https://doi.org/10.5194/egusphere-2026-2395
Preprint. Discussion started: 7 July 2026

CBAM-U-Net-Based Retrieval of Radar Composite Reflectivity from FY-4A Satellite Observations over Complex Terrain in Sichuan, China

Wen Kang[1], Hao Wang[1,2], Qiangyu Zeng[1], Tiantian Yu[1], Jiafeng Zheng[1], and Zhi Li[3]

[1]College of Atmospheric Sounding, Chengdu University of Information Technology, Chengdu, 610225, China
[2]Wenjiang National Climatology Observatory, Sichuan Provincial Meteorological Service, Chengdu, 610072, China
[3]Institute of Tibetan Plateau Meteorology, Chinese Academy of Meteorological Sciences, Chengdu, 610081, China

Correspondence: Hao Wang (wh@cuit.edu.cn)

Abstract. To address radar coverage blind spots in complex terrain, this study proposes an end-to-end deep learning framework to retrieve Radar Composite Reflectivity (RCRF) from FY-4A satellite multi-channel observations. We introduce CBAM-U-Net, embedding a lightweight Convolutional Block Attention Module into a U-Net backbone. This dual-dimensional mechanism adaptively filters critical infrared spectral bands and precisely localizes intense convective cores. Evaluated on a comprehensively matched satellite-radar dataset (14,023 samples) from Sichuan Province (May–November 2023), CBAM-U-Net significantly outperforms mainstream CNN and Transformer baselines in retrieval accuracy (RMSE = 6.8290 dBZ, R^2 = 0.6277) and structural fidelity (SSIM = 0.7894). Crucially, within the challenging severe echo regime (45–70 dBZ), the model achieves optimal Probability of Detection (POD = 0.5296) and Critical Success Index (CSI = 0.4384). Furthermore, cross-sensor evaluations using FY-4B data demonstrate its robust zero-shot generalization against observational domain shifts. This research highlights the efficacy of integrating satellite multispectral features with attention-augmented networks to compensate for radar blind spots, providing reliable support for severe convective weather monitoring.

1 Introduction

China, situated in the East Asian monsoon region, frequently experiences extreme precipitation and severe convective weather, which often trigger secondary disasters such as flash floods, landslides, and urban waterlogging, posing significant threats to socioeconomic stability (Yan et al., 2022). Consequently, establishing a meteorological monitoring system characterized by high spatiotemporal resolution and robust reliability is critical for disaster prevention. Currently, while radiosondes provide precise atmospheric vertical profiling, their sparse spatial distribution and low temporal frequency (typically twice daily) severely hinder the real-time capture of rapidly evolving mesoscale convective systems (Wang et al., 2020b). Therefore, operational nowcasting predominantly relies on weather radars. By providing high-resolution backscattering information of precipitation particles, radar networks serve as the fundamental data source for severe convection monitoring. However, radar networks suffer from intrinsic spatiotemporal limitations. In topographically complex regions—such as the western Sichuan Plateau, which features dramatic elevation fluctuations exceeding 3,000 meters—radar beams are severely blocked on the leeward sides

https://doi.org/10.5194/egusphere-2026-2395
Preprint. Discussion started: 7 July 2026

of mountains (Yu et al., 2023b). This topographic blockage, coupled with the sparse distribution of stations and the recurrent "attenuation shadow" effect caused by heavy precipitation, inevitably creates extensive observation blind zones (Wang et al., 2013, 2020b; Yu et al., 2023b; Wang et al., 2026). Thus, achieving continuous monitoring of severe convective systems relying solely on ground-based radar networks in complex terrain remains exceedingly difficult. In contrast, geostationary meteorological satellites effectively overcome these geographical barriers. Leveraging their geosynchronous orbit, they facilitate large-scale, all-weather continuous observations at a minute-level temporal resolution. Multi-channel satellite radiative data can effectively characterize the macroscopic spatial distribution and evolution of cloud systems. Although satellites primarily detect cloud-top thermal radiation and microphysical properties, whereas radars measure the scattering characteristics of precipitation particles within and beneath clouds, a strong physical and statistical correlation exists between the two during severe convective events (Roberts and Rutledge, 2003; Wang et al., 2022b). Specifically, the pronounced decline in satellite infrared brightness temperature—indicative of deep convection and intense updrafts—aligns spatiotemporally with corresponding increases in radar composite reflectivity. This intrinsic physical linkage establishes a theoretical foundation for translating multi-channel satellite features into radar reflectivity fields. However, constructing such a precise mapping relationship poses substantial challenges. Traditional physical retrieval techniques and simplistic statistical models (e.g., linear regression, polynomial fitting) frequently fail due to the highly complex, nonlinear, and ill-posed relationship between 2D cloud-top radiative characteristics and 3D intra-cloud structures. Because conventional models rely heavily on manually crafted feature extraction rules, they struggle to fully exploit the high-dimensional spatiotemporal correlations embedded in multi-channel satellite datasets. Consequently, recent studies have increasingly turned to advanced deep learning techniques. Acting as intelligent expert systems with robust automatic feature extraction capabilities, Convolutional Neural Networks (CNNs)—particularly encoder-decoder architectures like U-Net—have demonstrated significant potential in meteorological image reconstruction due to their superior multi-scale feature extraction capabilities (Veillette et al., 2018; Duan et al., 2021b; Hilburn et al., 2021; Wang et al., 2025). Despite these substantial methodological advancements, the high-precision reconstruction of radar composite reflectivity over complex topographic regions remains constrained by several interconnected challenges (Chen et al., 2025; He et al., 2025; Gao et al., 2025; Biswas and Chandrasekar, 2018; Wang et al., 2022a; Hilburn et al., 2021; Zeng et al., 2025b; Yang et al., 2024; Sun et al., 2021b). Foremost among these is the detailed characterization of multi-scale echo structures. Severe convective echoes exhibit profound spatial heterogeneity, ranging from expansive mesoscale systems spanning hundreds of kilometers to highly localized, intense convective cores of merely a few kilometers. Standard U-Net architectures, while capable of recovering macroscopic morphology, frequently suffer from a persistent "smoothing effect," leading to the severe blurring of convective boundaries and the systematic underestimation of peak intensities within strong echo cores. Compounding this architectural limitation are the inherent difficulties associated with spatiotemporal alignment and data quality control. Geostationary satellites and radars inherently differ in spatial resolution (approximately 4 km versus 1 km) and observation latency. Inadequate spatiotemporal registration and the inclusion of low-quality samples (e.g., radar mosaic gaps) introduce severe systematic biases into the learning process. Furthermore, existing methodologies frequently struggle to efficiently utilize multi-channel radiometric information (Wang et al., 2020a; Sun et al., 2021a; Wang et al., 2025; Yang et al., 2023; Kou et al., 2022; Duan et al., 2021a; Liu et al., 2026; Yu et al., 2023a; He et al., 2022). While satellite data provide rich cloud-top ther-

https://doi.org/10.5194/egusphere-2026-2395
Preprint. Discussion started: 7 July 2026

modynamic signatures—for instance, water vapor channels are highly sensitive to deep convection, whereas infrared window channels correlate with cloud thickness—current retrieval networks often lack a lightweight, synergistic attention mechanism capable of adaptively integrating channel-wise spectral importance with precise spatial-wise core localization within a unified feature representation. To address the aforementioned challenges, this study proposes an intelligent, end-to-end reconstruction framework, termed CBAM-U-Net, designed specifically to compensate for radar blind spots over complex terrain utilizing multi-channel observations from the FY-4A geostationary satellite. Focusing on the topographically complex Sichuan Province during the 2023 extended flood season, we seamlessly integrate the Convolutional Block Attention Module (CBAM) into a U-Net backbone. By serially combining channel and spatial attention mechanisms within the same feature layer, the framework adaptively recalibrates the responses of critical spectral bands (e.g., water vapor channels, split-window brightness temperature differences) while precisely localizing strong echo cores and their high-gradient regions. The primary contributions of this paper are summarized as follows:

- An end-to-end multimodal retrieval framework. We developed a robust data-driven expert system that successfully maps FY-4A multi-channel radiometric signatures to high-fidelity radar composite reflectivity. By effectively bridging critical observational gaps over mountainous regions, this system empowers meteorological agencies to achieve continuous monitoring of severe convective weather.
- Dual-dimensional attention enhancement. By embedding a lightweight CBAM module within the U-Net architecture, the model fundamentally mitigates the traditional "smoothing effect." It achieves superior structural fidelity and accurate peak intensity estimation without incurring extensive computational overhead, ensuring high efficiency for real-time operational deployment.
- Rigorous operational validation. Evaluated comprehensively on a massive dataset (14,023 samples), the proposed model achieves state-of-the-art detection performance in challenging severe echo regimes (45–70 dBZ). Furthermore, it demonstrates robust cross-sensor zero-shot generalization (successfully migrating to FY-4B data), underscoring its immense value for real-time disaster early warning operations.

