


RESEARCH LETTER

Open Access



Downscaling Taiwan precipitation with a residual deep learning approach

Li-Huan Hsu^{1*} , Chou-Chun Chiang¹, Kuan-Ling Lin¹, Hsin-Hung Lin¹, Jung-Lien Chu¹, Yi-Chiang Yu¹ and Chin-Shyung Fahn²

Abstract

In response to the growing demand for high-resolution rainfall data to support disaster prevention in Taiwan, this study presents an innovative approach for downscaling precipitation data. We employed a hierarchical architecture of Multi-Scale Residual Networks (MSRN) to downscale rainfall from a coarse 0.25-degree resolution to a fine 0.0125-degree resolution, representing a substantial challenge due to a resolution increase of over 20 times. Our results demonstrate that the hierarchical MSRN outperforms both the one-step MSRN and linear interpolation methods when reconstructing high-resolution daily rainfall. It surpasses the linear interpolation method by 15.1 and 9.1% in terms of mean absolute error and root mean square error, respectively. Furthermore, the hierarchical MSRN excels in accurately reproducing high-resolution rainfall for various rainfall thresholds, displaying minimal biases. The threat score (TS) highlights the hierarchical MSRN's capability to replicate extreme rainfall events, achieving TS scores exceeding 0.54 and 0.46 at rainfall thresholds of 350 and 500 mm per day, outperforming alternative methods. This method is also applied to an operational global model, the ECMWF's daily rainfall forecasts over Taiwan. The evaluation results indicate that our approach is effective at improving rainfall forecasts for thresholds greater than 100 mm per day, with more significant improvement for the 1- to 3-day lead forecast. This approach also offers a realistic visual representation of fine-grained rainfall distribution, showing promise for making significant contributions to disaster preparedness and weather forecasting in Taiwan.

Keywords Downscaling, Super-resolution, Deep learning, Residual network, Precipitation forecasting

Introduction

Since the 1950s, heavy precipitation has increased over most land areas, and it is likely that precipitation extremes will become more frequent and intense under projected scenarios (Masson-Delmotte et al. 2021). The demand of weather and climate model rainfall outputs for precipitation impact assessment has risen across

various disciplines in recent years. Specifically, local disaster pre-warning information, such as flooding, for example, heavily relies on high-resolution forecast rainfall data during evolving disastrous weather systems (Yu et al. 2016). However, due to the limitation of computational and storage resources, physical approximations or model parameterizations, most global models have spatial resolutions of 10–100 km or even coarser (Taylor et al. 2012; Bauer et al. 2015; Li et al. 2022). Users generally access pre-interpolated data with a resolution of 0.25–0.5 degrees through public data repositories. These resolutions are inadequate to resolve the convective scale systems associated with heavy rainfall and assess their local impacts (Schiermeier 2010; D'Onofrio et al. 2014). In Taiwan, with its cramped land area and high population density, rainfall data of spatial resolution smaller

*Correspondence:

Li-Huan Hsu
lhhsu@ncdr.nat.gov.tw

¹ National Science and Technology Center for Disaster Reduction, New Taipei City 231007, Taiwan

² Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei 106335, Taiwan

than 5 km is frequently required for effective disaster prevention in local areas. This poses a challenge for meteorological models to meet the demand.

National weather centers typically employ regional climate models (RCM) within a general circulation model (GCM) or use a regional weather model with initial and boundary conditions from a GCM to provide better representations of high-resolution precipitations (von Storch et al. 2000; Maraun et al. 2010; Giorgi and Gutowski 2015; Prein et al. 2015; Kaur et al. 2020). These dynamical downscaling approaches generate more small-scale regional forecast information, but they require significant computational resources on super-computing clusters and need to be developed individually. However, the performance of dynamical downscaling is still affected by the global forecasting model and may contain its own model errors (May 2004; Xue et al. 2014; Hong and Kanamitsu 2014; Lucas-Picher et al. 2021). Statistical downscaling projects low-resolution data to construct higher-resolution data in a more computationally efficient manner compared to dynamical downscaling. Recent years have witnessed significant progress in this field by employing deep learning techniques, specifically those derived from the single-image super-resolution (SR) problem. Various studies have demonstrated that deep learning models based on Convolutional Neural Networks (CNNs) or Generative Adversarial Networks (GANs) can generate realistic high-resolution deterministic or stochastic precipitation forecasting from single or multiple low-resolution meteorological fields.

Dong et al. (2015) proposed the SRCNN model, which applies deep CNN to the single-image SR reconstruction. They used three convolutional layers to extract image features and reconstruct a high-resolution image with the resolution increased up to a factor of 4 times. Vandal et al. (2017) used a stacked SRCNN framework, called the DeepSD, for downscaling of climate variables. Precipitation and elevation are used as input channels to generate high-resolution daily precipitation. The DeepSD was applied to GCMs' climate change scenarios projections in downscaling daily precipitation from 100 to 12.5 km (in a scaling factor of 8 times) over the Continental United States. Kumar et al. (2021) applied the DeepSD to the Indian region. Their results showed that the DeepSD could obtain reliable 0.25-degree rainfall data from 1-degree rainfall data derived from India Meteorological Department (IMD) and the Tropical Rainfall Measuring Mission (TRMM). The DeepSD method is also used to generate 0.125-degree rainfall from 0.25-degree ERA5 reanalysis rainfall over the Indian region. Some recent works considered both downscaling and nowcasting, in which the models contain CNN and Recurrent Neural Networks (RNN) parts. Shi et al. (2015) presented a

Convolutional Long Short-Term Memory (ConvLSTM), using convolutions to extract large-scale weather features and a recurrent structure for precipitation nowcasting. Adewoyin et al. (2021) presented a Temporal Recurrent U-Net (TRU-NET) model, which uses convolutional-recurrent layers to model multi-scale spatiotemporal weather processes. Their model predicts 8.5-km-resolution precipitation over the United Kingdom (UK) from several 65-km-resolution simulated weather variables (approximately an increase of 8 times resolution).

Another approach for image SR involves the utilization of GANs. Ledig et al. (2017) present SRGAN to get realistic images with 4× higher resolution. Kumar et al. (2023) showed that SRGAN outperforms DeepSD and ConvLSTM in downscaling 0.25-degree IMD rainfall to 0.0625-degree rainfall (4 times higher) over the India region. In the probabilistic forecasting sense, GANs are used as stochastically downscaling techniques to generate an ensemble of high-resolution precipitations from coarsen-resolution meteorological variables, which present the small-scale samples from the same large-scale distribution. Price and Rasp (2022) used a deep generative model to correct and downscale global ensemble forecasts over the Continental United States. Multiple meteorological variables from the European Centre for Medium-Range Weather Forecasts (ECMWF) ensemble forecasts system at 32 km resolution are used as input data to produce stochastic 4-km-resolution precipitation. Harris et al. (2022) used a GAN model to learn the mapping from low-resolution model outputs to high-resolution radar data. They produced realistic precipitation forecast data than the input data from approximately 10 km ECMWF model fields to 1 km precipitation over the UK (10X resolution higher). Their model produced an ensemble of predictions with reasonable statistical properties to estimate the uncertainty of their forecasts.

This paper uses CNN-based multi-scale residual network (MSRN; Li et al. 2018) downscaling approach to generate high-resolution precipitation data over Taiwan. In the SR problem, the residual network assumes that low-frequency information in the input and output images is similar. The model training efficiency can be improved by learning only the high-frequency information residuals between the two images (He et al. 2016; Kim et al. 2016). Li et al. (2018) used convolutional kernels of different sizes to design multi-scale residual blocks that extract multi-scale features. MSRN has a simple training structure and can be applied to any resolution ratio. The model performs well, improving the resolution by 2–8 times while using the multi-scale features of low-resolution images to reconstruct high-resolution images. In this work, we aim to downscale precipitation data from a 0.25-degree resolution to a

0.0125-degree resolution, thereby producing rainfall data with 20 times higher resolution. Faced with the challenge posed by the large scaling factor of 20, we utilize the advantage of multi-scale residual blocks to design a hierarchical architecture for the MSRN. We compare the effectiveness of this hierarchical approach with a direct downscaling MSRN approach in reconstructing high-resolution images under this large scaling factor condition. The model is trained using high-resolution radar retrieval precipitation, and we apply it to GCM’s low-resolution rainfall forecasts over Taiwan. The evaluation matrices associated with deterministic forecast are used for exploring the performance of MSRN on downscaling rainfall in Taiwan.

