The model ~~consists of two steps~~has two main components: the encoding ~~network and the decoding network. Where the encoding network contains a series of down-sampling layers of initial features based on several~~ and decoding networks. The encoding network uses multiple 3DMFF modules ~~, it plays a role in feature extraction and information compression. As for~~ to extract features and compress information, while the decoding network ~~, it applies several~~ involves feature restoration and up-sampling using 3D transpose convolutions and 2DMFF modules~~for feature up-sampling and feature restoration. Note that the~~ . The 3D transpose convolutions ~~also~~ generate a tensor (see P in Fig. 3) ~~which can be deemed~~ that acts as the probability matrix~~. To fully restore the features of the decoding networkto~~ , retaining the intensity information of radar echo for predicting various precipitation systems such as light rain, moderate rain, and heavy rain. To restore the ~~incipient~~ decoding network's features to the input's features, we perform two Hadamard product operations~~are performed: one is~~ . The first operation multiplies the output features of the 2DMFF ~~multiplied~~ by the probability matrix (see $m \odot P$ in Fig. 3)~~while another is the~~ , while the second operation multiplies the output features of the 3D transpose convolutions ~~multiply~~ by $1 - P$ (see $(1 - P) \odot f_1$ in Fig. 3). ~~The action of the probability matrix is that it retains the most of intensity information of radar echo as much as possible so that various precipitation systems (e.g. light rain, moderate rain, and heavy rain) can be well predicted. Because the~~ Since the outputs from 3D transpose convolutions lack edge information~~(since the use of the padding strategy of 0), so these outputs are also concatenated~~ , we concatenate them with the outputs from the 2DMFF modules to reduce information loss. ~~Finnally~~Finally, we apply a 3D convolution operation to adjust the channel of the product outputs and ~~further~~ generate the precipitation nowcasting results.

By drawing lessons from 'MetNet' (Sønderby et al., 2020), suppose the target weather condition is y, and the input condition is x, therefore,

$$p(y|x) = DNN_\theta(x) \tag{1}$$

Where $p(y|x)$ is a conditional probability over the output target y given the input x, $DNN_\theta(x)$ is a deep neural network with parameters θ~~, the model introduces uncertainties due to the calculation of the~~ . The model used in this case introduces uncertainties because it calculates the probability distribution over possible outcomes and does not provide a deterministic output. In most cases, the radar echo reflectivity is a continuous variable, ~~hence we~~ and we need to discretize the variable into a series of intervals ~~and then~~ to approximate the probability density function of the variable. ~~Because the~~ By using a discrete probability model, we can reduce uncertainties~~, therefore,~~ . Therefore, the combination of discrete probability and radar echo reflectivity ~~will further~~ can significantly reduce uncertainties of extrapolative radar echo. Here, we ~~invoke~~ use a mechanism of discrete probability as follows:

$$y^{[\tau]} = \Sigma \sum_{i=1}^{c} p_i^{[\tau]}(y^{[\tau]}|x) \cdot x_i \tag{2}$$

Where $y^{[\tau]}$ is the output at a given time τ , x is the input condition, c is the number of channels, $p_i^{[\tau]}(y^{[\tau]}|x)$ is a conditional probability~~at a channel at time τ~~. Eq. (2) shows the information of multiple channels at ~~a certain~~ time τ . Here, ~~one channel corresponds to one probability valuesuggesting that the probability is assigned to each channel to conduct better feature extraction. We multiply the~~ each channel has its own probability value, which is used to extract better features. The conditional

probability $p_i^{[\tau]}$ is multiplied by related channel information ~~x_i and then calculate their summation~~ xi, their sum is calculated over all channels ~~so that~~ to abtain more realistic radar echo reflectivities~~are achieved. As can be seen that~~ . The mechanism of discrete probability is used by both $m \odot P$ and ~~$1 - P$ (see $(1 - P) \odot f_1$ in~~ $(1 - P) \odot f_1$ (see Fig. 3~~use the mechanism of discrete probability~~).

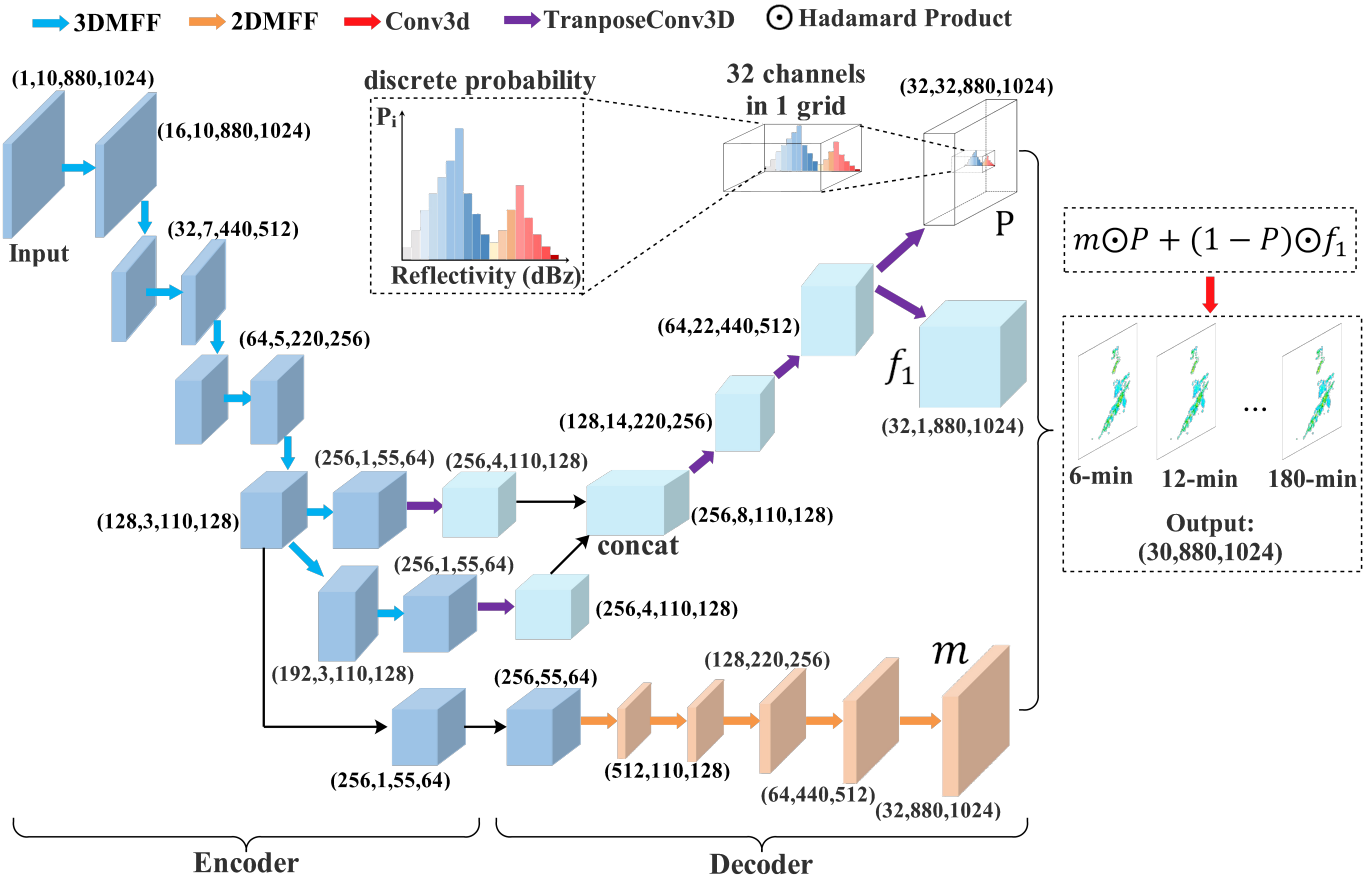

Figure 3. The framework of the nowcasting model based on Multi-scale Feature Fusion (MFF).

Overall, the ~~framework of the nowcasting model with~~ nowcasting model has a deep and hierarchical encoding-decoding backbone ~~is instrumental in extracting the~~ that helps to extract essential features from ~~the inputs, while~~ inputs. It also has several plug-and-play modules ~~are suitable to excavate the context informationor meticulous texture featuresof the inputs and reduce background noises of the inputs, making the model effectively to investigate~~ suitable for excavating context information, reducing background noise and identifying texture features. This makes the model effective in investigating the movement vector features of precipitation ~~system (e.g.~~ systems such as shape, movement direction, and moving speed~~) in the practical nowcasting. Moreover, the model introduces~~ . The model also uses the mechanism of discrete probability to ~~skillfully~~ reduce

uncertainties and forecast errors, ~~making the model~~ which helps to postpone the declining rate of strong-intensity echoes to some extent. Therefore, the model can produce heavy rains with longer lead times.