The remainder of this paper is organized as follows: Section 2 introduces the multi-source datasets, physical feature selection, and rigorous preprocessing procedures. Section 3 details the architecture of the proposed CBAM-U-Net expert system and its training strategies. Section 4 presents comprehensive comparative experiments, staged evaluations during the flood season, and typical case studies. Finally, Section 5 summarizes the findings, discusses limitations, and outlines future research directions.

2 Methodology

2.1 Study Area and Dataset Preparation

This study focuses on Sichuan Province and its adjacent regions in southwestern China (26°N–34.5°N, 97°E–109°E) as the primary study area. Positioned at the eastern foothills of the Qinghai-Tibet Plateau, this region serves as a critical geographic

https://doi.org/10.5194/egusphere-2026-2395
Preprint. Discussion started: 7 July 2026

transition zone connecting the first-step terrain (the Qinghai-Tibet Plateau) and the second-step terrain (the Sichuan Basin and surrounding mountains) of China's topographic macro-structure. As illustrated in Figure 1, the landforms within this domain are exceptionally complex and highly heterogeneous.

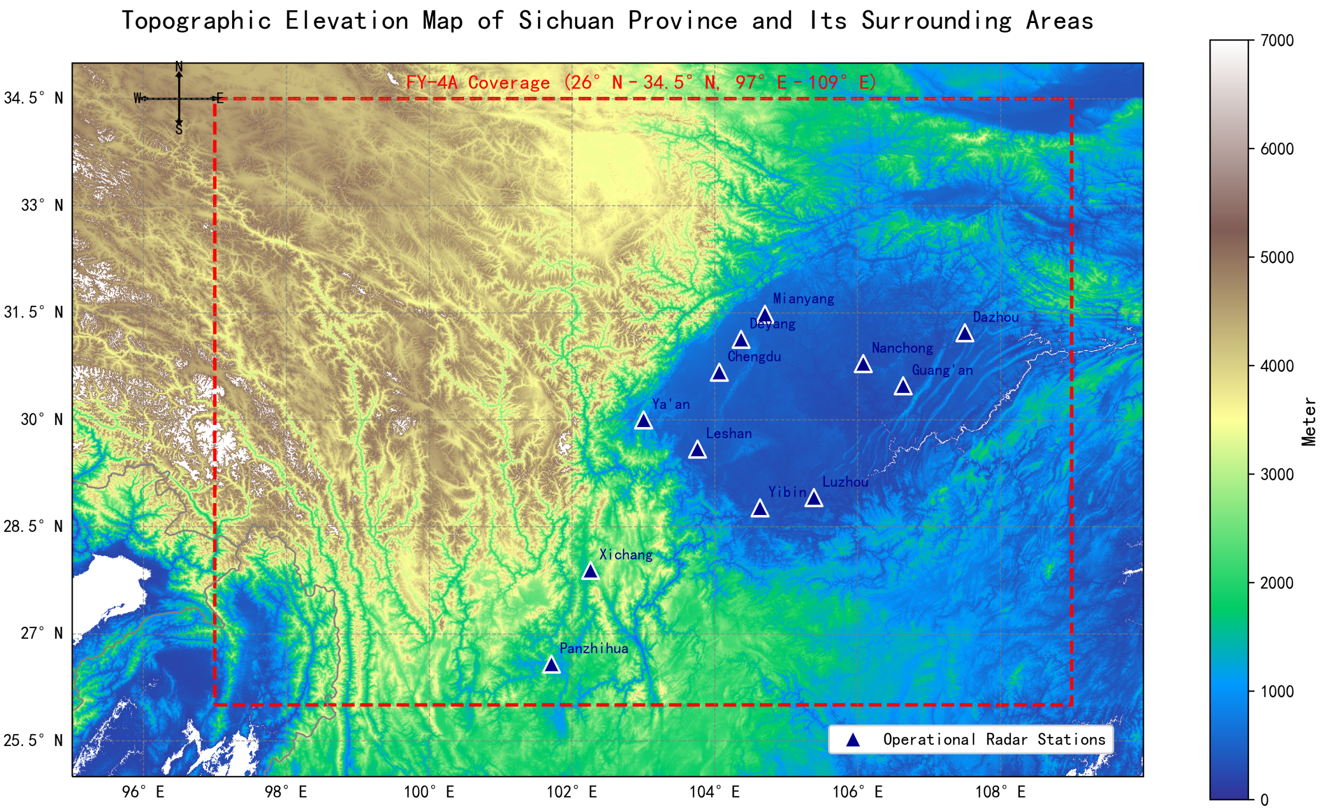

Figure 1. Topographic elevation map of Sichuan Province and surrounding areas

Topographically, the study area exhibits a pronounced stepped elevation gradient, descending sharply from west to east. The western sector is dominated by the Western Sichuan Plateau and the Hengduan Mountains, characterized by rugged terrain, deep intersecting gorges, and an average elevation exceeding 3,000 meters. Conversely, the eastern sector encompasses the typical Sichuan Basin, a relatively flat landscape with an average elevation of approximately 500 meters. This drastic topographic dichotomy, featuring an elevation drop of over 4,000 meters across a short horizontal distance, creates a highly unstable atmospheric environment. Such extreme terrain profoundly governs the genesis, evolution, and dissipation of local meso- and micro-scale weather systems. Dynamically, the massive topographic barrier of the Qinghai-Tibet Plateau induces profound dynamic blocking and thermal forcing, frequently generating cyclonic circulations on its leeward slopes (the western mountainous margins of the basin). These mechanisms provide optimal dynamic triggers for convective initiation. Thermodynamically, the steep mountainous slopes foster intense localized mountain-valley circulations. Coupled with abundant moisture convergence within the basin, these conditions form the renowned "Western Sichuan Rain Belt." Consequently, this region is highly prone

https://doi.org/10.5194/egusphere-2026-2395
Preprint. Discussion started: 7 July 2026

to abrupt, intense, and highly unpredictable severe convective events, including short-duration heavy rainfall and thunderstorm gales (Wang et al., 2023). Therefore, selecting Sichuan as the testbed for this research is highly strategic. It not only presents immense challenges for traditional satellite physical retrievals—owing to complex surface emissivity and severe terrain-induced cloud interference—but also serves as an ideal, rigorous scenario for evaluating the robustness and generalization capability of deep learning-based expert systems. Successfully reconstructing convective radar echoes over such intricate underlying surfaces convincingly demonstrates the model's operational viability in the most demanding meteorological environments.