Methodology

Multi-scale residual network model

Figure 1 depicts the hierarchical architecture of MSRN in this study. The input data are the low-resolution rainfall data at 0.25-degree resolution, with the objective of improving the resolution by 20 times to output high-resolution rainfall data at 0.0125-degree resolution in Taiwan. This study designs a hierarchical connection of MSRN models, two MSRN models are used to enhance the resolution by 4 and 5 times, respectively, achieving a 20-time increase in resolution. First, using the 0.25-degree precipitation data as input for the 4X-MSRN and obtaining the output at 0.0625-degree resolution. The 5X-MSRN is then set to take the 0.0625-degree precipitation data as input and produce the output precipitation data at 0.0125-degree resolution.

Following Li et al. (2018), each MSRN model has eight multi-scale residual blocks (MSRBs), each of which is designed with 3×3 and 5×5 convolutional kernels and

two convolutional layers to extract different sizes of feature maps. Afterward, the two scales of feature maps are concatenated via one convolutional layer (Fig. 3, Li et al. 2018, p.522). This design allows multiple MSRBs to be connected and trained individually for input and output residuals, enhancing the module’s training efficiency while extracting feature maps of different sizes. MSRN also incorporates a bottleneck layer with a 1×1 convolutional kernel (gray box in Fig. 1) that merges the first convolutional layer and the features extracted by the eight MSRBs before entering the high-resolution image reconstruction process. In the high-resolution image reconstruction process, a pixel shuffle layer (green box in Fig. 1) is designed to perform sub-pixel convolutional layer upsampling on the low-resolution feature map and obtain the high-resolution image (Shi et al. 2016).

Data preprocessing and training

The high-resolution gridded precipitation data from Taiwan Central Weather Administration’s Quantitative Precipitation Estimation and Segregation Using Multiple Sensors system (QPESUMS, Chang et al. 2021) is used as high-resolution ground truth in this study. QPESUMS is recorded every 10 min with a spatial resolution of 0.0125 degrees. This work first aggregates the QPESUMS 10-min rainfall record to obtain the daily precipitation from 2006 to 2021 which is used for training and evaluation. The training and testing data are chosen randomly with a ratio of 8:2, which is 3447 days and 861 days, respectively. Each of the data covers the region of 21.825–25.3125°N, 120.0–122.2375°E with a dimension of 280×180.

Regarding the training process (Fig. 2), low-resolution (0.25-degree) and mid-resolution (0.0625-degree) precipitation data need to be prepared for training the

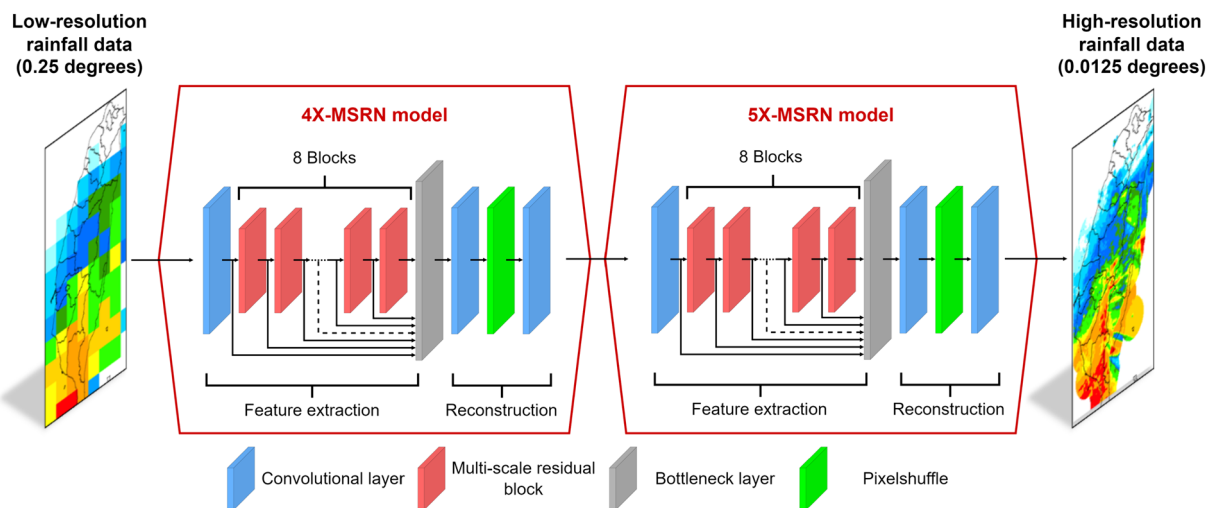


Fig. 1 The hierarchical architecture of multi-scale residual network (MSRN)