2.4 Comparative Models

To ~~have a comprehensive comparison,~~ provide a thorough comparison, here we also present three radar echo extrapolation models~~are also presented here~~.

2.4.1 Optical Flow (OF)

The ~~radar echo extrapolating problem can be regarded as moving object detection which separates the~~ problem of radar echo extrapolation can be seen as detecting moving objects, which involves separating targets from a continuous sequence of images. Gibson (1979) ~~proposed~~ introduced the concept of ~~OF characterizing an~~ optical flow (OF), which characterizes the instantaneous velocity of pixel motion of a space object in an imaging plane. ~~Specifically, the OF uses~~ The OF method employs the variation of a pixel of the image sequence in the time domain and the correlation between two adjacent frames ~~, thereby investigating~~ to investigate the movement information of objects between consecutive frames. ~~Generally~~Essentially, the transient variation of a pixel on a certain coordinate of the 2D imaging plane is defined as an optical flow vector. The OF method ~~satisfies two basic hypotheses: the~~ relies on two basic assumptions: grey-scale ~~invariant and the tiny~~ invariance and the small movement of pixels between consecutive frames.

Let $I(x,y,t)$ be the grey-scale value of the pixel at position (x,y) and time t, it moves (dx,dy) units of distances using dt units of time. Based on the grey-scale invariant hypothesis, the grey-scale value remains unchanged between two adjacent times, so the following equation holds:

$$I(x,y,t) = I(x+dx, y+dy, t+dt) \tag{3}$$

Using Taylor expansion, the right term of Eq. (3) becomes:

$$I(x,y,t) = I(x,y,t) + \frac{\partial I}{\partial x}dx + \frac{\partial I}{\partial y}dy + \frac{\partial I}{\partial t}dt + \epsilon \tag{4}$$

Where ϵ represents the infinitesimal of the second order which is negligible. Then substitute Eq. (4) into Eq. (3) and divide by dt, therefore we have:

$$\frac{\partial I}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial I}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial I}{\partial t} = 0 \tag{5}$$

Suppose $u = dx/dy$ and $v = dy/dy$ are two velocity vectors of optical flow along the x-axis and the y-axis, respectively. Let $I_x = \frac{\partial I}{\partial x}$, $I_y = \frac{\partial I}{\partial y}$ and $I_t = \frac{\partial I}{\partial t}$ are the partial derivative of the grey-scale of pixels along the x-axis, the y-axis, and the t-axis, respectively. Therefore, Eq. (5) turns into:

$$I_x u + I_y v + I_t = 0 \tag{6}$$

Where I_x, I_y, and I_t can be calculated from the original image data, while (u,v) are two unknown vectors. Because Eq. (6) is a constraint equation but has two unknown variables. Therefore, it is necessary to add other constraint conditions to

calculate (u, v). Currently, there are two common algorithms used ~~by solving~~ to solve this problem: global optical flow (Horn and Schunck, 1981) and local optical flow (Lucas and Kanade, 1981), detailed mathematical derivations of the two algorithms do not expatiate here.

2.4.2 UNet

The second comparative model is U-Net. ~~The biggest difference between~~ Unlike the MFF model~~and the U-Net model is the latter~~ , the UNet model uses general 2D convolution ~~in place~~ instead of the 'MFF module'. ~~There are mainly three parts in the U-Net model~~It consists of three main parts. The first part, called the encoder module, is a backbone network ~~(encoder module) used for~~ that performs down-sampling and feature extraction~~, and is stacked by~~ . It is composed of several convolution layers and max-pooling layers. ~~Based on the output features from the first part, the second part (decoder module)~~ The second part, called the decoder module, uses several up-convolution layers and convolution layers to conduct up-sampling and strengthen feature extraction~~, so that the features can be fused more effectively. The~~ . This allows for effective feature fusion based on the output features from the first part. Finally, the third part is a prediction module ~~which~~ that is used for a specific task, such as regression and segmentation. ~~In addition~~Additionally, to ensure the down-sampling feature's size matches the up-sampling feature's size~~and further reserves~~ , and to further preserve more original information, the ~~operation of~~ 'feature copying and cropping' ~~is also needed~~operation is also required.

2.4.3 Time Series Residual Convolution (TSRC)

The third comparative model is ~~the TSRC model proposed by~~ TSRC proposed n our previous study (Huang et al., 2023)~~, detailed mathematical derivations of the TSRC are omitted here. The core idea of TSRC is that it compensates~~ . The model compensates for the current local radar echo features with previous features during convolution processes on a spatial scale. ~~Moreover, the model implants~~ It also incorporates 'time series convolution' to ~~ease the~~ minimize dependencies on spatial-temporal scales~~so that~~ , resulting in the preservation of more contextual information and ~~less uncertain features are reserved~~ fewer uncertain features in the hierarchical architecture. ~~Especially, the model exhibits good performance in dealing with~~ The model has shown excellent performance in handling the smoothing effect of the precipitation field and the degenerate effect of the echo intensity. For detailed mathematical derivations of the TSRC model, please refer to our previous study.

2.5 Evaluation Metrics

We utilize five ~~evaluation metrics to examine the forecast skills of the three extrapolative models, including the~~ metrics to assess the forecast accuracy of three models: probability of detection (POD), false alarm rate (FAR), mean absolute error (MAE), radially averaged power spectral (RAPS), and structural similarity index (SSIM).

$$POD = \frac{successful\ forecast}{successful\ forecast + missing\ forecast} \tag{7}$$

$$FAR = \frac{null\ forecast}{successful\ forecast + null\ forecast} \tag{8}$$

Where successful forecast, missing forecastIn practical precipitation tasks, it is common to encounter successful forecasts, missing forecasts, and null forecast typically appear in practical precipitation tasks. The above three values forecasts, which are determined by the comparison between comparing the ground true value (GTV), forecast value (FV), and threshold value (TV). In practical precipitation tasks (Huang et al., 2023), Here, the threshold of 20 dBz represents those is used to represent reflectivity values greater than 20 dBz (hereafter referred to as '∼20 dBz'). Similarly, the term '∼30 dBz' and '∼40 dBz' can be abbreviated. In this study , we adopt three thresholds which are set as This study adopts three thresholds (20, 30, and 40 dBz. For example,). To determine the occurrence of successful forecast events, mark one if GTV $\geq$ TV and FV $\geq$ TV, then mark one successful forecast event; . For missing forecast events, mark one if GTV $\geq$ TV and FV < TVthen mark one missing forecast event; . For null forecast events, mark one if GTV < TV and FV $\geq$ TV, then mark one null forecast event. Both POD and FAR intuitively describe the . The performance of growth and dissipation forecasting tasks . can be intuitively described by both POD and FAR.

$$MAE = \frac{1}{n}\Sigma\sum_{i=1}^{n}|Y_i^g - Y_i^f| \tag{9}$$

Where Y_i^g and Y_i^f are the ground truth value and forecast value in the $i-th$ pixel of the related echo image, and n is the total number of pixels. This metric describes the performance of each forecast model at different precipitation intensity levels.

This studyregards the grey-scale of radar echoas In this study, the radar echo's greyscale is considered a signal. The power spectrum describes the magnitude of different signal frequency componentsdifferent frequency components' magnitudes of a 2D image , therefore it is treated by signal. Therefore, we use the Fourier transform to convert it from the spatial domain into to the frequency domain (Braga et al., 2014). Different frequency components within in the power spectra are located at different varying distances and directions from the base point on the frequency plane. High-frequency components are located more distant farther from the base point, and different directions from the base point indicate different orientations of the data features. Here, we use RAPS (Sinclair, 2005; Ruzanski, 2011) to To investigate the smoothing effect of forecast radar echo maps and discuss the forecast skill on local convection. Detailed , we use RAPS (Sinclair, 2005; Ruzanski, 2011). Here are the mathematical derivations of RAPS are omitted here.

in detail. First, we perform a 2D Fourier Transform on a 2D input image.

$$F(u,v) = \mathcal{F}\{f(x,y)\} \tag{10}$$

where $F(u,v)$ is the representation of complex domain after Fourier Transform, $\mathcal{F}$ is the Fourier operator, $f(x,y)$ is the input image. And then we calculate the power spectral P and radial coordinate r in the frequency domain.

$$P(u,v) = |F(u,v)|^2, r(u,v) = \sqrt{u^2+v^2} \tag{11}$$

Last, the power spectrum are grouped according to the radial coordinate of frequency; subsequently, take the average. For each radius r_k, its corresponding radially averaged power spectral P_k is

$$P_k = \frac{1}{N_k} \sum_i P(u_i, v_i) \tag{12}$$

Where $r(u_i, v_i) \approx r_k$, and N_k is number of frequency point falling within the radius of r_k.