2.2 Data Sources and Feature Engineering

2.2.1 FY-4A Geostationary Satellite Data

Fengyun-4A (FY-4A), the first flight unit of China's second-generation geostationary meteorological satellite network, was successfully launched in December 2016. Its primary payload, the Advanced Geosynchronous Radiation Imager (AGRI), features 14 spectral bands spanning a broad spectrum from visible to long-wave infrared. Compared to legacy sensors, AGRI offers substantial advancements in both spatiotemporal resolution and radiometric calibration precision (Shen et al., 2024). Crucially for this study, the AGRI infrared channels provide a spatial resolution of 4 km and a temporal scanning interval over the Chinese domain of approximately 5 minutes. This high temporal frequency closely aligns with the volumetric scanning cycle of operational weather radars, thereby providing an optimal temporal match to capture the rapid evolutionary dynamics of severe convective clouds (Cheng et al., 2024; Gao et al., 2024; Zhang et al., 2019). Furthermore, FY-4A Level-2 auxiliary products, such as cloud masks and cloud-top heights, supply supplementary information for rigorous data quality control and physical mechanism interpretation (Zhang et al., 2019). In data-driven predictive modeling, constructing a physically meaningful input space is paramount. The feature selection methodology employed in this investigation is deeply rooted in the integration of atmospheric radiative transfer theory and cloud microphysics. The objective is to establish a robust mapping relationship to infer the 3D "intra-cloud precipitation structure" from the 2D "cloud-top radiation field." Recognizing that strong radar reflectivity is primarily induced by the backscattering of large hydrometeors (e.g., raindrops, graupel, and hailstones), and acknowledging the limited penetration depth of individual infrared bands, this study meticulously selects five convection-sensitive channels alongside two physically derived Brightness Temperature Difference (BTD) features. Specifically, the 10.8 μm (Ch12) and 12.0 μm (Ch13) bands serve as critical split-window channels that directly quantify cloud-top brightness temperature and optical thickness, acting as proxies for the vertical development of convective systems. The 6.25 μm (Ch09) water vapor channel incorporates background moisture distribution in the mid-to-upper troposphere, outlining the environmental kinematic conditions surrounding deep convection. Additionally, the 3.72 μm (Ch08) and 8.5 μm (Ch11) channels leverage the refractive index disparities of hydrometeors across the short-wave and mid-wave infrared spectra. These channels provide vital microphysical indicators regarding cloud-top particle sizes and phase states (liquid water versus ice). To further augment the model's feature representation, two BTD formulations (BTD1 and BTD2) are engineered. These composite features significantly enhance the model's sensitivity to "overshooting tops" and mixed-phase regions while effectively suppressing background interference inherent to single-channel observations. This physically informed feature engineering approach provides a solid foundation for the deep learning framework to accurately extract and learn the nonlinear microphysical signatures embedded within the core regions of intense convection. A comprehensive overview of the selected channel parameters and their corresponding physical significance is presented in Table 1 (Yang et al., 2023).

Table 1. Parameters of input features (central wavelength and physical significance)

Category	Feature/Channel	Physical Significance
Infrared channel	Ch08 ($\approx$3.72 μm)	Cloud top texture, night low cloud/fog
Infrared channel	Ch09 ($\approx$6.25 μm)	Mid-upper tropospheric water vapor, deep convection
Infrared channel	Ch11 ($\approx$8.5 μm)	Water/ice cloud phase differences
Infrared channel	Ch12 ($\approx$10.8 μm)	Cloud top temperature (CTT), convective intensity
Infrared channel	Ch13 ($\approx$12.0 μm)	Split-window info, cloud particle size/thickness
Brightness temp diff	BTD1 = Ch12 − Ch09	Overshooting cloud tops, strong convection
Brightness temp diff	BTD2 = Ch13 + Ch11 − 2×Ch12	Sensitivity to ice/liquid water cloud particle phases

2.2.2 Radar Composite Reflectivity Data

The Radar Composite Reflectivity (RCRF) product utilized in this study is derived from the Severe Weather Automatic Nowcasting (SWAN) system, operated by the China Meteorological Administration (CMA). Serving as the cornerstone of the national operational short-term and nowcasting system for severe convective weather, SWAN integrates observational data from a dense network of over 200 S-band and C-band Doppler weather radars across China. Through rigorous quality control algorithms—encompassing ground clutter mitigation, outlier smoothing, and multi-radar 3D mosaicking—SWAN generates high-spatiotemporal-resolution radar products. The RCRF product covers a broad spatial domain (73.0°E–135.0°E, 12.2°N–54.2°N) with a temporal resolution of 6 minutes and a spatial resolution of 0.01° (approximately 1 km). These attributes provide high-frequency, high-fidelity reference labels, rendering them ideal for the supervised training of deep learning models aimed at refined regional precipitation monitoring. To ensure strict spatiotemporal alignment with the FY-4A satellite observations, the SWAN data were systematically cropped to the study area and resampled. Mathematically and physically, RCRF (unit: dBZ) is a two-dimensional projection obtained by extracting the maximum reflectivity factor from multiple elevation sweeps within a single volumetric scan onto a horizontal plane. Its fundamental objective is to capture the strongest echo signal along

https://doi.org/10.5194/egusphere-2026-2395
Preprint. Discussion started: 7 July 2026

the vertical atmospheric column, thereby precisely delineating the convective cores and precipitation intensity. Let $Z(\theta, r, \phi)$ denote the reflectivity factor at elevation angle θ, slant range r, and azimuth ϕ. The composite reflectivity is defined as:

$$RCRF(r,\varphi) = \max_{\theta \in \{\theta_1, \theta_2, \ldots, \theta_n\}} Z(\theta, r, \varphi) \tag{1}$$

where $\{\theta_1, \theta_2, \ldots, \theta_n\}$ represents the set of elevation angles included in a single volume scan, and n denotes the total number of elevation layers. Compared to reflectivity at a single elevation angle (e.g., 0.5° or 1.5°) or a Constant Altitude Plan Position Indicator (CAPPI, e.g., 3 km), RCRF consolidates the peak intensity information throughout the entire vertical structure of the cloud system. This attribute makes it significantly more effective for identifying severe convective systems characterized by profound vertical development and their associated heavy precipitation centers. Consequently, this study designates the RCRF as the retrieval target (Ground Truth) for the proposed deep learning expert system. This selection provides a rigorous benchmark to evaluate the capability of FY-4A multi-channel observations in accurately characterizing and reconstructing precipitation echo structures across complex topography.

2.2.3 Dataset Construction and Preprocessing Pipeline

This study selected the extended flood season (May to November 2023) in the Sichuan region as the research period. During this timeframe, FY-4A/AGRI Level-1 multi-channel observations and their corresponding RCRF mosaic products were systematically collected to construct the training and testing datasets. The spatial domain was rigorously cropped to 97°E–109°E and 26°N–34.5°N, encompassing highly complex terrain units, including the Sichuan Basin, the Western Sichuan Plateau, and the mountainous Panxi region. To reconcile the intrinsic disparities in spatial resolution, map projections, and observation timestamps inherent to multi-source heterogeneous data, a rigorous data preprocessing pipeline was engineered. This pipeline integrates geometric correction, precise spatiotemporal matching, and stringent quality control protocols. Ultimately, it generates high-fidelity "satellite-radar" image pairs, which are indispensable for the supervised training of the proposed deep learning framework. The comprehensive dataset construction workflow is illustrated in Figure 2, comprising the following principal steps:

(1) FY-4A AGRI L1 geometric positioning and latitude-longitude conversion

The FY-4A/AGRI Level 1 data are typically structured according to line and column indices within the framework of a geostationary orbit projection. This organization necessitates the conversion of pixel line and column indices into corresponding latitude and longitude coordinates through the application of navigation parameters. In alignment with the data format and navigation specifications outlined for FY-4A, this paper delineates the subsequent procedure for the inverse calculation of latitude and longitude coordinates. Initially, an intermediate angular variable (expressed in radians) is defined, derived from the row and column indices:

$$x = \frac{\pi \cdot (c - \mathrm{COFF})}{180 \cdot 2^{-16} \cdot \mathrm{CFAC}} \tag{2}$$