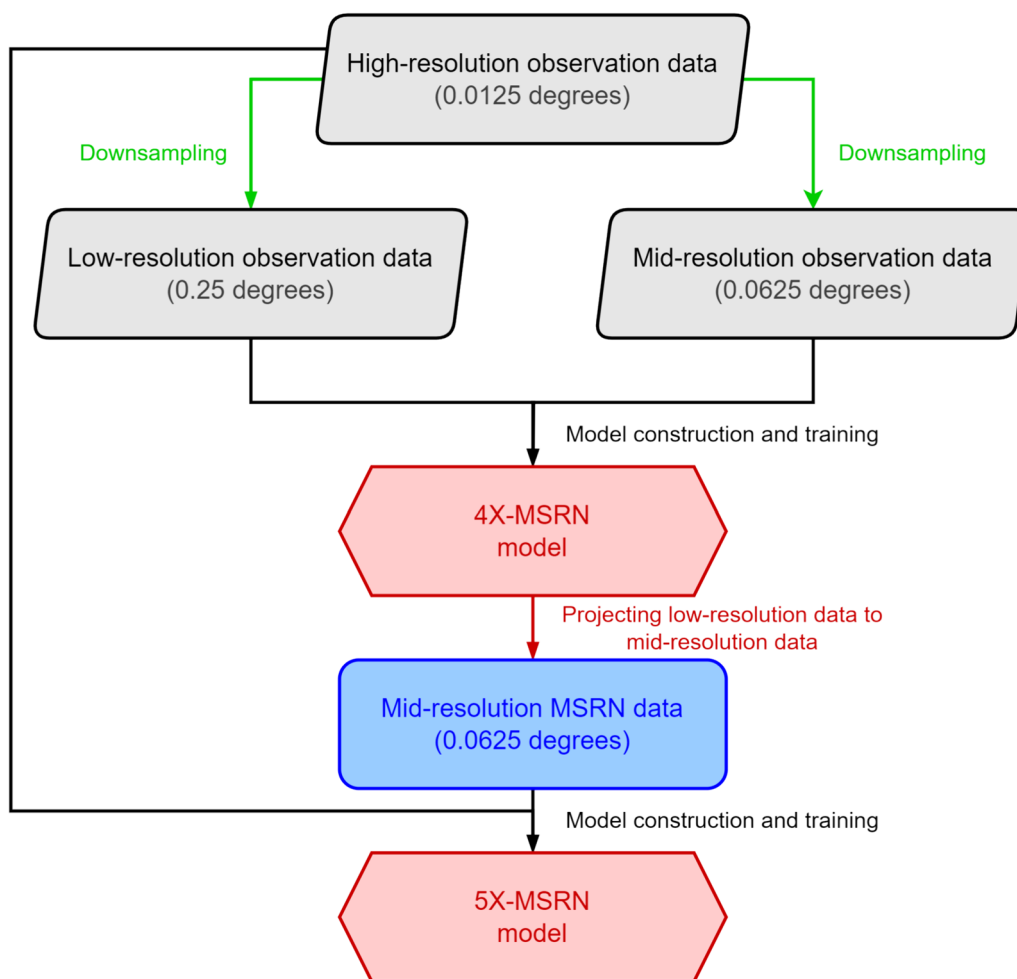


Fig. 2 Workflow of training the hierarchical MSRN module

4X-MSRN model. First, we use a 3×3 mean filter to reduce the magnitude of the high-frequency components of the high-resolution (0.0125-degree) precipitation data. Then the 0.0125-degree data are downsampled via storing the 20×20 -grid-averaged rainfall (5×5 -grid-averaged rainfall) to the corresponding cell of a 14×9 array (56×36 array) to construct the low-resolution dataset (mid-resolution dataset). For training the 4X-MSRN model, the 0.25-degree data are used as the input data, and the 0.0625-degree data are used as the learning target. The patch sizes are set to 4×4 and 16×16 and strides are set to 3 and 12 on 0.25° to 0.0625° scale. After the training of the 4X-MSRN model is completed, a set of 0.0625-degree data can be produced. For training the 5X-MSRN model, the 0.0625-degree data produced by the 4X-MSRN model is used as the input, with the observed 0.0125-degree data as the learning target. The patch sizes are set to 10×10 and 50×50 and strides are set to 5 and 25 on 0.0625 – 0.0125° scale. The mean

absolute error (MAE) is used as the loss function for both MSRN models.

Evaluation indices

In this study, we focus on the MSRN performance over Taiwan. Therefore, the reconstructed high-resolution results are compared with the 0.0125-degree QPESUMS daily precipitation data only on the 20485 land grids in our domain. We use mean absolute error (MAE) and root mean square error (RMSE) to evaluate the forecast errors, and compared the occurrence frequency of rainfall intensity between downscaled rainfall and their observation pairs. For evaluating the performance of reconstructed high-resolution rainfall that exceeds different rainfall thresholds, correspond contingency tables are constructed, which shows the frequencies of true positive (TP), false positive (FP), false negative (FN), and true negative (TN) for all land grids and all events. Five evaluation indices, namely

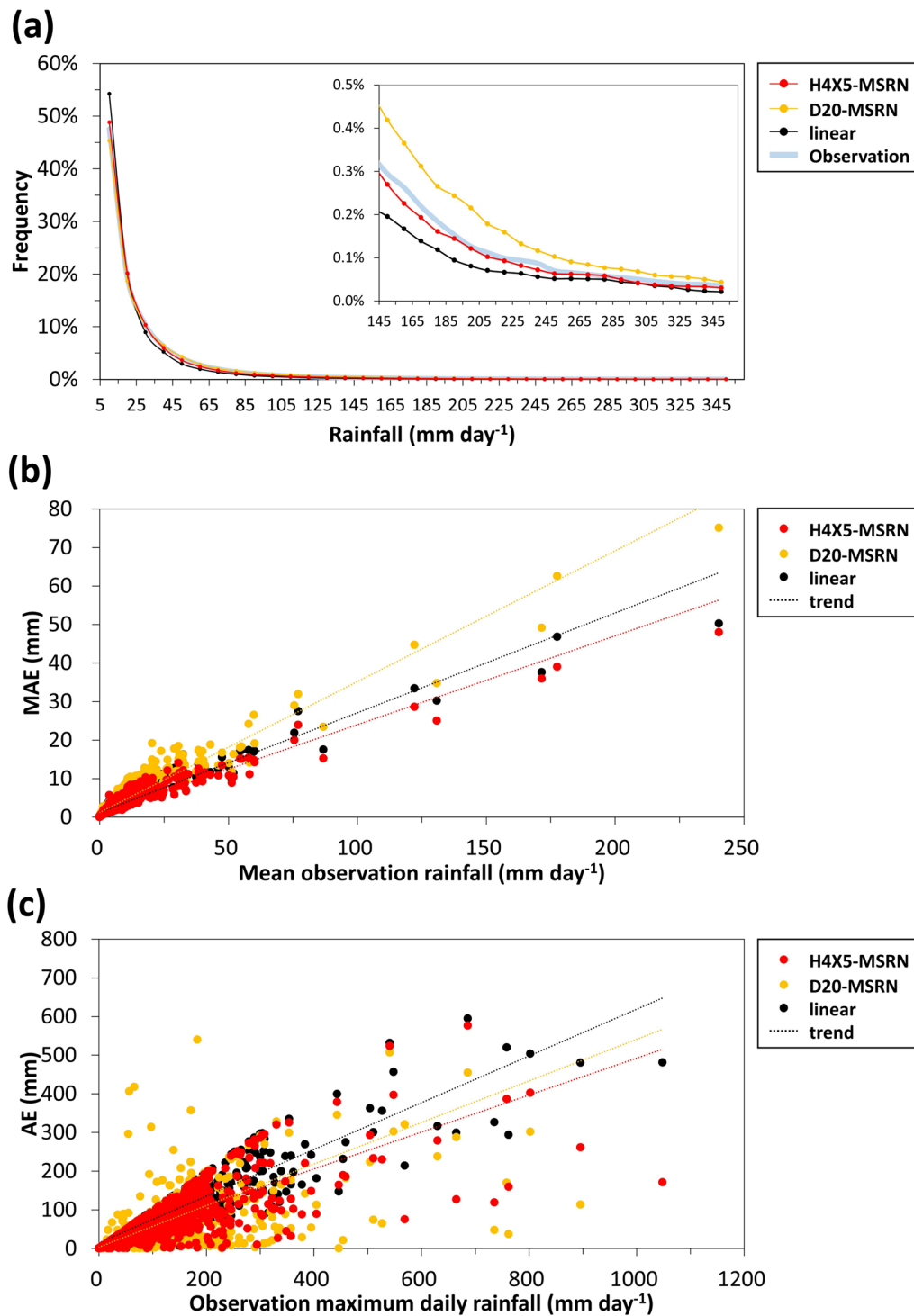


Fig. 3 **a** The relative frequency for which rainfall $\geq 5 \text{ mm day}^{-1}$. Each bin width is 10 mm. Right panel enlarge the scale of frequency within the rainfall range of 150–350 mm day^{-1} . The total counts for observation, H4X5-MSRN, D20-MSRN, and linear are 3875473, 4026543, 2853880, and 4486465, respectively. **b** The mean absolute errors (MAE) against with the mean observational daily rainfall. **c** The absolute errors of maximum daily rainfall. Red, yellow, and black lines and dots represent the results from H4X5-MSRN, D20-MSRN, and linear interpolation. **a** Also denotes the observational frequency with a light blue line. The dotted lines in **b** and **c** are the linear trends

true positive rate (*TPR*), F_1 score (F_1), bias score (*BS*), false positive rate (*FPR*), threat score (*TS*) can be derived (Jolliffe and Stephenson 2012):

$$TPR = \frac{TP}{TP + FN} \quad (1)$$

$$F_1 = \frac{2TP}{2TP + FP + FN} \quad (2)$$

$$BS = \frac{TP + FP}{TP + FN} \quad (3)$$

$$FPR = \frac{FP}{FP + TN} \quad (4)$$

$$TS = \frac{TP}{TP + FN + FP} \quad (5)$$

TPR ranges from 0 to 1 and represents the proportion of correctly predicted (hits) events to all occurred observations, also called sensitivity. F_1 ranges from 0 to 1 and represents the harmonic mean of precision and sensitivity thus measures the accuracy. *BS* ranges from 0 to ∞ and represents the relative frequency of positive forecasts compare to the occurred observations. The perfect score is 1. The forecast tends to underestimate (overestimate) the occurred frequencies while the *BS* is less (greater) than 1. *FPR* ranges from 0 to 1 and represents the proportion of false alarm events to all negative observed events. *TS* ranges from 0 to 1 and quantifies the similarity between occurred observations and positive predicted events by the ratio of the intersection size to the union size of the two samples.