Besides, we calculate the SSIM (Wang et al., 2004) to examine the similarity of precipitation fields between ground true and forecasting radar echo maps.

$$SSIM = \frac{(2\mu_g\mu_f + c_1)(2\sigma_{gf} + c_2)}{(\mu_g^2 + \mu_f^2 + c_1)(\sigma_g^2 + \sigma_f^2 + c_2)} \tag{13}$$

Where μ_g and μ_f are the means of ground truth and forecasting radar echo map, σ_g and σ_f are the related standard deviation, σ_{gf} is the covariance, respectively. c_1 and c_2 are two constants. This metrics reflects the movement of precipitation ~~field~~ fields between ground truth and forecasting radar echo map.

3 Results

3.1 Overall Forecast Performances on Testing Data

We use four years' data (~~2018-2011~~2018-2021) for model training and one year's data from 2022 for model testing. In Fig. 4~~shows~~, we show the four evaluation metrics: POD, FAR, MAE, and SSIM in three reflectivity intervals of ∼20, ∼30, and ∼40 dBz. Overall, we observed that POD in the four models consistently ~~plunges~~ decreases with increased forecast lead time for all reflectivity ~~internals, while it is conversely for FAR~~ intervals, while FAR increases. The rankings of POD (or FAR) are quite different ~~from~~ for the three reflectivity intervals. ~~For example, in~~ In the ∼20 dBz reflectivity interval, MFF ranks the highest in POD during the ~~whole~~ entire forecast period, ~~it remains~~ remaining stable ranging from 0.6 to 0.8, which is almost twice that of TSRC, OF, and UNet after the ~~2 h lead time; however~~2-hour lead time. However, MFF and TSRC ~~hold the~~ have nearly equal FAR~~which are~~, which is roughly half of that ~~from~~ of OF and UNet. In the ∼30 dBz reflectivity interval, TSRC ranks ~~the highest~~ highest in POD, followed by MFF~~; coincidentally~~. Coincidentally, TSRC also ranks ~~the highest~~ highest in FAR before the 1-hour forecast time, while both MFF and TSRC obtain relatively low FAR compared ~~with~~ to that of OF and UNet. In the ∼40 dBz reflectivity interval, POD in TSRC is ahead of ~~that of~~ the other three models, especially before the 1-hour lead time, and it degrades into that of MFF at the longer lead time~~;~~. POD in both OF and UNet ~~are~~ remains lower than 0.2 during the ~~whole~~ entire forecast period and nearly ~~decline~~ declines to 0 after ~~2 h~~ the 2-hour lead time; MFF reports the lowest FAR during the ~~whole~~ entire forecast period. While the value of FAR climbs from about 0.1 to 0.9~~;~~, TSRC has a relatively stable FAR, while the value of FAR is higher than 0.5 during the ~~whole forecast period;~~ entire forecast period. FAR in OF and UNet rapidly ~~increase~~ increases from about 0.1 to 0.8 ~~at~~ in the first 90 ~~min;~~ minutes. FAR in all models ~~are~~ is greater than 0.8.

Although MFF produces relatively low POD in high reflectivity (∼30 and ∼40 dBz) intervals compared ~~with TSRC, however,~~ to TSRC, it obtains relatively low FAR at the same time. ~~It is evident from~~ From the definition of POD/FAR, it can

be understood that both more ~~'successful forecasts' and more 'null forecasts' ~~"successful forecasts" and more "null forecasts" occur in TSRC, ~~and conversely, fewer 'successful forecast' ~~while fewer "successful forecast" events and fewer ~~'null forecast' ~~"null forecast" events occur in MFF compared ~~with ~~to TSRC for high reflectivity intervals. ~~Taking ~~If we take POD=0.6 as a dividing point, it is ~~obvious ~~clear that MFF yields ~~'successful forecast' ~~"successful forecast" events for the whole forecast period in ∼20 dBz reflectivity~~; ~~, while TSRC, OF, and UNet gain ~~'successful forecast' ~~"successful forecast" events only before 60 ~~min~~minutes, 24 ~~min~~minutes, and 36 ~~min~~minutes, respectively. Similarly, ~~taking ~~if we take FAR=0.5 as a dividing point in ∼40 dBz reflectivity, it can be found that MFF, OF, and UNet report FAR<0.5 only before 2 ~~h~~hours, 30 ~~min~~minutes, and 30 ~~min~~minutes, respectively, suggesting that at least the three models can avoid half of the ~~'null forecast' ~~"null forecast" events before 2 ~~h~~hours, 30 ~~min~~minutes, and 30 ~~min~~minutes, respectively. However, TSRC is unavoidable ~~to produce 'null forecast' ~~in producing "null forecast" events for the whole forecast period in ∼40 dBz reflectivity.

~~In addition, we examine the MAE and SSIM between ~~We also analyze the Mean Absolute Error (MAE) and Structural Similarity Index Metric (SSIM) between the nowcasting and ground truth. ~~Overall, the MAE gradually rises ~~The MAE gradually increases with the forecast time for all models. ~~Specifically, in terms of MAE~~Out of the three models, MFF has the smallest MAE (~~about ~~around 15 dBz)~~which is ahead of ~~, which is better than both TSRC and UNet by ~~about ~~approximately 2 dBz reflectivity after ~~90-min, while ~~90 minutes. On the other hand, OF has the ~~largest MAE, especially ~~highest MAE, particularly in long forecast ~~time. These suggest ~~times. This indicates that MFF reproduces the precipitation intensity with relatively less overestimation or underestimation compared ~~with the other three ~~to the other models, while OF shows little capacity to do so, especially in a long forecast time. In terms of SSIM, ~~an important finding is that MFF keeps ~~it can be found that MFF performs well and maintains an upward trend, ~~OF enjoys a steadfast position, while ~~while OF remains consistent throughout the forecast time. However, TSRC and UNet show a downward trend, especially after 90 ~~min. These indicate ~~minutes. This indicates that MFF is ~~suitable to capture ~~capable of capturing the shapes of precipitation fields with ~~high with ~~relatively high similarity, and its forecast performance ~~increases ~~improves with forecast times.

Understandably, ~~MFF has a strong ability to dredge the movement vector features ~~the MFF model can identify movement vectors of precipitation systems and reduce uncertainties in ~~multi-scale and alleviate the issue of information loss (reduce uncertainties) in ~~high reflectivity intervals. ~~Therefore, by enforcing ~~By using the mechanism of discrete probability, the model ~~shows favorable superiority especially ~~is particularly effective for high-intensity precipitation systems even at longer forecast times. ~~By using the strategy of compensated information in time series~~However, the TSRC model ~~might ~~may struggle to replenish the information on precipitation intensity~~but inevitably brings the issue of ~~, leading to overestimation for the ~~whole forecast period , which leads to producing more 'null forecasts' events. OF ~~entire forecast period and producing more "null forecasts" events. On the other hand, the OF model produces precipitation fields based on the ~~grey-scale ~~gray-scale invariant and the ~~tiny ~~slight movement of the precipitation system, ~~so it is difficult to dig the fast changes of ~~making it difficult to detect fast changes in precipitation fields, especially in longer forecast ~~time, and also ~~times. Additionally, it tends to overestimate or underestimate ~~the ~~high-intensity precipitations. ~~UNet ~~Lastly, the UNet model only performs feature extractions ~~only ~~on the spatial scale~~which results in ~~, leading to information loss and ~~is unable to excavate the fast changes of ~~an inability to detect fast

changes in precipitation fields on the temporal scale and ~~the~~ high-intensity precipitations~~, therefore,~~. Consequently, it has poor ~~ability in nowcasting during the whole~~ performance in nowcasting throughout the forecast period.