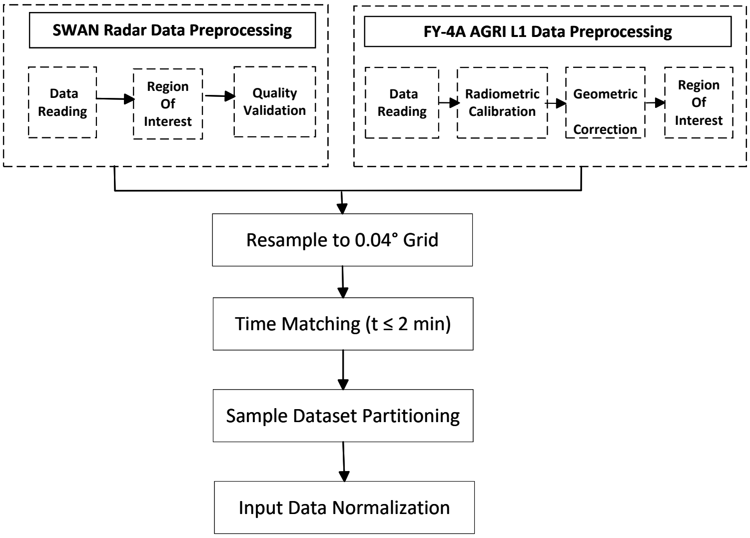

Figure 2. Workflow of dataset construction for FY-4A and radar composite reflectivity data

$$y = \frac{\pi(1 - \mathrm{LOFF})}{180 \cdot 2^{-16} \cdot \mathrm{CFAC}} \tag{3}$$

Where c and l represent the column and row numbers of the pixel, respectively; COFF/CFAC denote the column-direction offset and scale factor, and LOFF/LFAC represent the row-direction offset and scale factor (in practical applications, it is advisable to first read the navigation parameters within each data file to avoid manual hardcoding). For the FY-4A AGRI 4 km resolution product, typical values commonly reported in open research are $\mathrm{COFF} = \mathrm{LOFF} = 1373.5$ and $\mathrm{CFAC} = \mathrm{LFAC} = 10233137$, after which other intermediate variables can be calculated as follows:

$$S_d = \sqrt{(h\cos(x)\cos(y))^2 - \left(\cos^2(y) + \frac{ea^2}{eb^2}\sin^2(y)\right)(h^2 - ea^2)} \tag{4}$$

$$S_n = \frac{h\cos(x)\cos(y) - S_d}{\cos^2(y) + \frac{ea^2}{eb^2}\sin^2(y)} \tag{5}$$

$$S_1 = h - S_n\cos(x)\cos(y) \tag{6}$$

$$S_2 = S_n\sin(x)\cos(y) \tag{7}$$

$$S_3 = -S_n \sin(y) \tag{8}$$

$$S_{xy} = \sqrt{S_1^2 + S_2^2} \tag{9}$$

where h denotes the distance from the geocenter to the center of mass of the satellite, and e_a and e_b represent the Earth's semi-major and semi-minor axes, respectively. The specific values of these parameters are set as follows: $h = 42164$ km, $e_a = 6378.137$ km, and $e_b = 6356.7523$ km. Finally, the conversion formulas for the satellite's longitude and latitude are established based on these intermediate variables, where λ_D refers to the longitude of the satellite's subsatellite point.

$$lon = \frac{180}{\pi} \cdot \arctan\left(\frac{S_2}{S_1}\right) + \lambda_D \tag{10}$$

$$lat = \frac{180}{\pi} \cdot \arctan\left(\frac{e_a^2}{e_b^2} \cdot \frac{S_3}{S_{xy}}\right) \tag{11}$$

(2) Satellite data resampling

To achieve spatial congruence with radar tags, this study resamples the multi-channel observations obtained from FY-4A, which have been geolocated using latitude and longitude coordinates, onto a grid with a resolution of 0.04°×0.04° in both latitude and longitude. The resampling procedure employs nearest-neighbor interpolation to mitigate the numerical smoothing of radiometric values, thereby preserving the textural characteristics of cloud regions. This grid configuration is congruent with the subsequent aggregation of radar data, which varies in resolution from 0.01° to 0.04°.

(3) Radar data quality screening and effective proportion definition

The original spatial resolution of the radar composite reflectivity is 0.01°. Given that radar mosaicking and areas of missing measurements may result in the presence of invalid pixels, this study computes the "effective data proportion" for each radar sample and subsequently applies a filtering process to the samples. The effective proportion is defined as follows:

$$P_{\text{valid}} = \frac{N_{\text{valid}}}{N_{\text{total}}} \times 100\% \tag{12}$$

where N_{valid} denotes the number of valid pixels in the sample, and N_{total} represents the total number of pixels within the study area. Taking August 2023 as an example, the number of original samples was 6039; after removing low-quality samples with an excessively low valid ratio (e.g., 0–10%), this study retained samples with $P_{\text{valid}} > 20\%$ for model training and evaluation. This approach ensures the usability of labels while preserving the structural characteristics of convective echoes as much as possible. After screening according to this criterion, a total of 14,023 valid samples were finally obtained. The distribution of the valid ratio is presented in Figure 3.

(4) Space matching

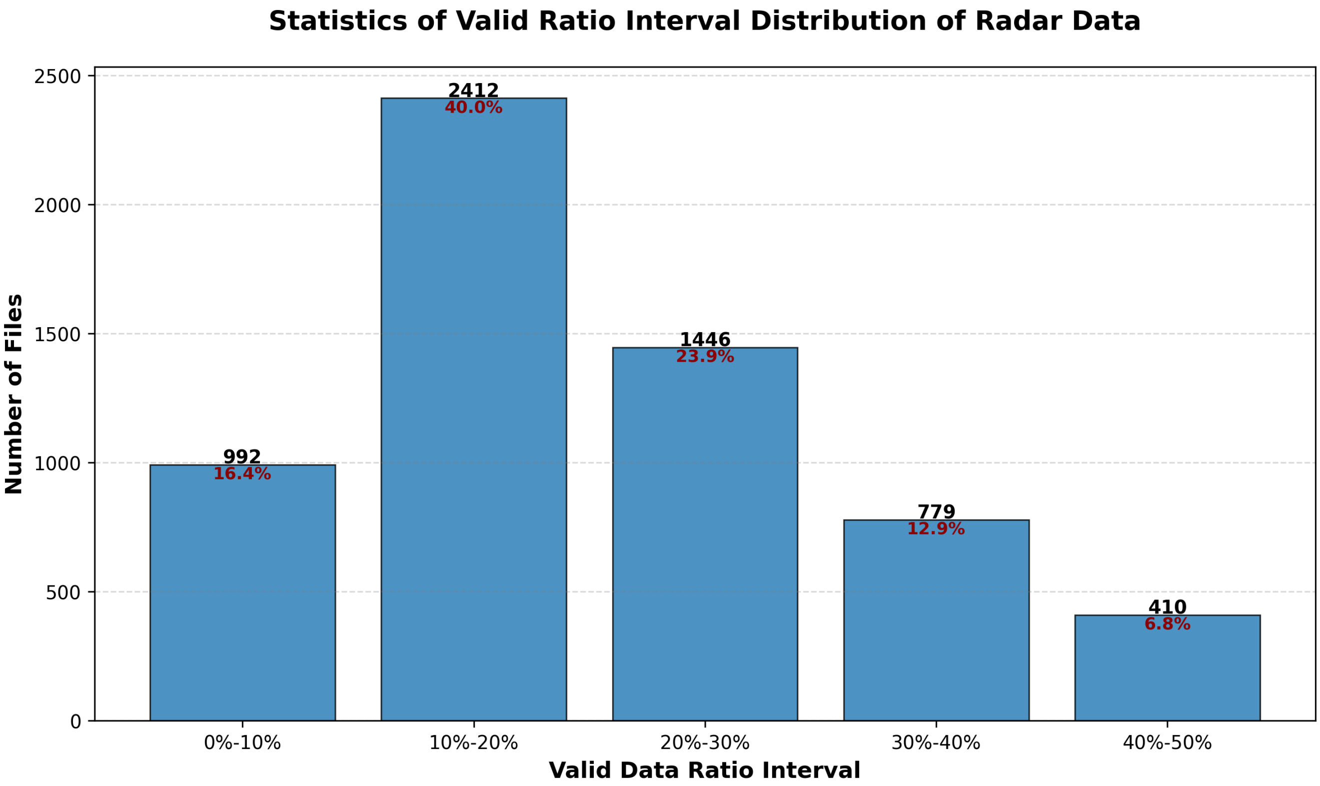

Figure 3. Distribution of effective proportion of radar combined reflectivity in Sichuan region (August 2023)