In addition to above grid-to-grid evaluation methods, the Fractions Skill Score (*FSS*) (Roberts and Lean 2008) is one of the popular skill scores that evaluates the high-resolution rainfall predictions at a given spatial scale. For a given rainfall threshold and a given size of neighborhood, the forecast and observed fraction of grids exceed the threshold in the neighborhood of each grid are compared. The *FSS* of daily rainfall can be obtained by:

$$FSS = 1 - \frac{\frac{1}{N} \sum_N (P_f - P_o)^2}{\frac{1}{N} (\sum_N P_f^2 + \sum_N P_o^2)} \quad (6)$$

where P_f is the forecast fraction, P_o is the observed fraction, N is the number of evaluated grids. Considering the sizes of Taiwan counties and cities, a neighborhood of 4×4 grids (approximately 5×5 km²) is empirically chosen to calculate the fractions in this work.

Results

Evaluations of testing data

We compare the hierarchical MSRN approach (H4X5-MSRN) with the direct downscaling approach of MSRN (D20-MSRN) and the linear interpolation method (linear). Figure 3a shows the relative frequency for which rainfall is equal to and greater than 5 mm day⁻¹. The results indicate that the H4X5-MSRN exhibits the closest rainfall frequency distribution to the observation data. The D20-MSRN has a tendency to overestimate the frequency of rainfall amounts ≥ 30 mm day⁻¹. This overestimation is likely due to a significant decrease in the total number of grid points with rainfall > 5 mm day⁻¹ along with an increase in grid points with rainfall < 5 mm day⁻¹. The linear interpolation method underestimates the frequency of rainfall ≥ 20 mm day⁻¹ while overestimating the frequency of rainfall < 20 mm day⁻¹. Figure 3b displays the daily MAE against the mean observed rainfall. H4X5-MSRN tends to have smaller errors among the three methods, while D20-MSRN tends to produce larger errors. Overall, the total averaged MAE (RMSE) is 2.595 mm day⁻¹ (5.542 mm day⁻¹) for H4X5-MSRN, which represents an improvement of 15.1% in MAE (9.1% in RMSE) compared to the linear interpolation method (Table 1). The improvements of the three methods regarding the maximum daily rainfall are examined. Figure 3c reveals the absolute errors of the three methods against the observational maximum daily rainfall. H4X5-MSRN generally exhibits smaller errors in the maximum rainfall values. Table 2 further calculates the MAE of the maximum daily rainfall based on three rainfall ranges: < 100 mm, 100–300 mm, and ≥ 300 mm. D20-MSRN has the smallest MAE of 94.86 mm day⁻¹ in the range of 100–300 mm, while H4X5-MSRN has the smaller MAEs of 31.25 mm day⁻¹ and 199.42 mm day⁻¹ in the range of < 100 mm and ≥ 300 mm, respectively. Notably, the calculation of MAE for maximum daily rainfall does not consider the correct locations of maximum rainfall, meaning that the maximum rainfall may occur in

Table 1 Mean absolute errors (MAE) and root mean square errors (RMSE) for H4X5-MSRN, D20-MSRN and linear interpolation method

	Linear	D20-MSRN	H4X5-MSRN
MAE (mm day ⁻¹)	3.057	3.617	2.595
MAE improve ratio (%)		- 18.3	15.1
RMSE (mm day ⁻¹)	6.095	7.848	5.542
RMSE improve ratio (%)		- 28.8	9.1

The second row and fourth row show the improve ratios of MAE and RMSE relative to the linear interpolation

The best score for each metric is highlighted in bold

Table 2 Mean absolute errors (MAE) of maximum daily rainfall for H4X5-MSRN, D20-MSRN and linear interpolation

Maximum daily rainfall range (mm day ⁻¹)	Linear (MAE/improve ratio %)	D20-MSRN (MAE/improve ratio %)	H4X5-MSRN (MAE/improve ratio %)
< 100	39.630	33.648/15.1	31.253/21.1
100–300	123.181	94.863/23.0	106.048/13.9
≥ 300	274.196	209.502/23.6	199.418/27.3

The MAE and improve ratio relative to the linear interpolation are calculated in three criteria: maximum daily rainfall < 100 mm, 100 mm ≤ maximum daily rainfall < 100 mm, and maximum daily rainfall ≥ 300 mm

The best score for each metric is highlighted in bold

the wrong places. Next, we use the grid-to-grid evaluation indices to examine our results.

Figure 4 examines grid-to-grid evaluation indices at rainfall thresholds of 1, 10, 40, 80, 100, 150, 200, 350, and 500 mm day⁻¹. In Fig. 4a, we depict the *TPR* and *F₁* scores. H4X5-MSRN consistently outperforms the other two methods for rainfall thresholds exceeding 40 mm day⁻¹. *TPR* and *F₁* score remain close to 0.8 until the rainfall threshold surpasses 200 mm day⁻¹. For thresholds of 350 and 500 mm day⁻¹, the *F₁* scores still exceed 0.7 and 0.6, respectively. Figure 4b analyzes the *BS* and indicates that H4X5-MSRN exhibits nearly no bias (*BS*=1), while the linear method (D20-MSRN) tends to underestimate (overestimate) at higher rainfall thresholds. The *FPR* analysis also shows that D20-MSRN

tends to produce more false alarms for higher rainfall thresholds (≥ 40 mm day⁻¹), while the linear method tends to produce more false alarms at thresholds of 1 and 10 mm day⁻¹. Overall, the *TS* indicates that H4X5-MSRN outperforms the other two methods (as shown in Fig. 4c), particularly at rainfall thresholds of 350 and 500 mm day⁻¹. At these thresholds, the *TS* scores for H4X5-MSRN surpass 0.54 and 0.46, respectively, signifying a significant enhancement in performance compared to the other two methods in cases of extreme rainfall. D20-MSRN consistently scores the lowest *TS* in almost all thresholds due to its higher *FPR*, except for the 500 mm day⁻¹ threshold, where the linear method performs worst due to its lowest *TPR*. Figure 4d reveals the results of daily averaged *FSS* for the three methods.