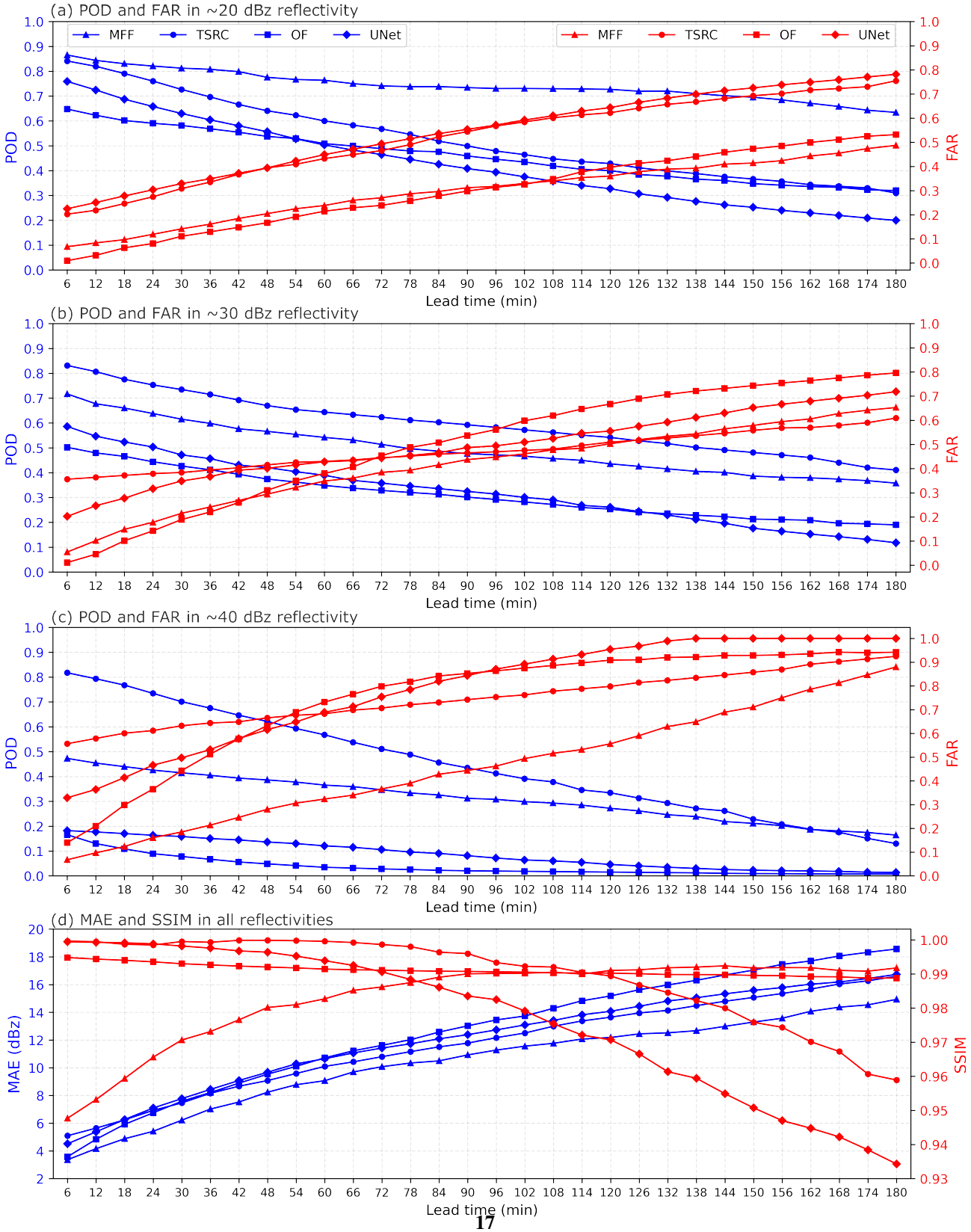

Figure 4. The forecast results in terms of four evaluation metrics on testing data ~~on~~ for the whole year of 2022.

3.2 Results of Case Study

Here we ~~show some forecast results of~~ present forecast results for two real precipitation cases (see Fig. 5) to ~~further understand the forecast~~ better understand the performance of the four models.

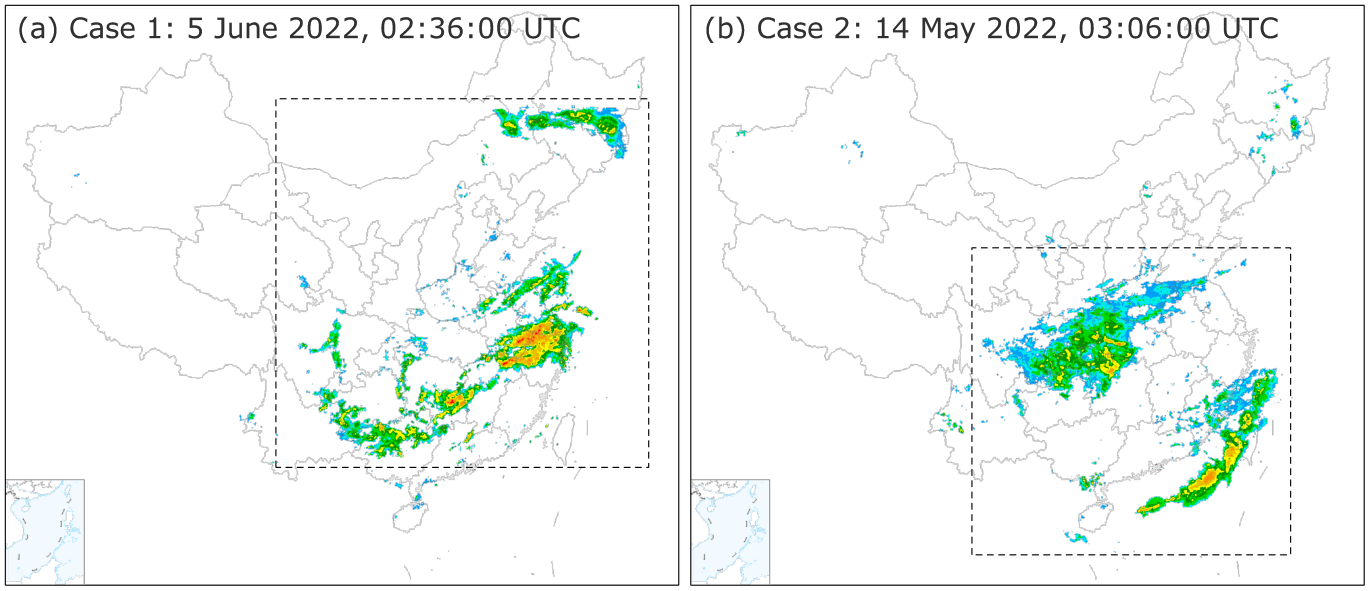

Figure 5. The overview of two precipitation cases over China.

3.2.1 Case 1

The first case is a large-scale precipitation process over China on 5 June ~~2025~~2022, 02:36:00 UTC (Fig. 5a), it contains the north part over northeastern China affected by a cold vortex and the south part (also known as 'dragon-boat rain') over southern China affected by warm-humid air. Fig. 6 presents the forecast results of this case. In the ground truth (GT), the whole precipitation area keeps a sluggish enlarging trend with the increased lead time, but the precipitation area of the high-intensity (e.g. greater than 35 dBz) echoes narrows gradually. As an important finding, MFF shows the best forecast performances since it can predict high-intensity (e.g. greater than 45 dBz) echoes even at the longest lead time (T+180 min). Comparatively, TSRC and UNet produce these echoes only at the short-range forecast time and miss them at the longest lead time. In terms of the precipitation field, both MFF and TSRC roughly capture the precipitation area, especially for low-intensity (e.g. less than 30 dBz) echoes at short-range lead times; OF draws an obvious dragged trajectory of the precipitation field in longer lead times, indicating the model simply creates precipitation fields with symbolic replications from the first frame to the last frame (T+180 min) at a horizontal scale and always misses the local changes of the precipitation system; UNet is ~~definitely~~ difficult to grasp the whole precipitation filed, not to mention the heavy precipitation system and its precipitation filed narrows gradually with the increasing lead time and finally disappears. The above analysis seems to be in accord with previous results that the high POD is reported in MFF and TSRC for low-intensity echoes (Fig. 4a), while the relatively steady POD in OF and UNet for

high-intensity echoes (Fig. 4b and 4c). Overall, MFF outperforms the other three models in predicting the precipitation field and the heavy precipitation.

On June 5th, 2022 at 02:36:00 UTC, a large-scale precipitation process occurred over China, affecting northeastern China and southern China differently. The northeastern part was affected by a cold vortex, while the southern part experienced warm-humid air, also known as 'dragon-boat rain'. In Fig. 5a, the forecast results of this case are presented, with the ground truth showing a sluggish enlargement of the whole precipitation area, but a gradual narrowing of high-intensity echoes (greater than 35 dBz). An important finding is that MFF has the best forecast performances, predicting high-intensity echoes even at the longest lead time (T+180 minutes). In contrast, TSRC and UNet only produce these echoes at short-range forecast times and missed them at the longest lead time. In terms of the precipitation field, both MFF and TSRC roughly capture it, especially for low-intensity echoes (less than 30 dBz) at short-range lead times. OF draws an obvious dragged trajectory of the precipitation field in longer lead times, indicating the model simply creates precipitation fields with symbolic replications from the first frame to the last frame (T+180 minutes) at a horizontal scale, always missing the local changes of the precipitation system. UNet is definitely difficult to grasp the whole precipitation field, not to mention the heavy precipitation system, and its precipitation field narrows gradually with the increasing lead time and eventually disappears. The above analysis is in line with previous results that MFF and TSRC have high POD for low-intensity echoes (see Fig. 4a), while OF and UNet have relatively steady POD for high-intensity echoes (see Fig. 4b and 4c). Overall, MFF outperforms the other three models in predicting the precipitation field and heavy precipitation.