Regarding spatial resolution, the native grid of FY-4A satellite data is approximately $0.04°$, whereas the grid of SWAN radar products is $0.01°$. This implies that one satellite pixel corresponds to $4 \times 4 = 16$ radar sub-pixels. To preserve the core characteristics of strong convective systems during the upscaling process and avoid the smoothing of extreme values caused by averaging, this study adopts a max-pooling strategy for spatial matching. Specifically, the maximum value among the 16 radar observations at $0.01°$ resolution within a $0.04°$ grid cell is assigned as the label value for that grid cell (Yongguang et al., 2015):

$$\mathrm{RCRF}_{0.04^\circ}(i,j) = \max_{(m,n)\in\Omega_{i,j}} \mathrm{RCRF}_{0.01^\circ}(m,n) \tag{13}$$

where $\Omega_{i,j}$ denotes the set of 16 radar sub-pixels projected within the range of the (i,j)-th satellite pixel. This strategy can effectively preserve the peak information of strong echoes, thereby improving the accuracy of the model in locating heavy precipitation centers and optimizing its Probability of Detection (POD) and Critical Success Index (CSI) in the threshold range of strong echoes (e.g., > 45 dBZ).

(5) Time matching

https://doi.org/10.5194/egusphere-2026-2395
Preprint. Discussion started: 7 July 2026

In the temporal dimension, considering the discrepancy between the scan cycle of FY-4A over China (approximately 5 minutes) and the volume scan cycle of the radar (approximately 6 minutes), this study sets a strict temporal matching window. Taking the satellite observation time T_{sat} as the reference, the nearest radar observation time T_{rad} is retrieved within the time window $[T_{\text{sat}} - 2\,\text{min}, T_{\text{sat}} + 2\,\text{min}]$. If valid radar data exist within this window, the time step with the smallest time difference $\Delta T = |T_{\text{sat}} - T_{\text{rad}}|$ is selected as the matching label; otherwise, the satellite sample at that time is discarded. This temporal tolerance setting aims to balance the number of samples and the temporal error, minimizing the "satellite-radar" structural bias caused by the rapid movement, generation, or dissipation of cloud systems.

(6) Data standardization

To accelerate model convergence and improve training stability, this study applies Z-Score normalization to all input features (FY-4A multi-channel brightness temperatures and BTDs). Specifically, the channel-wise mean (μ) and standard deviation (σ) are calculated over the entire training set, and the input data are transformed to a zero-mean, unit-variance distribution:

$$x' = \frac{x - \mu}{\sigma} \tag{14}$$

This step mitigates the discrepancies in magnitude and numerical scale across different channels (e.g., the distinct value ranges associated with water vapor channels versus window region channels), thereby enhancing the search efficiency of the optimization algorithm on the loss surface. It is important to specifically note that the normalization of inputs in this context falls under the category of data preprocessing, with the objective of improving the distribution characteristics of the input data. In contrast, Batch Normalization (BN), which is frequently employed in deep learning architectures, constitutes a learnable operation within the layers of the network, specifically designed to address the issue of internal covariate shift (Ioffe and Szegedy, 2015).

2.3 Evaluation Method

To thoroughly assess model performance, this study establishes a comprehensive evaluation framework encompassing three dimensions: numerical regression accuracy, structural consistency of images, and the capability to capture hierarchical events. (detailed formulas are provided in Table 2).

Table 2. Evaluation Metrics and Corresponding Formulas

Metric Formula	**Parameter Description**	**Value Range**	**Physical Meaning**
$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - y_{\mathrm{pred}})^2}$	y_i: Ground truth; y_{pred}: Prediction; N: Total samples; $\bar{y}_i$: Mean of ground truth	$[0, +\infty)$	Penalizes large errors, sensitive to outliers, widely used for regression evaluation.
$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\lvert y_i - y_{\mathrm{pred}}\rvert$		$[0, +\infty)$	Calculates average absolute error, robust to outliers with stable performance.
$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - y_{\mathrm{pred}})^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2}$		$(-\infty, 1]$	Quantifies goodness of fit; closer to 1 indicates stronger regression ability.
$\mathrm{PSNR} = 10\log_{10}\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right)$	MAX: Peak image signal value; MSE: Mean Square Error	$20 \sim 40$ dB	Higher value indicates smaller distortion and better reconstruction quality.
$\mathrm{SSIM} = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$	μ_x, μ_y: Mean brightness; σ_x^2, σ_y^2: Variance; σ_{xy}: Covariance; c_1, c_2: Stability constants	$[0, 1]$	Measures similarity in brightness, contrast and structure; 1 = perfect consistency, 0 = no similarity.
$\mathrm{POD} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}$	TP: True Positive;FN: False Negative; FP: False Positive	$[0, 1]$	Reflects event capture ability; closer to 1 indicates higher hit rate.
$\mathrm{FAR} = \frac{\mathrm{FP}}{\mathrm{TP}+\mathrm{FP}}$		$[0, 1]$	Quantifies false alarm level; lower value indicates more reliable predictions.
$\mathrm{CSI} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}+\mathrm{FN}}$		$[0, 1]$	Comprehensively balances hit rate and false alarm rate, core metric for event evaluation.

https://doi.org/10.5194/egusphere-2026-2395
Preprint. Discussion started: 7 July 2026

3 Model

3.1 Model Architecture and Experimental Setup

Extensive research demonstrates that U-Net-style encoder-decoder architectures, owing to their profound proficiency in multi-scale feature extraction, are highly effective at reconstructing the macroscopic morphology and spatial distribution of meteorological fields (Ronneberger et al., 2015; Yu et al., 2023b; Veillette et al., 2020). However, severe convective echoes exhibit highly complex structural characteristics, typically featuring pronounced central intensity gradients and intricate edge textures. Confronted with these patterns, conventional U-Net models frequently struggle to reconcile macroscopic global structures with localized peak details. While various attention mechanisms have been introduced to help models focus on salient regions, conventional single-dimensional attention strategies—whether purely channel-wise or spatial-wise—often fail to simultaneously achieve discriminability in the spectral dimension and selectivity in the spatial domain. Furthermore, although some parallel or hierarchical dual-attention frameworks address this gap, they inherently impose substantial computational overhead. This increased complexity not only hampers operational efficiency but also significantly elevates the risk of overfitting, particularly when training on finite spatiotemporal meteorological datasets. To overcome these architectural bottlenecks, this study proposes an advanced, lightweight deep learning expert system, designated as CBAM-U-Net. By seamlessly integrating the Convolutional Block Attention Module (CBAM) into the standard U-Net backbone, this framework is specifically engineered to process the 7-channel input tensor derived from FY-4A/AGRI observations (comprising five independent infrared bands and two composite brightness temperature difference features). The primary objective is to establish a robust, end-to-end nonlinear mapping from multidimensional satellite cloud-top characteristics to the target Radar Composite Reflectivity (RCRF). To ensure rigorous and reproducible model evaluation, the rigorously preprocessed dataset was partitioned into training, validation, and testing subsets following a standard 7:2:1 ratio. In the comparative experiments detailed in Section 4, a comprehensive evaluation of CBAM-U-Net was conducted against several state-of-the-art CNN- and Transformer-based baselines. A specific emphasis was placed on validating the proposed model's robustness, generalization capabilities, and enhancement effects across varying precipitation intensity thresholds.