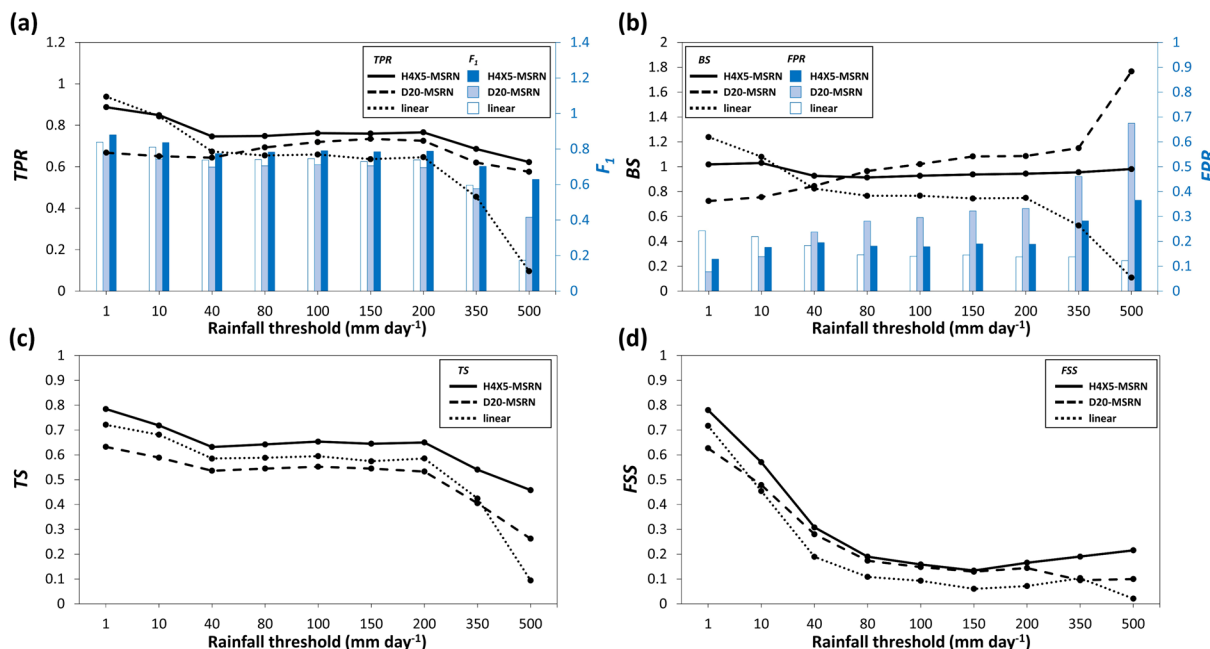


Fig. 4 Evaluation scores at rainfall thresholds of 1, 10, 40, 80, 100, 150, 200, 350, and 500 mm day⁻¹. **a** *TPR* (lines) and *F₁* scores (bars). **b** *BS* (lines) and *FPR* (bars). **c** *TS*. **d** Daily averaged *FSS*. Solid lines and dark blue bars represent H4X5-MSRN. Dashed lines and light blue bars represent D20-MSRN. Dotted lines and white bars show the results of linear interpolation

FSS allows the forecast to exceed the threshold within a given neighborhood size (approximately $5 \times 5 \text{ km}^2$ in this study). Therefore, the scores of D20-MSRN are not significantly different from those of H4X5-MSRN for thresholds between 40 and 200 mm day^{-1} . However, H4X5-MSRN consistently outperforms the others, even when we increase the neighborhood size to approximately $11 \times 11 \text{ km}^2$ (figure not shown), especially for rainfall thresholds higher than 350 mm day^{-1} .

Figure 5 presents examples of H4X5-MSRN and D20-MSRN, illustrating the various precipitation distributions associated with different meteorological conditions. The results depicted in Fig. 5 unmistakably emphasize H4X5-MSRN's capability to accurately recreate fine-grained rainfall distribution and intense rainfall areas. This results in a more realistic visual representation due to its sharply defined spatial structure. In contrast, while D20-MSRN can also generate intense rainfall, it occasionally falls short in precisely reconstructing the spatial distribution. For instance, on October 11, 2009, and May 19, 2014, D20-MSRN successfully reproduced the local rainfall maxima in northeastern and mid-western Taiwan, respectively. However, it encounters challenges in faithfully reproducing the rainfall distribution in southeastern and southwestern Taiwan. It is notable that the D20-MSRN method appears to artificially generate many light rain areas from the low-resolution data. This is due to its tendency to increase the number of grid points with rainfall $< 5 \text{ mm day}^{-1}$.

Evaluations of applying hierarchical MSRN to GCM precipitation data

This section evaluates the performance of the hierarchical MSRN approach (H4X5-MSRN) in improving the rainfall forecast for Taiwan produced by a global model. We utilize real-time, open data downloaded from the ECMWF website (www.ecmwf.int). The data are covered by the Creative Commons Attribution 4.0 International (CC BY 4.0) license (<https://creativecommons.org/licenses/by/4.0/>). This open data is a subset of the ECMWF real-time products based on a high-resolution single forecast model (HRES) and ensemble forecasts (Copyright © 2023 by the European Centre for Medium-Range Weather Forecasts (ECMWF)). The HRES model has an approximately 9 km horizontal resolution, while the open data are produced at a 0.4-degree resolution. To assess H4X5-MSRN's effectiveness, we collected ECMWF's daily rainfall forecasts initialized from 00 UTC for December 2022 to November 2023. The rainfall forecast data are first interpolated to 0.25-degree mesh covering Taiwan to match the input format of the H4X5-MSRN model. Figure 6 displays rainfall maps for some example cases using ECMWF's day-1 forecast data. These

examples indicate that the H4X5-MSRN model generates more detailed structures and a more realistic visual representation of rainfall distribution compared to the other methods. The H4X5-MSRN model is capable of improving the forecasted rainfall and clearly enhances the intense rainfall areas in comparison with the linearly interpolated rainfall and the low-resolution rainfall.

The *TPR*, *FAR*, F_1 score and *TS* at various rainfall thresholds for lead times of 1–3, 4–6, and 7–9 days are shown in Fig. 7. Figure 7a indicates that the H4X5-MSRN improves the *TPR* at rainfall thresholds greater than 10 mm day^{-1} overall. Specifically, the most significant improvement occurs at rainfall thresholds greater than 100 mm day^{-1} for the 1- to 3-day lead time. While the H4X5-MSRN improves the 1- to 3-day *TPR*, it only slightly increases the *FAR* (Fig. 7b). Note that both the *TPR* and *FAR* are zero for the 1- to 3-day linear experiment, indicating that the linear experiment cannot forecast 500 mm day^{-1} cases. For the 4- to 6- and 7- to 9-day leads, the linear method shows low *TPR* and high *FAR* at large rainfall thresholds, consistent with decreased model predictability at longer lead times. Thus, the H4X5-MSRN improvement is less significant compared to the 1- to 3-day lead results. Overall, the two aggregative indices, F_1 score and *TS*, show similar characteristics (Fig. 7c, d). The H4X5-MSRN is more effective for rainfall greater than 100 mm day^{-1} and a 1- to 3-day lead, also depending on the predictability of the GCM results. If the GCM forecast is completely inaccurate, the H4X5-MSRN cannot correct the rainfall forecast to a very different scenario.

Conclusions and discussions

In recent years, there has been a growing demand for rainfall data with a spatial resolution smaller than 5 km to effectively support disaster prevention efforts in various regions of Taiwan. It is important to note that users typically access GCM's rainfall data with a resolution ranging from 25 to 80 km. Therefore, the use of downscaling techniques has become crucial for generating higher-resolution rainfall data from coarser reanalysis or forecast rainfall datasets. We employed a CNN-based Multi-Scale Residual Network (MSRN) approach to downscale precipitation from a 0.25-degree resolution to a much finer 0.0125-degree resolution. To tackle the significant challenge of increasing resolution by over 20 times, we developed a hierarchical architecture for MSRN. The model was trained using high-resolution radar-derived precipitation data and subsequently applied to GCM's low-resolution rainfall forecasts for Taiwan.

The hierarchical MSRN (H4X5-MSRN) outperforms both the one-step MSRN (D20-MSRN) and linear interpolation methods. The MAE and RMSE of daily rainfall