~~The precipitation maps of case 1 (5 June 2025, 02:36:00 UTC); the first row: ground truth (GT); the second row: Multi-scale Feature Fusion (MFF); the third row: time series residual convolution (TSRC); the fourth row: optical flow (OF); the fifth row: UNet; the sixth row: the radially averaged power spectral density.~~

~~The RAPS is an important metric to intuitively~~ The RAPS metric is used to examine the smoothing and blurry precipitation fields~~, while the lower the power spectral , the smoother the precipitation fields~~. A lower power spectral indicates a smoother precipitation field. Conspicuously, OF enjoys a relatively high power spectral ~~which is close~~ that is comparable to that of the ground truth for the ~~whole~~ entire wavelength range. At first glance~~OF can predict the meticulous local-convection~~ , OF can accurately predict local convection activity and the evolution of the precipitation system. However, the ~~precipitation field is shifted by the model~~ model shifts the precipitation field from the first frame to the last~~frame and is to raise 'successful forecast' or 'null forecast' events, indicating the~~ , which results in poor forecast performance ~~of the model~~ at longer lead times. The other three DL-based models (MFF, TSRC, and UNet) have relatively low power spectra, ~~suggesting they inevitably~~ indicating that they introduce a smooth precipitation field to some extent~~, but~~ . But they might describe the evolution of the precipitation system more reasonably. ~~Specifically, MFF reports inconspicuous~~ MFF has a low power spectral under 4 km wavelength~~which is~~ , higher than that of TSRC and UNet~~, while this difference becomes increasingly apparent~~ . This difference becomes more significant at longer lead times, suggesting ~~the model has even more advantages to describe the local-convection~~ that the MFF model is better at describing local convection activity on a small scale~~and the model did~~ . Overall, the MFF model does at least ease the smoothing effect.

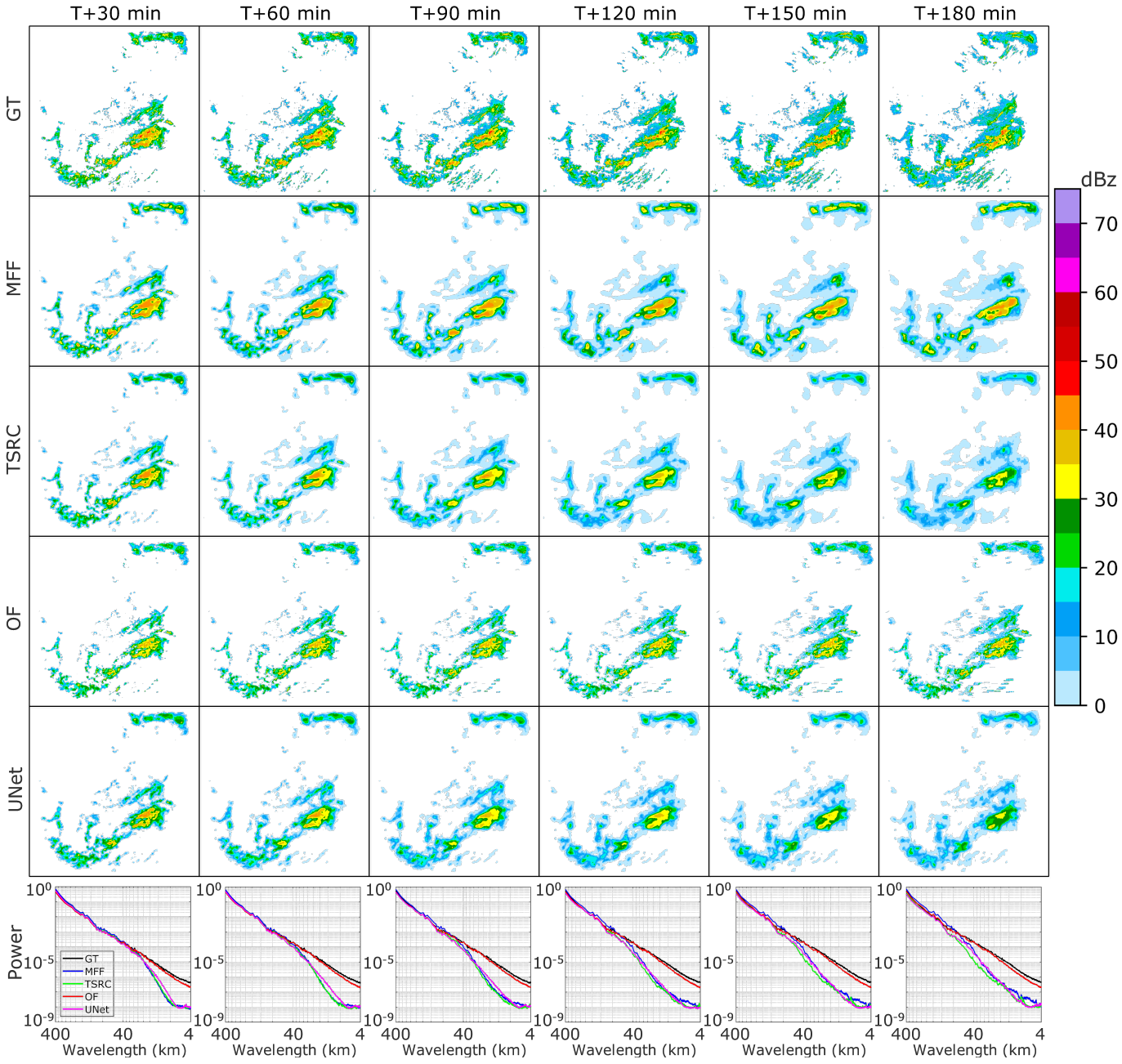

Figure 6. The precipitation maps of case 1 (5 June 2025, 02:36:00 UTC); the first row: ground truth (GT); the second row: Multi-scale Feature Fusion (MFF); the third row: time series residual convolution (TSRC); the fourth row: optical flow (OF); the fifth row: UNet; the sixth row: the radially averaged power spectral density.

3.2.2 Case 2

The second case occurred on 15 May 2022, On May 15, 2022 at 03:06:00 UTC, a significant precipitation event occurred in central China and offshore China (Fig. 5b). It was a large-scale precipitation process over central China and offshore of China, and was affected The event was influenced by the upper-level westerly trough, the southwest vortex, and the lower-level shear. Fig. 7 presents the forecast results of this case. As can be found from the ground truth, the precipitation field gradually expandsThe forecast results for this event are presented in Fig. 7, The precipitation area gradually expands, but the high-intensity area shrinks decreases with lead times. Overall, both Both the MFF and TSRC roughly exhibit models roughly follow the shape of the precipitation areas (on the land and the ocean) on land and sea and provide the evolutionary trend of the precipitation system. OF still However, the OF model shifts the precipitation field from the first frame to the last frameand certainly misses , missing the evolutionary trend of the precipitation systemespecially , particularly for longer lead times, which further proves the poor ability of the model in long-range precipitation forecasting. As the precipitation field grows smaller in UNet which is opposite to . This confirms the model's poor ability to forecast precipitation over long ranges. The UNet model proved to be the most challenging to use in capturing the ground truth, the model is quite difficult to capture the evolutionary trend of the precipitation system. In terms of echo intensity, the models have , as it reproduces the small precipitation field, which is the opposite of the ground truth. The four models show different forecasting performances in terms of echo intensity. MFF overestimates these strong-intensity echoes (greater than 30 dBz)but at the same time it enlarges , but it also increases the area of echoes at all lead times; TSRC is unable to produce these . TSRC can not produce strong echo intensities after the 120-min 120 minutes lead time. Although OF can predict these strong-intensity echoes at longer lead times, however, they are almost the same as the first framewhich indicates that the model performs poorly but the model's prediction is almost identical to the first frame, indicating a poor performance in predicting the evolution of strong-intensity echoes. Unfortunately, UNet shows the worst forecast performance since it underestimates these strong intensities at shorter lead times and can not produce deduce these strong-intensity echoes at longer lead times.