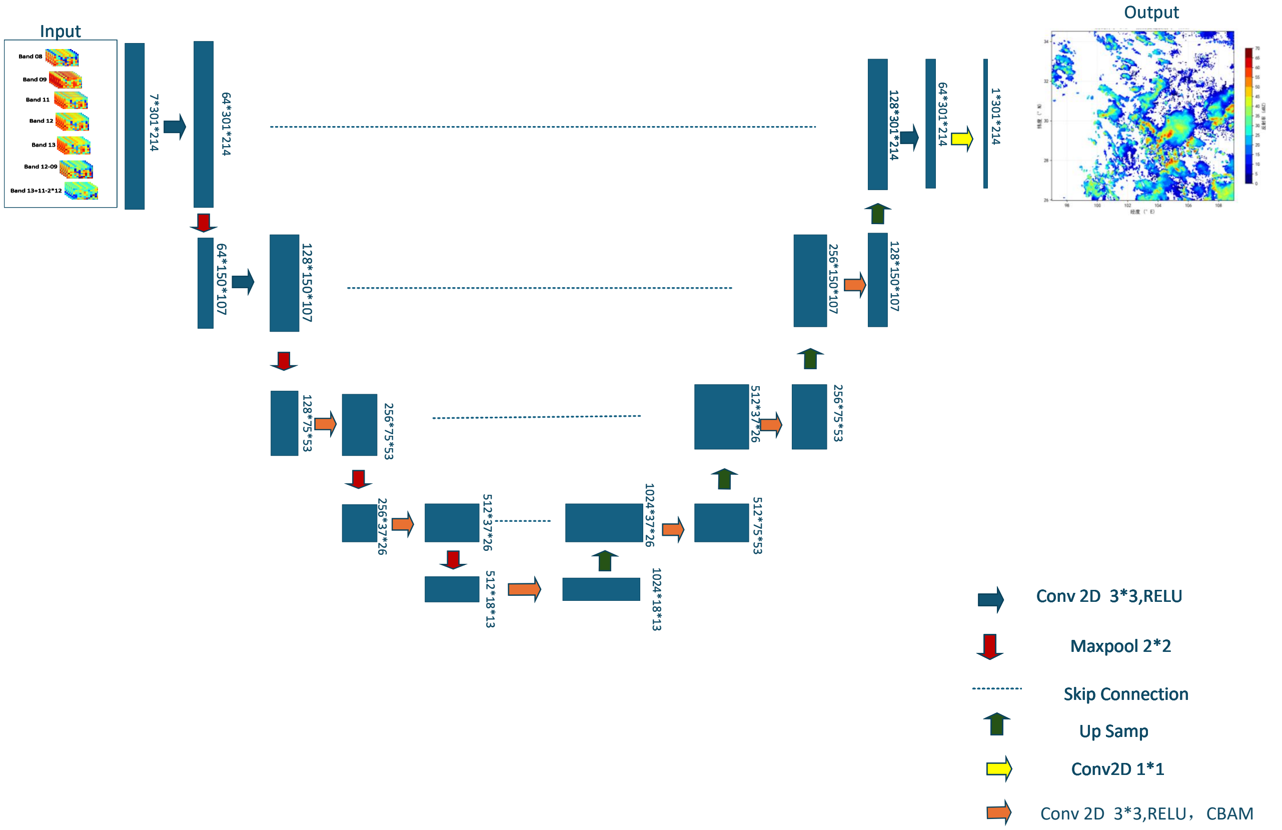

Figure 4. Framework of the FY-4A-driven CBAM-U-Net for radar composite reflectivity reconstruction over complex terrain

3.2 Channel Attention Mechanism

In Convolutional Neural Networks (CNNs), feature maps across various channels inherently encapsulate distinct, multi-level semantic responses. During the reconstruction of radar reflectivity from multi-band satellite imagery, certain channels are intrinsically more informative than others. For instance, specific channels capture high-frequency thermal responses at severe convective cores, while others (e.g., water vapor channels) delineate critical environmental moisture textures. To empower the network to adaptively discern and prioritize these highly diagnostic feature channels, this study seamlessly integrates the Channel Attention Module (CAM) (Woo et al., 2018). The CAM explicitly models the inter-channel relationships of the input feature maps, dynamically learning the importance weights for each channel to achieve comprehensive feature recalibration. The specific computational workflow is meticulously designed as follows: First, rather than relying on a traditional single-pooling strategy, the module simultaneously executes both Global Average Pooling (GAP) and Global Max Pooling (GMP) across the spatial dimensions of the input feature map. This dual-pooling mechanism is critical: GAP captures the global statistical response intensity of the channel, while GMP distinctly preserves the localized, highly salient extrema (such as the intense, isolated peaks of convective storm cores). Consequently, this step achieves a robust, complementary characterization of both global context and localized prominent features. Subsequently, the two resulting spatial-context descriptors are independently fed into a shared Multi-Layer Perceptron (MLP) network. To balance computational efficiency with representational capacity, this shared MLP is architecturally designed with a bottleneck structure—initially compressing the dimensionality (via a reduction ratio) to minimize computational overhead, and subsequently restoring the original channel dimensions. This process facilitates the learning of nonlinear cross-channel interactions. Finally, the output feature vectors from the shared MLP are merged via element-wise summation and subjected to a Sigmoid activation function to generate the final channel attention map (a normalized weight vector). These learned weights are then applied multiplicatively, channel-by-channel, to the original input feature map. Through this adaptive recalibration process, the semantic responses of channels most critical to radar echo retrieval are selectively amplified, while less informative or noisy background channels are effectively suppressed.

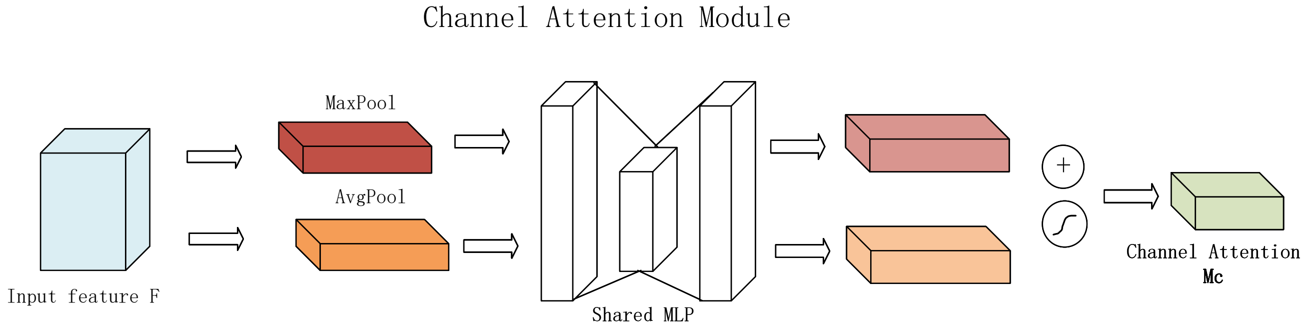

Figure 5. Channel Attention Module (CAM)

3.3 Spatial Attention Mechanism

Complementary to the channel attention, the spatial attention mechanism employed in this research utilizes the Spatial Attention Module (SAM) (Woo et al., 2018). While CAM focuses on "what" features are meaningful, SAM is explicitly designed to determine "where" these critical features are located. This is particularly vital for severe convective weather applications, as it empowers the model to dynamically focus on specific spatial locations—such as intense convective cells or structural cloud walls—while effectively suppressing widespread irrelevant background regions (e.g., clear skies or non-precipitating cirrus clouds). The primary operational workflow of SAM is outlined as follows: Initially, to aggregate rich spatial information while significantly reducing computational overhead, average pooling and max pooling operations are applied simultaneously along the channel axis of the channel-refined feature map (the output from CAM). This procedure generates two distinct 2D spatial feature descriptors, intuitively representing the average spatial distribution of cloud properties and the peak spatial locations of intense convective signatures across all channels. Subsequently, these two diverse spatial descriptors are concatenated along the channel dimension. The fused feature map is then processed through a large receptive field convolution—specifically, a

7×7 convolutional layer. This large kernel size is deliberately chosen to effectively capture spatial dependencies and contextual information over a broader region of the satellite imagery. The convolution operation simultaneously compresses the channel dimension from two to one, culminating in a single-channel spatial map while meticulously preserving the original spatial resolution (H×W). Finally, the compressed spatial feature map is normalized using a Sigmoid activation function to derive the spatial attention weights (ranging from 0 to 1). These normalized weights are subsequently broadcast and applied via element-wise multiplication to the input channel-refined feature map. This spatial recalibration process dynamically emphasizes the structurally significant regions of precipitation systems, culminating in the final, robust attention-enhanced feature representation ready for subsequent multi-scale decoding.