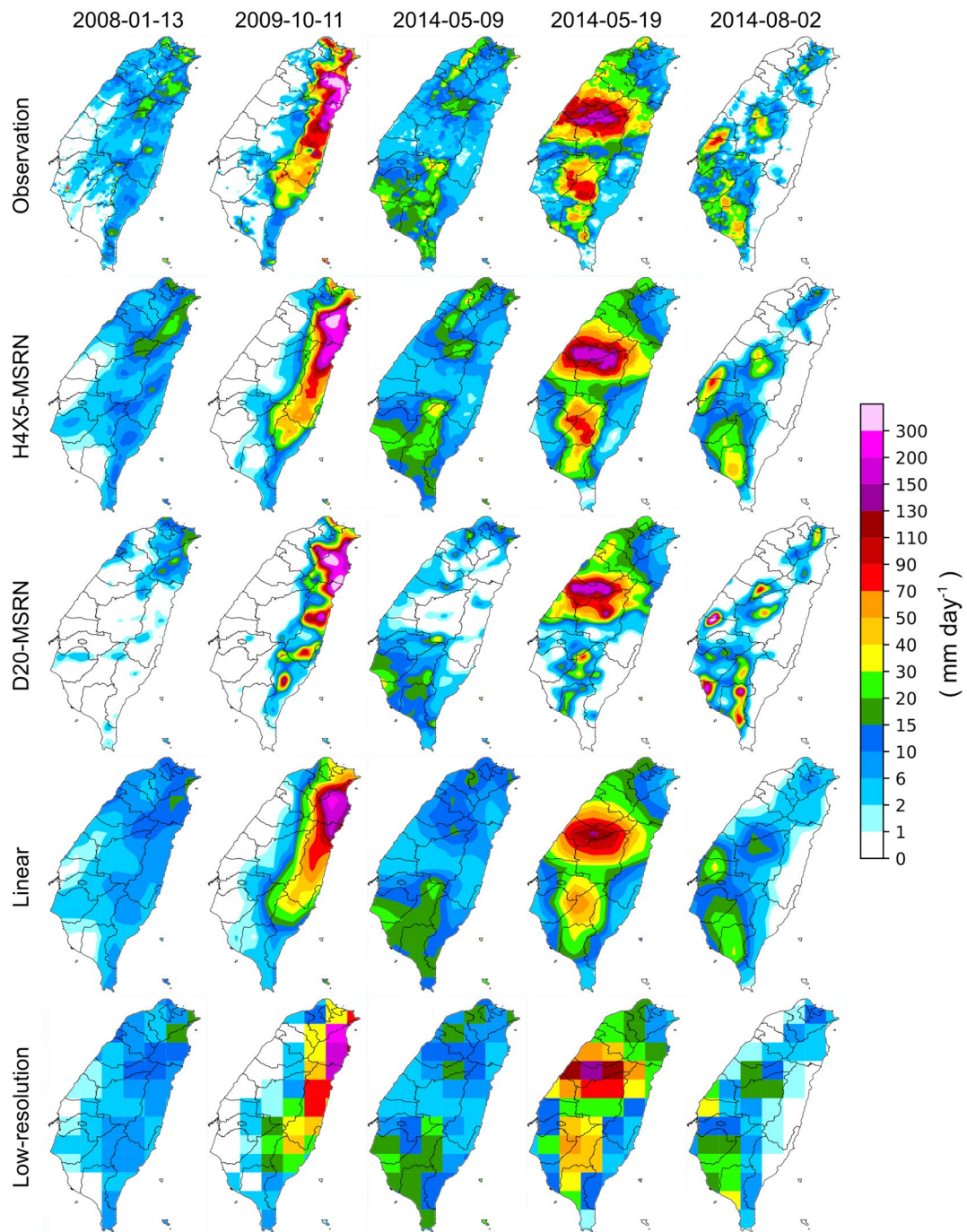


Fig. 5 Comparison of reconstructed high-resolution rainfall maps with low-resolution (0.25-degree) rainfall maps for five selected cases. Rainfall maps from top to bottom are: the high-resolution observational rainfall data, the H4X5-MSRN reconstructed rainfall data, the D20-MSRN reconstructed rainfall data, the linear interpolated rainfall data, and the low-resolution rainfall data

for the H4X5-MSRN are 2.60 and 5.54 mm, respectively. This represents a 15.1% improvement in MAE and a 9.1% improvement in RMSE compared to the linear interpolation method. We also examined the evaluation metrics based on the error matrix for various rainfall thresholds.

The H4X5-MSRN consistently outperforms the other two methods in accurately reconstructing high-resolution rainfall, showing nearly no biases for all rainfall thresholds. The *TS* indicates that H4X5-MSRN excels in reproducing extreme rainfall events. *TS* scores for

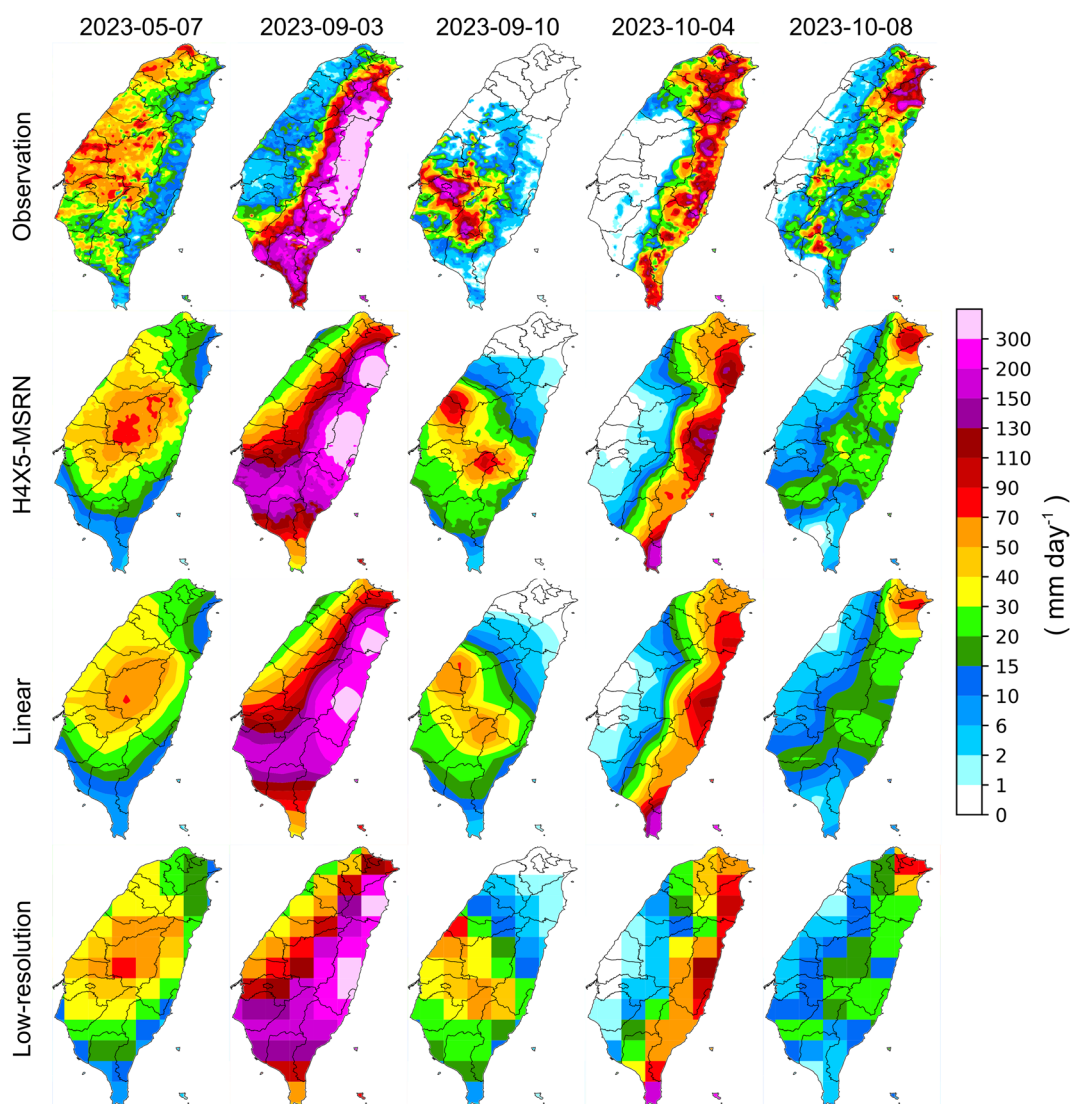


Fig. 6 Comparison of reconstructed high-resolution rainfall maps with low-resolution (0.25-degree) rainfall maps using ECMWF's day-1 forecast data for five selected cases. The rainfall maps from top to bottom are: the high-resolution observational rainfall data, the H4X5-MSRN reconstructed rainfall data, the linear interpolated rainfall data, and the low-resolution rainfall data

H4X5-MSRN exceed 0.54 and 0.46 at rainfall thresholds of 350 and 500 mm per day, respectively, significantly outperforming the other two methods (0.42/0.41 and 0.09/0.26 for linear/D20-MSRN experiments). Furthermore, H4X5-MSRN accurately reproduces fine-grained rainfall distribution, resulting in a more realistic visual representation compared to the other two methods. The ECMWF daily rainfall forecasts over Taiwan from December 2022 to November 2023 are used to examine the effectiveness of the H4X5-MSRN. The evaluation results indicate that the H4X5-MSRN is effective at improving rainfall forecasts for thresholds greater than 100 mm per day. However, the performance is also

affected by the predictability of the original GCM results, with more significant improvement for 1- to 3-day lead forecast compared to longer lead times.