The three DL-based models report relatively low power spectra before the 90 min minutes lead time. OF obtain obtains relatively high power spectra, which are almost equal to that of the ground truth at all lead times, for the same reason that . This is because OF shifts the precipitation filed by field using an extrapolative technique. It is noteworthy that the power spectral in MFF is slightly greater than that in TSRC and UNet at longer lead timeswhich suggests , suggesting that the smoothing effect is further improved by MFF, therefore, this model is making it more suitable for precipitation forecasting both on the land and the oceanland and sea.

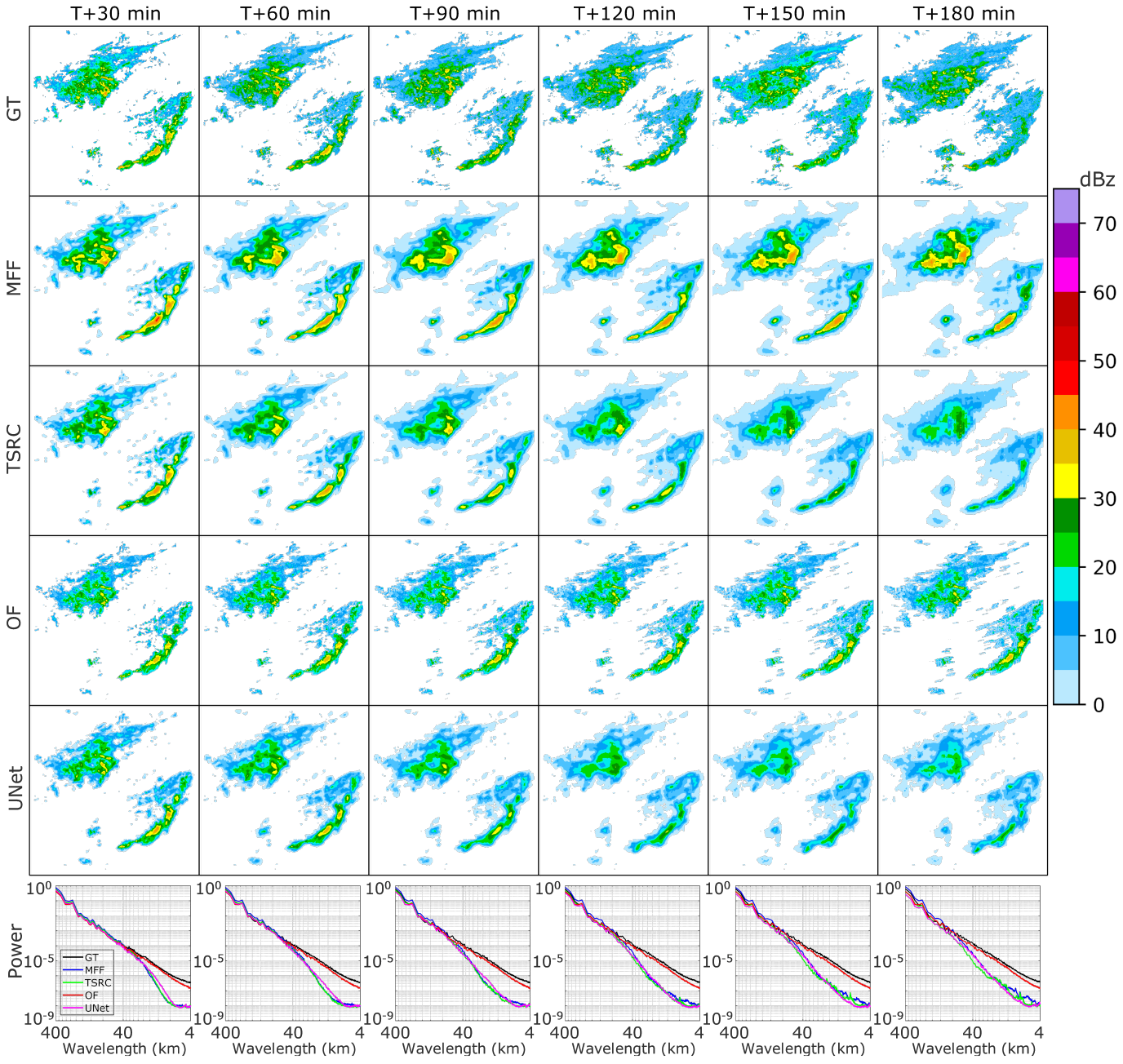

Figure 7. Same as in Fig. 6, but for the precipitation case 2 (15 May 2022, 03:06:00 UTC).

3.3 Discussions

Here we summarize the advantages and disadvantages of the four models in precipitation nowcasting.

3.3.1 MFF

The purpose of MFF is to improve the ~~forecast skill of heavy precipitations, especially~~ accuracy of precipitation forecasting, particularly at longer lead times. Current DL-based models for precipitation ~~nowcasting are facing two challenges: one is poor forecast skills when there are~~ forecasting are faced with two major challenges: the poor forecast skill when different precipitation systems with ~~various scales~~ varying scales are present; and the ~~other is the~~ low predictive accuracy when ~~various precipitation targets(e.g.~~ different precipitation targets, such as light rain, moderate rain, and heavy rain~~) ,~~ are densely distributed ~~at~~ in a certain area of interest and also introduce noises. From a qualitative perspective, ~~this study proposes MFF which~~ MFF proposes a deep and hierarchical encoding-decoding architecture that can make full use of the receptive fields to efficiently detect different precipitation systems in multi-scales and predict various precipitation targets. This superiority is unable to be achieved by the traditional single-scale receptive fields. However, ~~while this deep and hierarchical encoding-decoding~~ this architecture shows strong ability in feature extraction ~~, it~~ but might also account for the issue of information redundancy. Therefore, the model ~~introduces several crafts (e.g.~~ employs several techniques, such as channel shuffle, feature concatenation, and spatial-temporal convolution~~) ,~~ to enhance the feature interaction ability among multi-scales and further ease information redundancy. ~~The above operations~~ These techniques do obtain considerable forecast ~~performances~~ performance in several evaluation metrics~~: ~~, including POD, FAR, MAE, and SSIM. ~~In addition~~Additionally, the model skillfully applies the mechanism of discrete probability, which mathematically allocates the probability information into each channel and can reduce uncertainties and forecast errors to the most extent. The results of the case study further prove that only this model can produce heavy precipitations such as those greater than 45 dBz reflectivity radar echoes even at the 3 ~~h~~ hours lead time. It is noteworthy that two tricky issues~~(~~, namely the smoothing effect and cumulated error~~) ,~~ are still inevitably reported in the model~~, of course, they ~~. However, these issues are not specific to MFF~~while most DL-based ~~, as most deep learning-based models are also confronted~~with the same issues~~. The principal reasons account for ~~the issues~~ them include: the convolution strives to smooth multi-scale ~~feature~~ features in receptive fields to minimize fitting errors~~and the iterative discrepancy between ~~; and there are iterative discrepancies between the training processes and targets. Encouragingly, by introducing ~~Multi-scale Feature Fusion~~ the MFF and the mechanism of discrete probability, at least our models have some improvements~~which offer much promise for tackling many ~~. This offers a lot of promise for handling practical tasks such as precipitation growth and dissipation, fast-moving precipitation ~~system~~systems, heavy precipitation, ~~local convection activity, etc.~~and local convection activity...In any event, MFF is a DL-based and data-driven radar extrapolative model without any consideration of physical constraints and atmospheric dynamics~~, hopefully, ~~. Hopefully, the model can be further improved by ~~combining~~ adding multi-source data and combining ingenious DL architectures.

Furthermore, like other convolution-based DL models, the MFF model also requires highlighting the "inductive bias" to improve its generalization ability. Inductive bias can be thought of as a sort of "local prior". In the case of image analysis, the inductive bias in the MFF model mainly consists of two aspects. The first aspect is "spatial locality", which assumes that adjacent regions in a radar echo image always have relevant precipitation features. For example, the region of strong-intensity echoes is usually accompanied by the region of moderate-intensity echoes. However, this inductive bias may sometimes

overestimate the precipitation intensity (see case 2 in Figure 7), or enlarge the precipitation field, leading to accuracy issues. The other aspect is "translation equivariance", which means that when the precipitation field in the input map is translated, the precipitation field is also translated due to the use of local connection and weight-sharing in the multi-scale convolution process. This feature does allow the MFF model to trace the moving precipitation system. Therefore, as a widely-concerned weather phenomenon, extreme precipitations (e.g. 1/100 year rainfall events) may also be extrapolated and predicted by using inductive bias in the MFF model if both the training dataset and testing input provide precipitation events with very strong radar echoes. Conversely, it is also very challenging for the MFF model to tackle such a forecasting task.