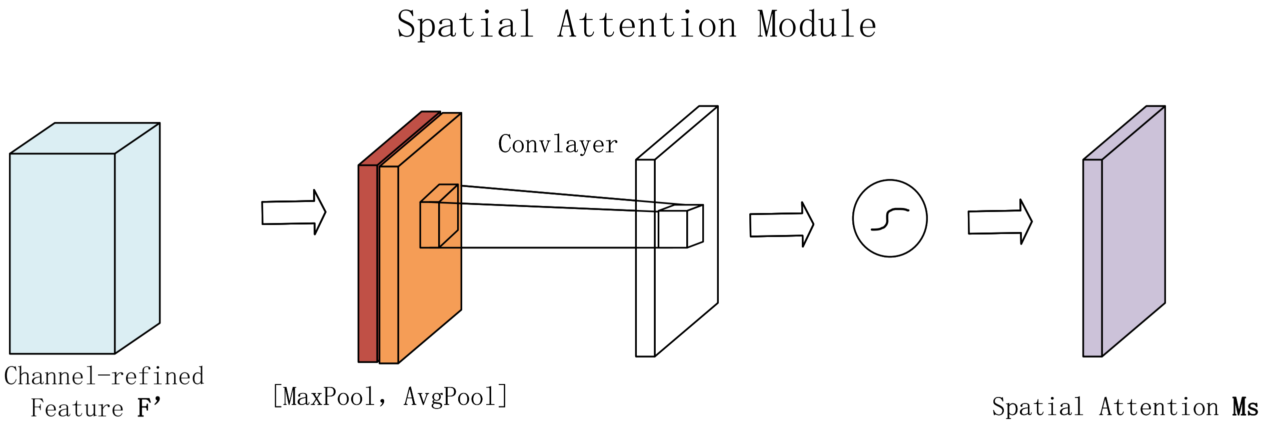

Figure 6. Spatial Attention Module (SAM)

3.4 Baseline Models for Comparison

To comprehensively evaluate the effectiveness, robustness, and generalization capability of the proposed CBAM-U-Net in retrieving Radar Composite Reflectivity (RCRF) from FY-4A multi-channel observations, this study establishes a robust set of comparative baselines. These baselines span a spectrum of architectural paradigms, encompassing classical CNNs, attention-augmented CNNs, and advanced Transformer-based global modeling frameworks. The CNN family is anchored by the foundational encoder-decoder architecture, U-Net (Yu et al., 2023b). Renowned for its elegant multi-scale feature fusion mechanism via skip connections, U-Net serves as the standard baseline for meteorological image reconstruction. Building upon this foundation, Attention U-Net (Chen et al., 2023) incorporates a gating mechanism designed to implicitly suppress irrelevant background noise while highlighting salient features. Furthermore, MCDA-U-Net (Zeng et al., 2025a) refines local feature representations through a more complex, multi-scale coupled attention strategy. In stark contrast to these complex variants, the CBAM-U-Net proposed herein champions a lightweight, cascaded dual-attention strategy. It strives to achieve a synergistic enhancement of spectral filtering (channel) and spatial focusing (spatial) with minimal computational overhead. To further investigate the potential benefits of global self-attention modeling, this study integrates two prominent Transformer-based architectures into the baseline suite: SegFormer (Li et al., 2025) and SRViT (Stock et al., 2024). The former exhibits remarkable capability in capturing extensive spatial contextual relationships without positional encoding, while the latter excels in the high-resolution restoration of structural consistency. Through a rigorous comparative analysis of these distinct architectural paradigms, this research aims to ascertain whether CBAM-U-Net can achieve an optimal equilibrium between computational efficiency and the high-fidelity reconstruction of intricate structural details—particularly in real-world scenarios constrained by limited sample availability. Ultimately, this investigation aspires to demonstrate that the proposed CBAM-U-Net offers superior robustness and operational viability in retrieving severe convective echoes compared to increasingly complex, computationally intensive global modeling methodologies. A comprehensive analysis summarizing the core architectural attributes, respective strengths, and inherent limitations of each model evaluated in this study is detailed in Table 3.

Table 3. Comparison of the advantages and disadvantages of different models.

Model	Advantages	Disadvantages
U-Net	Simple structure, easy to reproduce and train; fewer parameters and high inference speed.	No attention mechanism and weak long-distance dependency modelling; limited performance for complex boundaries and small targets.
Attention-U-Net	Introduces attention gates to focus on key regions; suppresses background noise and improves retrieval accuracy.	Increased computational cost caused by attention calculation; limited global modelling capability.
MCDA-U-Net	Uses a multi-scale coupled attention mechanism; improves local detail representation.	More complex structure; larger number of parameters and slower training speed.
CBAM-U-Net	Lightweight structure with low computational cost; jointly enhances channel and spatial features; suppresses background noise and improves edges and small targets; stable training and low risk of overfitting.	Slight redundancy of attention weights may occur during training; performance still depends on the backbone network, and CNN-based local convolution has limited long-distance dependency modelling capability.
SegFormer	Lightweight Transformer structure with high efficiency; hierarchical feature extraction without positional encoding.	Requires a large amount of data and is prone to overfitting with limited samples; relatively slow training speed.
SRViT	Strong global modelling capability; suitable for fine-grained high-resolution reconstruction.	Large number of parameters and high computational requirements; prone to overfitting with limited data.

4 Results and Discussion

To comprehensively assess the retrieval robustness and operational applicability of the proposed model across diverse meteorological backgrounds, the test dataset (spanning May to November 2023) was systematically partitioned into three distinct climatic phases. This temporal stratification aligns seamlessly with the climatological characteristics and operational meteorological standards of the Sichuan region: the pre-flood onset phase (May–June), the primary flood season (July–August), and

https://doi.org/10.5194/egusphere-2026-2395
Preprint. Discussion started: 7 July 2026

the post-flood transition phase (September–November). Guided by this temporal framework, a rigorous, two-tiered evaluation strategy was executed. First, a comprehensive statistical evaluation was conducted on the aggregate test set encompassing the entire extended flood season. Subsequently, in-depth quantitative assessments and visual analyses of representative case studies from each distinct phase were performed. This methodology aims to profoundly explore the model's performance variations and generalization capabilities under varying synoptic regimes.

4.1 Comprehensive Statistical Evaluation

To rigorously benchmark the competing models on the complete test dataset, a multi-dimensional evaluation metric system was established, assessing model performance from three primary perspectives:

- Pixel-wise regression metrics. Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2). These metrics are utilized to quantify the absolute, pixel-level intensity deviations between the deep learning retrieved echoes and the ground-truth radar observations.
- Structural and perceptual fidelity. Peak Signal-to-Noise Ratio (PSNR) and the Structural Similarity Index (SSIM). Borrowed from computer vision, these indices are employed to evaluate the macroscopic morphological consistency and spatial texture preservation of the reconstructed convective cloud systems.
- Threshold-based meteorological verification. Probability of Detection (POD), False Alarm Ratio (FAR), and the Critical Success Index (CSI). These represent the standard categorical verification metrics in radar meteorology. They are calculated across various echo intensity thresholds to evaluate the model's diagnostic accuracy and operational reliability in capturing convective events of differing severities.

4.1.1 Overall regression and structural performance analysis

Table 4 presents the comparison of regression accuracy and image reconstruction quality across different models on the full test set.