Many previous studies have also considered bias correction during downscaling. In our work, we trained the model using high-resolution radar-derived precipitation data, indicating that the learning processes did not correct the biases specific to numerical models. Model biases may arise not only from differences in horizontal resolutions of data and topography, but also from assumptions made in various physic schemes and energy conservation during integration. We regard high-resolution radar-derived precipitation data as our learning target, with

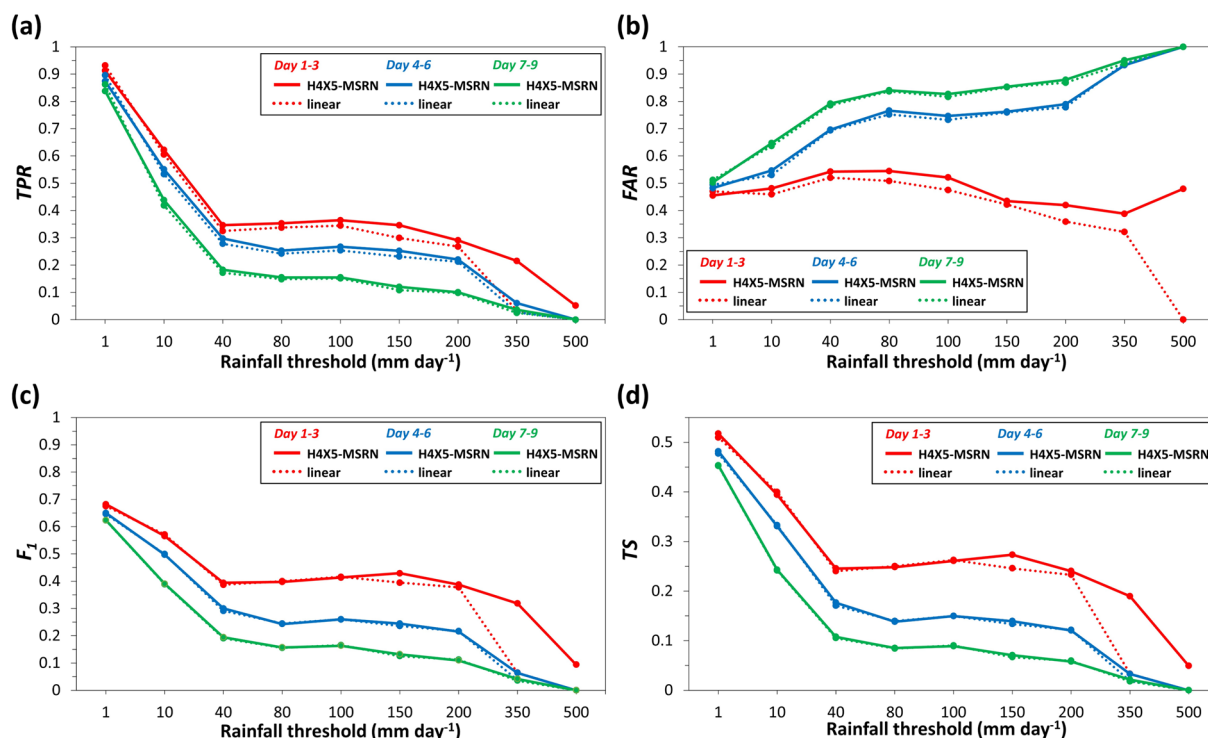


Fig. 7 **a** TPR, **b** F1 scores and **c** TS at rainfall thresholds of 1, 10, 40, 80, 100, 150, 200, 350, and 500 mm day⁻¹ for applying H4X5-MSRN (solid lines) to ECMWF HRES daily rainfall forecast over Taiwan. Dotted lines show the results of linear interpolation. Red, blue, and green represent the results of 1- to 3-day lead, 4- to 6-day lead, and 7- to 9-day lead, respectively

the expectation that the model can learn the processes related to varying horizontal data resolutions, including the inherent relationships between high-resolution observational data and topography. Harris et al. (2022) addressed this problem by separating it into two parts: pure super-resolution and bias/spread correction. Their model aims to both increase the resolution of the original forecast and provide error correction in a probabilistic sense. In practice, users typically access rainfall data from a modern ensemble forecast system, which may include results generated by various dynamic models. Correcting biases for each individual model can be challenging for users. Our method provides a universal approach for efficiently downscaling ensemble daily rainfall data. The availability of high-resolution rainfall data can further enhance its application in downstream hydrological modeling for disaster impact assessment. One current limitation is that our method, designed for daily data, does not generate hourly rainfall outputs. Thus, it may be unsuitable for applications requiring higher temporal rainfall data resolution. Moving forward, it is practical to integrate coarse- or fine-scale bias correction mechanisms into our downscaling process for future research.

Abbreviations

BS Bias score

CNN	Convolutional neural network
ConvLSTM	Convolutional long short-term memory
ECMWF	European centre for medium-range weather forecasts
F ₁	F ₁ score
FN	False negative
FP	False positive
FPR	False positive rate
FSS	Fractions skill score
GAN	Generative adversarial network
GCM	General circulation model
HRES	High-resolution single forecast model
IMD	India meteorological department
MAE	Mean absolute error
MSRN	Multi-scale residual networks
MSRB	Multi-scale residual block
QPESUMS	Quantitative precipitation estimation and segregation using multiple sensors
RCM	Regional climate model
RMSE	Root mean square error
RNN	Recurrent Neural Network
SR	Super-resolution
TN	True negative
TP	True positive
TPR	True positive rate
TRMM	Tropical rainfall measuring mission
TRU-NET	Temporal recurrent U-Net
TS	Threat score

Acknowledgements

This research was supported by the National Science and Technology Center for Disaster Reduction, Taiwan. The computational and storage resources were provided by the National Center for High-performance Computing (NCHC),

Taiwan. The authors would like to thank Central Weather Administration Taiwan for providing the QPESUMS precipitation data.

Author contributions

Conceptualization, L.-H. H., C.-C. C. and J.-L. C.; methodology, L.-H. H., C.-C. C. and C.-S. F.; software, C.-C. C. and C.-S. F.; validation, C.-C. C.; formal analysis, C.-C. C.; investigation, L.-H. H.; data curation, K.-L. L., H.-H. L.; writing—original draft preparation, L.-H. H.; writing—review and editing, L.-H. H. and J.-L. C.; visualization, C.-C. C.; supervision, L.-H. H. and J.-L. C.; project administration, Y.-C. Y.; funding, acquisition, Y.-C. Y. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Availability of data and materials

The rainfall data are provided by the Taiwan Central Weather Administration's Quantitative Precipitation Estimation and Segregation Using Multiple Sensors system (QPESUMS). The European Centre for Medium-Range Weather Forecasts (ECMWF) open data are available from their website: <https://www.ecmwf.int> (accessed on Nov. 2023). This data set is covered by the Creative Commons Attribution 4.0 International (CC BY 4.0) license (<https://creativecommons.org/licenses/by/4.0/>). (Copyright © 2023 by the European Centre for Medium-Range Weather Forecasts (ECMWF)).