3.3.2 TSRC

Essentially, TSRC is a reinforced ~~'encoding-deconding' architecture , it appends previous features into~~ "encoding-decoding" architecture that adds previous features to current feature planes on temporal scales during convolution processes~~, so~~. This allows more contextual information and less uncertain features ~~could~~ to remain in deep networks. The model ~~fully considers~~ takes into account the correlation of radar echo features on a temporal scale, ~~therefore, it should theoretically alleviate~~ which theoretically reduces the problem of information loss and the degenerate effect intensity. However, ~~those~~ the compensatory features in the architecture may lack specificity and carry noises, ~~resulting in the model mindlessly increasing~~ which causes the model to increase the precipitation intensities at the whole forecast lead times ~~. Understandably,~~ mindlessly. Although the model has relatively high POD~~though,~~ , it has high FAR and MAE~~can be found especially~~ , particularly for heavy precipitations. ~~Undoubtedly, the~~ The model increases the depth of the hierarchical architecture with different learnable parameters, excavates the dependencies of echo features on both temporal ~~scales~~ and spatial scales, and ~~also uses several crafts (e.g.~~ uses several techniques, such as feature concatenation, residual connection, ~~attention mechanism) to retard~~ and attention mechanism, to prevent the declining rate of intensity and the smoothing effect. The ~~results of testing data show great advantages of the model at those~~ testing data shows that the model has great advantages at real forecast tasks, such as low-intensity precipitation systems, and slow-change precipitation systems, especially for short lead times. However, the model lacks consideration of multi-scale features due to the fixed/unique receptive field on spatial scales~~, it lacks consideration of multi-scale feature which~~ . This leads to great difficulty in many real forecast tasks, such as local-convection activity, growth and dissipation, fast-moving precipitation systems, and rapid changes ~~of~~ in the rainfall field. ~~This has sparked speculation~~ Therefore, it is speculated that the model ~~will~~ can be further improved by implementing feature extraction on multi-spatial scales.

3.3.3 OF

The ~~core idea of~~ main principle behind OF is to ~~calculate the change of pixels from the image sequence~~ observe the variation of pixels in a sequence of images in the time domain~~and the dependence in~~ . By examining the correlation of two adjacent frames, ~~and thus investigate the information of moving objects. The forecast results from testing data show that OF is only suitable for the forecast of precipitation systems with slow change at very~~ the algorithm can detect the movement of objects. However, OF can only accurately forecast precipitation systems that have slow changes even at short lead times. The ~~reason lies in the two basic hypothesesof the model which are the~~ model relies on two fundamental hypotheses, namely the grey-scale ~~invariant~~

invariance and the tiny movement of pixels. ~~For the~~ The grey-scale ~~invariant, it makes the model difficult~~ invariance feature renders the model challenging to deal with ~~that precipitation system with rapid intensity changes. As for~~ precipitation systems characterized by swift intensity fluctuations. For the tiny movement of pixels, it can hardly satisfy the forecast of a fast-moving precipitation system..Unlike the DL-based models, OF ~~produces~~ generates precipitation fields based on Lagrangian persistence and smooth motion, which also fail to recognize both ~~the local features of echo and~~ local and multi-scale features of the ~~multi-scales features of echo, resulting in the~~ precipitation system. This leads to poor forecasting ability at longer lead times~~. Therefore, it is easy to understand OF yields forecast results by simply~~ , and the model often produces inaccurate results by merely shifting the precipitation fields~~, while the timeliness of precipitation forecasting and the accuracy are hard to guarantee. Even with further improved methods such as~~ . Although improved methods like the semi-Lagrangian method, which relies on the advection field, ~~it is still difficult to expound~~ have been developed, they still struggle to explain the complex features of the precipitation system.

3.3.4 UNet

~~There are three key steps in the UNet architecture which are~~ The UNet architecture involves three important steps, namely encoding, decoding, and skip connection. The encoding ~~part uses several~~ procedure uses multiple convolution layers for down-sampling and features compression~~, allowing~~ . This allows the contracting path to capture more context information. ~~Conversely~~On the other hand, the decoding ~~part~~ process applies several deconvolution layers for up-sampling and feature restoration~~, allowing~~ . This allows the expanding path to locate different features. The skip connection part fuses the pixel-level features and semantic-level features to achieve feature segmentation and reduce information loss. ~~Meanwhile,~~ While the bottom of the hierarchical architecture collects low-frequency information in the form of greater receptive fields~~but~~ , it fails to capture high-frequency information. ~~Therefore, when confronted with~~ As a result, when it comes to forecast tasks, the model may focus on those global (abstract or essential) features of precipitation systems but omit those exquisite changes in precipitation systems at spatial-temporal scales. ~~The~~ Since radar echoes usually have variability at ~~multi-scales, so it is not~~ multiple scales, it is insufficient for UNet to capture complex features of the precipitation system. The results from the case study also confirm that the model has poor forecast skills in fast-moving precipitations, high-intensity precipitations, growth, and dissipation, ~~and~~ as well as long-term forecasting. In ~~a word, UNet ranks~~ summary, UNet has the worst forecast performance among the three DL-based models.

4 Conclusions

~~In this study , we present~~ This study presents MFF, a ~~DL-based model~~ deep learning model designed for large-scale precipitation nowcasting with a lead time of up to 3 ~~h~~hours. The model aims to investigate the movement features of precipitation systems on ~~multi-scales. Moreover~~multiple scales. To reduce uncertainties and forecast errors, we introduce the mechanism of discrete probability in the model~~to reduce uncertainties and forecast errors. Three~~ . We compare our model with three existing radar echo extrapolative models~~which are~~ , namely TSRC, OF, and UNet~~are compared with our study~~. The comprehensive

analyses of testing data further prove the impressive forecast skills of MFF under four evaluation metrics: POD, FAR, MAE, and SSIM. ~~In addition, from the results of case studies at least,~~ MFF is the only extrapolative model that produces heavy precipitations even at the 3 ~~h~~ hours lead time, and the smoothing effect of the precipitation field is improved~~by our models~~. From an early warning perspective, the model shows a promising application prospect.

It is well known that data always determine the upper limit of a machine learning model, while algorithms only attempt to approximate this limit. Regretfully, the current study only considers the radar echo data as the model inputs. Therefore, ~~it is highly suggested to consider~~ we highly recommend considering more meteorological variables ~~(e.g.,~~ such as temperature, pressure, humidity, wind, etc.~~) ,~~ and ground elevations in future ~~work, while these data~~ works. These data can come from various sources such as radar observations, satellite sounding, reanalysis, real-time observation, NWP downscaling, etc. We believe that these multi-source data ~~would~~ can fortify some kind of physical or thermodynamic constraint for a pure data-driven extrapolative model. ~~As one of the thorny problems in most DL-based models, the smoothing effect~~ The smoothing effect remains a challenging task, and it is still reported as long as the convolution procedure is performed~~and remains a challenging task~~. Therefore, future works will focus on the structural adjustments of the network and the combinations with numerical models to further improve the forecast accuracy of heavy precipitations at longer lead times.

Code and data availability. The source codes and pretrained models are available on a Zenodo repository https://doi.org/10.5281/zenodo.8105573 (last access: 1 July 2023; Tan, 2023).

Author contributions. Conceptualization was performed by JT and SC; QH and JT contributed to the methodology, software and investigation, as well as the preparation of the original draft; QH and SC contributed to the resources and data curation, as well as visualizations and project administration; JT and SC contributed to the review and editing of the manuscript, supervision of the project and funding acquisition. All authors have read and agreed to the published version of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Financial support. This research was funded by GuangDong Basic and Applied Basic Research Foundation (No. 2020A1515110457), the China Postdoctoral Science Foundation (No. 2021M693584), the Guangxi Key R&D Program (No. 2021AB40108, 2021AB40137), the Innovation Group Project of Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai) (No. 311021001), and the Opening Foundation of Key Laboratory of Environment Change and Resources Use in Beibu Gulf (Ministry of Education) (Nanning Normal University, Grant No. NNNU-KLOP-K2103).

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. The authors would like to thank the reviewers for their valuable suggestions that increased the quality of this paper.

References

Ayzel G, Heistermann M, Winterrath T. Optical flow models as an open benchmark for radar-based precipitation nowcasting (rainymotion v0. 1)[J]. Geoscientific Model Development, 2019, 12(4): 1387-1402.

Ayzel G, Scheffer T, Heistermann M. RainNet v1. 0: a convolutional neural network for radar-based precipitation nowcasting[J]. Geoscientific Model Development, 2020, 13(6): 2631-2644.

~~BiK~~

Bi, K., Xie, L., Zhang, H., ~~Xie L, Zhang H,~~ et al. ~~Pangu-Weather: A~~ Accurate medium-range global weather forecasting with 3D ~~High-Resolution Model for Fast and Accurate Global Weather ForecastJ. arXiv preprint arXiv:2211.02556, 2022.~~ neural networks. Nature 619, 533–538 (2023).

Braga M A, Endo I, Galbiatti H F, et al. 3D full tensor gradiometry and Falcon Systems data analysis for iron ore exploration: Bau Mine, Quadrilatero Ferrifero, Minas Gerais, Brazil[J]. Geophysics, 2014, 79(5): B213-B220.

Chen K, Han T, Gong J, et al. FengWu: Pushing the Skillful Global Medium-range Weather Forecast beyond 10 Days Lead[J]. arXiv preprint arXiv:2304.02948, 2023.

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

Czibula G, Mihai A, Albu A I, et al. AutoNowP: An Approach Using Deep Autoencoders for Precipitation Nowcasting Based on Weather Radar Reflectivity Prediction[J]. Mathematics, 2021, 9(14): 1653.

Dupuy F, Mestre O, Serrurier M, et al. ARPEGE cloud cover forecast postprocessing with convolutional neural network[J]. Weather and Forecasting, 2021, 36(2): 567-586.

Ehsani M R, Zarei A, Gupta H V, et al. Nowcasting-Nets: Deep Neural Network Structures for Precipitation Nowcasting Using IMERG[J]. arXiv preprint arXiv:2108.06868, 2021.

Gibson J J. Ecological Approach to Visual Perception: Classic Edition[J]. Taylor and Francis eBooks DRM Free Collection, 1979.

Gibson, J.J. The Ecological Approach to Visual Perception: Classic Edition; Psychology Press: New York, NY, USA, 2014.

Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[J]. Advances in neural information processing systems, 2014, 27.

Han L, Sun J, Zhang W. Convolutional neural network for convective storm nowcasting using 3-D Doppler weather radar data[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 58(2): 1487-1495.

Horn B K P, Schunck B G. Determining optical flow[J]. Artificial intelligence, 1981, 17(1-3): 185-203.

Huang Q, Chen S, Tan J. TSRC: A Deep Learning Model for Precipitation Short-Term Forecasting over China Using Radar Echo Data[J]. Remote Sensing, 2023, 15(1): 142.

Kim D K, Suezawa T, Mega T, et al. Improving precipitation nowcasting using a three-dimensional convolutional neural network model from Multi Parameter Phased Array Weather Radar observations[J]. Atmospheric Research, 2021, 262: 105774.

Lam R, Sanchez-Gonzalez A, Willson M, et al. GraphCast: Learning skillful medium-range global weather forecasting[J]. arXiv preprint arXiv:2212.12794, 2022.Ravuri S, Lenc K, Willson M, et al. Skilful precipitation nowcasting using deep generative models of radar[J]. Nature, 2021, 597(7878): 672-677.

Lebedev V, Ivashkin V, Rudenko I, et al. Precipitation nowcasting with satellite imagery[C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2019: 2680-2688.

Li D, Liu Y, Chen C. MSDM v1. 0: A machine learning model for precipitation nowcasting over eastern China using multisource data[J]. Geoscientific Model Development, 2021, 14(6): 4019-4034.

Li L, He Z, Chen S, et al. Subpixel-Based Precipitation Nowcasting with the Pyramid Lucas–Kanade Optical Flow Technique[J]. Atmosphere, 2018, 9(7): 260.

Liguori S, Rico-Ramirez M A. A review of current approaches to radar-based quantitative precipitation forecasts[J]. International Journal of River Basin Management, 2014, 12(4): 391-402.

Liu Y, Xi D G, Li Z L, et al. A new methodology for pixel-quantitative precipitation nowcasting using a pyramid Lucas Kanade optical flow approach[J]. Journal of Hydrology, 2015, 529: 354-364.

Lucas, B.D.; Kanade, T. An Iterative Image Registration Technique with an Application to Stereo Vision. In Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI '81), Vancouver, BC, Canada, 24–28 August 1981.

Luo H, Wang Z, Yang S, et al. Revisiting the impact of Asian large-scale orography on the summer precipitation in Northwest China and surrounding arid and semi-arid regions[J]. Climate Dynamics, 2023, 60(1-2): 33-46.

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

Min C, Chen S, Gourley J J, et al. Coverage of China new generation weather radar network[J]. Advances in Meteorology, 2019, 2019.

Nguyen T, Brandstetter J, Kapoor A, et al. ClimaX: A foundation model for weather and climate[J]. arXiv preprint arXiv:2301.10343, 2023.

Pan X, Lu Y, Zhao K, et al. Improving Nowcasting of Convective Development by Incorporating Polarimetric Radar Variables Into a Deep-Learning Model[J]. Geophysical Research Letters, 2021, 48(21): e2021GL095302.

Prudden R, Adams S, Kangin D, et al. A review of radar-based nowcasting of precipitation and applicable machine learning techniques[J]. arXiv preprint arXiv:2005.04988, 2020.

Pulkkinen S, Nerini D, Pérez Hortal A A, et al. Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1. 0)[J]. Geoscientific Model Development, 2019, 12(10): 4185-4219.

Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.

Ruzanski E, Chandrasekar V. Scale filtering for improved nowcasting performance in a high-resolution X-band radar network[J]. IEEE transactions on geoscience and remote sensing, 2011, 49(6): 2296-2307.

Sønderby C K, Espeholt L, Heek J, et al. Metnet: A neural weather model for precipitation forecasting[J]. arXiv preprint arXiv:2003.12140, 2020.

Sadeghi M, Nguyen P, Hsu K, et al. Improving near real-time precipitation estimation using a U-Net convolutional neural network and geographical information[J]. Environmental Modelling and Software, 2020, 134: 104856.

Shi X, Chen Z, Wang H, et al. Convolutional LSTM network: A machine learning approach for precipitation nowcasting[J]. Advances in neural information processing systems, 2015, 28.

Shi X, Gao Z, Lausen L, et al. Deep learning for precipitation nowcasting: A benchmark and a new model[J]. Advances in neural information processing systems, 2017, 30.

Sinclair S, Pegram G G S. Empirical Mode Decomposition in 2-D space and time: a tool for space-time rainfall analysis and nowcasting[J]. Hydrology and Earth System Sciences, 2005, 9(3): 127-137.

Singh M, Kumar B, Rao S, et al. Deep learning for improved global precipitation in numerical weather prediction systems[J]. arXiv preprint arXiv:2106.12045, 2021.

Su A, Li H, Cui L, et al. A convection nowcasting method based on machine learning[J]. Advances in Meteorology, 2020, 2020: 1-13.

Sun J, Xue M, Wilson J W, et al. Use of NWP for nowcasting convective precipitation: Recent progress and challenges[J]. Bulletin of the American Meteorological Society, 2014, 95(3): 409-426.

Tan. (2023). Tan/Multi-scale-Feature-Fusion-for-Precipitation-Nowcasting (v2.1). Zenodo. https://doi.org/10.5281/zenodo.8105573

Trebing K, Staczyk T, Mehrkanoon S. SmaAt-UNet: Precipitation nowcasting using a small attention-UNet architecture[J]. Pattern Recognition Letters, 2021, 145: 178-186.

Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE transactions on image processing, 2004, 13(4): 600-612.

Woo W, Wong W. Operational application of optical flow techniques to radar-based rainfall nowcasting[J]. Atmosphere, 2017, 8(3): 48.

Yan Q, Ji F, Miao K, et al. Convolutional residual-attention: a deep learning approach for precipitation nowcasting[J]. Advances in Meteorology, 2020.

Yano J I, Ziemiański M Z, Cullen M, et al. Scientific challenges of convective-scale numerical weather prediction[J]. Bulletin of the American Meteorological Society, 2018, 99(4): 699-710.

Yasuno T, Ishii A, Amakata M. Rain-Code Fusion: Code-to-Code ConvLSTM Forecasting Spatiotemporal Precipitation[C]//International Conference on Pattern Recognition. Springer, Cham, 2021: 20-34.

Zhang X, Zhou X, Lin M, et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 6848-6856.

Zheng K, Liu Y, Zhang J, et al. GAN–argcPredNet v1. 0: A generative adversarial model for radar echo extrapolation based on convolutional recurrent units[J]. Geoscientific Model Development, 2022, 15(4): 1467-1475.