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 9 January 2024 Accepted: 15 April 2024

Published online: 09 May 2024

References

- Adewoyin RA, Dueben P, Watson P, He Y, Dutta R (2021) TRU-NET: a deep learning approach to high resolution prediction of rainfall. *Mach Learn* 110:2035–2062. <https://doi.org/10.1007/s10994-021-06022-6>
- Bauer P, Thorpe A, Brunet G (2015) The quiet revolution of numerical weather prediction. *Nature* 525(7567):47–55. <https://doi.org/10.1038/nature14956>
- Chang PL, Zhang J, Tang YS, Tang L, Lin PF, Langston C, Kaney B, Chen CR, Howard K (2021) An operational multi-radar multi-sensor QPE system in Taiwan. *B Am Meteorol Soc* 102(3):E555–E577. <https://doi.org/10.1175/BAMS-D-20-0043.1>
- D'Onofrio D, Palazzi E, von Hardenberg J, Provenzale A, Calmanti S (2014) Stochastic rainfall downscaling of climate models. *J Hydrometeorol* 15(2):830–843. <https://doi.org/10.1175/JHM-D-13-096.1>
- Dong C, Loy CC, He K, Tang X (2015) Image super-resolution using deep convolutional networks. *IEEE T Pattern Anal* 38(2):295–307. <https://doi.org/10.1109/TPAMI.2015.2439281>
- Giorgi F, Gutowski WJ Jr (2015) Regional dynamical downscaling and the CORDEX initiative. *Annu Rev Environ Resour* 40:467–490. <https://doi.org/10.1146/annurev-environ-102014-021217>
- Harris L, McRae AT, Chantry M, Dueben PD, Palmer TN (2022) A generative deep learning approach to stochastic downscaling of precipitation forecasts. *J Adv Model Earth Sy*. <https://doi.org/10.1029/2022MS003120>
- He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 770–778. <https://doi.org/10.1109/CVPR.2016.90>
- Hong SY, Kanamitsu M (2014) Dynamical downscaling: fundamental issues from an NWP point of view and recommendations. *Asia-Pac J Atmos Sci* 50:83–104. <https://doi.org/10.1007/s13143-014-0029-2>
- Jolliffe IT, Stephenson DB (eds) (2012) Forecast verification: a practitioner's guide in atmospheric science, 2nd edn. John Wiley & Sons, Blackwell, Oxford
- Kaur M, Krishna RPM, Joseph S, Dey A, Mandal R, Chattopadhyay R, Sahai AK, Mukhopadhyay P, Abhilash S (2020) Dynamical downscaling of a multimodel ensemble prediction system: application to tropical cyclones. *Atmos Sci Lett* 21(8):e971. <https://doi.org/10.1002/asl.971>
- Kim J, Lee JK, Lee KM (2016) Accurate image super-resolution using very deep convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp 1646–1654. <https://doi.org/10.1109/CVPR.2016.182>
- Kumar B, Chattopadhyay R, Singh M, Chaudhari N, Kodari K, Barve A (2021) Deep learning-based downscaling of summer monsoon rainfall data over Indian region. *Theor Appl Climatol* 143:1145–1156. <https://doi.org/10.1007/s00704-020-03489-6>
- Kumar B, Atey K, Singh BB, Chattopadhyay R, Acharya N, Singh M, Nanjundiah RS, Rao SA (2023) On the modern deep learning approaches for precipitation downscaling. *Earth Sci Inform* 3:1–14. <https://doi.org/10.1007/s12145-023-00970-4>
- Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z, Shi W (2017) Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp 4681–4690.
- Li LL, Li J, Yu RC (2022) Evaluation of CMIP6 HighResMIP models in simulating precipitation over Central Asia. *Adv Clim Chang Res* 13(1):1–13. <https://doi.org/10.1016/j.jaccre.2021.09.009>
- Li J, Fang F, Mei K, Zhang G (2018) Multi-scale residual network for image super-resolution. In: Proceedings of the European conference on computer vision, pp 517–532. https://doi.org/10.1007/978-3-030-01237-3_32
- Lucas-Picher P, Argüeso D, Brisson E, Trambly Y, Berg P, Lemonsu A, Kotlarski S, Caillaud C (2021) Convection-permitting modeling with regional climate models: latest developments and next steps. *Wiley Interdiscip Rev: Clim Change* 12(6):e731. <https://doi.org/10.1002/wcc.731>
- Maraun D, Wetterhall F, Ireson AM, Chandler RE, Kendon EJ, Widmann M, Brienen S, Rust HW, Sauter T, Themeßl M, Venema VKC, Chun KP, Goodess CM, Jones RG, Onof C, Vrac M, Thiele-Eich I (2010) Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. *Rev Geophys*. <https://doi.org/10.1029/2009RG000314>
- Masson-Delmotte V, Zhai P, Pirani A, Connors SL, Péan C, Berger S, Caud N, Chen Y, Goldfarb L, Gomis MI, Huang M (2021) Climate change 2021: the physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change
- May W (2004) Simulation of the variability and extremes of daily rainfall during the Indian summer monsoon for present and future times in a global time-slice experiment. *Clim Dynam* 22(2):183–204. <https://doi.org/10.1007/s00382-003-0373-x>
- Prein AF, Langhans W, Fosser G, Ferrone A, Ban N, Goergen K, Keller M, Tölle M, Gutjahr O, Feser F, Brisson E (2015) A review on regional convection-permitting climate modeling: demonstrations, prospects, and challenges. *Rev Geophys* 53(2):323–361. <https://doi.org/10.1002/2014RG000475>
- Price I, Rasp S (2022) Increasing the accuracy and resolution of precipitation forecasts using deep generative models. In: Proceedings of the 25th International Conference on Artificial Intelligence and Statistics. PMLR 151:10555–10571
- Roberts NM, Lean W (2008) Scale-selective verification rainfall accumulations from high-resolution forecasts of convective events. *Mon Weather Rev* 136:78–97. <https://doi.org/10.1175/2007MWR2123.1>
- Schiermeier Q (2010) The real holes in climate science. *Nature* 463:284–287. <https://doi.org/10.1038/463284a>
- Shi X, Chen Z, Wang H, Yeung DY, Wong WK, Woo WC (2015) Convolutional LSTM network: a machine learning approach for precipitation nowcasting. *Adv Neur in* 28:802–810
- Shi W, Caballero J, Huszár F, Totz J, Aitken AP, Bishop R, Rueckert D, Wang Z (2016) Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 1874–1883. <https://doi.org/10.1109/CVPR.2016.207>
- Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. *B Am Meteorol Soc* 93(4):485–498. <https://doi.org/10.1175/BAMS-D-11-00094.1>
- Vandal T, Kodra E, Ganguly S, Michaelis A, Nemani R, Ganguly AR (2017) DeepSD: Generating high resolution climate change projections through single image super-resolution. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp 1663–1672. <https://doi.org/10.1145/3097983.3098004>

- von Storch H, Langenberg H, Feser F (2000) A spectral nudging technique for dynamical downscaling purposes. *Mon Weather Rev* 128(10):3664–3673. [https://doi.org/10.1175/1520-0493\(2000\)128%3c3664:ASNTFD%3e2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128%3c3664:ASNTFD%3e2.0.CO;2)
- Xue Y, Janjic Z, Dudhia J, Vasic R, De Sales F (2014) A review on regional dynamical downscaling in intraseasonal to seasonal simulation/prediction and major factors that affect downscaling ability. *Atmos Res* 147:68–85. <https://doi.org/10.1016/j.atmosres.2014.05.001>
- Yu W, Nakakita E, Kim S, Yamaguchi K (2016) Impact assessment of uncertainty propagation of ensemble NWP rainfall to flood forecasting with catchment scale. *Adv Meteorol* 2016:1384302. <https://doi.org/10.1155/2016/1384302>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.