Table 4. Comparison of Overall Evaluation Metrics for Different Models

Model	RMSE	MAE	R^2	PSNR	SSIM
U-Net	7.0043	5.0375	0.6083	19.9947	0.7506
Attention-U-Net	6.9997	4.9964	0.6200	20.1253	0.7578
MCDA-U-Net	7.0091	5.9418	0.5879	18.8303	0.7002
SegFormer	8.7098	6.6022	0.3944	18.1018	0.6525
SRViT	11.0277	8.7922	0.2092	16.0522	0.5574
CBAM-U-Net	**6.8290**	**4.8579**	**0.6277**	**25.2701**	**0.7894**

As shown in Table 4, CBAM-U-Net achieves the best performance across all evaluation metrics. In terms of regression accuracy, it attains the lowest RMSE (6.8290) and MAE (4.8579), along with an R^2 of 0.6277, indicating that the model can more accurately fit the numerical distribution of the RCRF field. Regarding structural fidelity, CBAM-U-Net achieves a notably high SSIM of 0.7894, significantly outperforming other comparative models. This demonstrates that the integration of the channel–spatial joint attention mechanism effectively enhances the model's capability to reconstruct the spatial structure, texture details, and boundary gradients of radar echoes. It is noteworthy that the overall performance of convolutional neural network (CNN)-based models significantly exceeds that of transformer-based models. This observation underscores the advantages of the U-Net architecture in pixel-level regression tasks, as it effectively utilizes an encoder to extract deep semantic features and a decoder to restore spatial resolution. Additionally, the incorporation of skip connections facilitates the integration of shallow details, rendering it particularly proficient in processing meteorological radar echo data that is characterized by fine-grained local features. Conversely, while SegFormer and SRViT exhibit robust global modeling capabilities, their performance is notably diminished in contexts involving limited sample sizes and the recovery of high-frequency details, as evidenced by a marked decline in performance metrics. This decline can be attributed to the inherent reliance of Vision Transformers on large datasets, coupled with their relative inadequacy in capturing local texture information.

4.1.2 Hierarchical echo event verification and analysis

To further evaluate the model's capability to capture precipitation events of different intensities, this study divides radar echo intensity (dBZ) into three intervals for event-based verification, with the results presented in Table 5.

Table 5. Event test results of different models under various echo intensity thresholds

Model	**POD**			**FAR**			**CSI**		
	0–25 dBZ	25–45 dBZ	45–70 dBZ	0–25 dBZ	25–45 dBZ	45–70 dBZ	0–25 dBZ	25–45 dBZ	45–70 dBZ
U-Net	0.9425	0.6241	0.2617	0.1094	0.2624	0.5954	0.8451	0.5107	0.3141
Attention-U-Net	0.9423	0.6060	0.3725	0.1106	0.2672	0.5069	0.8474	0.5054	0.3268
MCDA-U-Net	0.9216	0.5736	0.3572	0.1252	0.3393	0.4609	0.8142	0.4431	0.3278
SegFormer	0.8739	0.6199	0.1793	0.1176	0.4243	0.5382	0.7827	0.4255	0.1483
SRViT	0.8128	0.4751	0.0453	0.1703	0.5787	0.7330	0.6966	0.2875	0.0403
CBAM-U-Net	**0.9443**	**0.6417**	**0.5296**	**0.0998**	**0.2356**	**0.4357**	**0.8549**	**0.5256**	**0.4384**

As can be seen from Table 5, CBAM-U-Net demonstrates a comprehensive advantage characterized by a "high Probability of Detection (POD) coupled with a low False Alarm Rate (FAR)" across weak, moderate, and severe echo intervals, indicative of its exceptional inversion stability. Notably, within the severe echo interval (45–70 dBZ), which presents the greatest challenges for inversion, the performance enhancements of CBAM-U-Net are particularly pronounced: its POD achieves a value of 0.5296, the Critical Success Index (CSI) increases to 0.4384, and the FAR decreases to 0.4357. These results suggest that, due to the CBAM module's ability to filter key spectral channels and concentrate on the core regions of severe convection, the

https://doi.org/10.5194/egusphere-2026-2395
Preprint. Discussion started: 7 July 2026

model effectively mitigates the learning challenges associated with the scarcity of severe echo samples, thereby significantly enhancing its capacity to capture the central areas of severe convection. In conclusion, despite the comprehensive dataset encompassing a variety of complex meteorological patterns—including the pre-flood season in South China, the Meiyu period, high-temperature and low-rainfall conditions governed by the subtropical high, and autumn precipitation in Western China—CBAM-U-Net consistently exhibits optimal performance across all evaluative dimensions. This underscores the robustness and generalization potential of the proposed methodology in the face of intricate and variable meteorological conditions.

4.2 Analysis of flood season classification

4.2.1 The pre-flood onset phase (May-June)

In the Sichuan region, the interval from May to June signifies the pre-flood onset phase, shaped by the peripheral circulation associated with the pre-summer rainy season in South China and the leading edge of the Meiyu front. During this period, the atmospheric stratification exhibits instability, frequently resulting in localized short-duration heavy rainfall and sporadic light rain events. The weather phenomena during this timeframe are characterized by rapid convective initiation and dissipation, small spatial scales, and discrete spatial distributions, leading to substantial variability in radar echo data. This variability presents considerable challenges for feature extraction by predictive models. Notwithstanding the disruptions caused by high data variability, CBAM-U-Net, by leveraging its adaptive focusing capabilities on critical channels and spatial locations, consistently achieves the lowest inversion error and optimal image reconstruction quality.

(1) Quantitative indicator analysis

Tables 6 and 7 present the overall regression error and hierarchical event verification results of each model in this stage, respectively.

Table 6. Assessment value of comparison model for the pre-flood onset phase

Model	RMSE	MAE	R^2	PSNR	SSIM
U-Net	7.3726	6.3742	0.4586	19.5459	0.6941
Attention-U-Net	7.3252	5.8361	0.4729	19.7491	0.7192
MCDA-U-Net	7.5869	5.5666	0.4267	19.3007	0.6848
SegFormer	8.0911	6.1252	0.3479	18.7418	0.6468
SRViT	9.5448	7.5664	0.2925	17.3066	0.5800
CBAM-U-Net	**7.1048**	**5.0913**	**0.4972**	**24.8149**	**0.7617**

As can be seen from Table 6, CBAM-U-Net exhibits a prominent advantage in stability during the pre-flood period with large data fluctuations. It achieves the lowest RMSE (7.1048) and MAE (5.0913) among all comparative models, along with an SSIM of 0.7617, indicating that the model can well preserve the spatial structure integrity of radar echoes.

Table 7. Threshold-Based Event Verification of Comparative Models during the pre-flood onset phase

Model	**POD**			**FAR**			**CSI**		
	0–25 dBZ	25–45 dBZ	45–70 dBZ	0–25 dBZ	25–45 dBZ	45–70 dBZ	0–25 dBZ	25–45 dBZ	45–70 dBZ
U-Net	0.9168	0.4581	0.1613	0.0970	0.3054	0.4755	0.8758	0.3813	0.1470
Attention-U-Net	0.9280	0.4622	0.1979	0.0950	0.3535	0.4495	0.8836	0.3995	0.1798
MCDA-U-Net	0.9462	0.4878	0.1896	0.0927	0.3452	0.4335	0.8711	0.3880	0.1656
SegFormer	0.9422	0.4663	0.0629	0.0918	0.4190	0.5558	0.8546	0.3490	0.0583
SRViT	0.9274	0.3321	0.0090	0.1222	0.5544	0.7061	0.8214	0.2350	0.0089
CBAM-U-Net	**0.9690**	**0.5032**	**0.3477**	**0.0864**	**0.3023**	**0.4229**	**0.8877**	**0.4131**	**0.2495**

Table 7 further reveals the performance differences of models in handling local severe convection. For the severe echo interval (45–70 dBZ), which is the most critical for early warning, the CSI of CBAM-U-Net reaches 0.2495, far exceeding that of U-Net (0.1470) and Transformer-based models (<0.06). This demonstrates that under the background of rapid initiation and dissipation of convection in the pre-flood period, conventional CNNs struggle to capture transient high-frequency features, while Transformers suffer from the loss of strong echo centers due to the lack of local inductive bias. Only CBAM-U-Net effectively enhances the response capability to the core of strong echoes through the attention mechanism.

(2) Visual analysis of typical cases

To intuitively demonstrate the inversion performance of the models during the pre-flood active period, Figures 7 and 8 present a comparative analysis of a typical local convective event recorded at 20:19:17 (Beijing Time) on May 10, 2023.

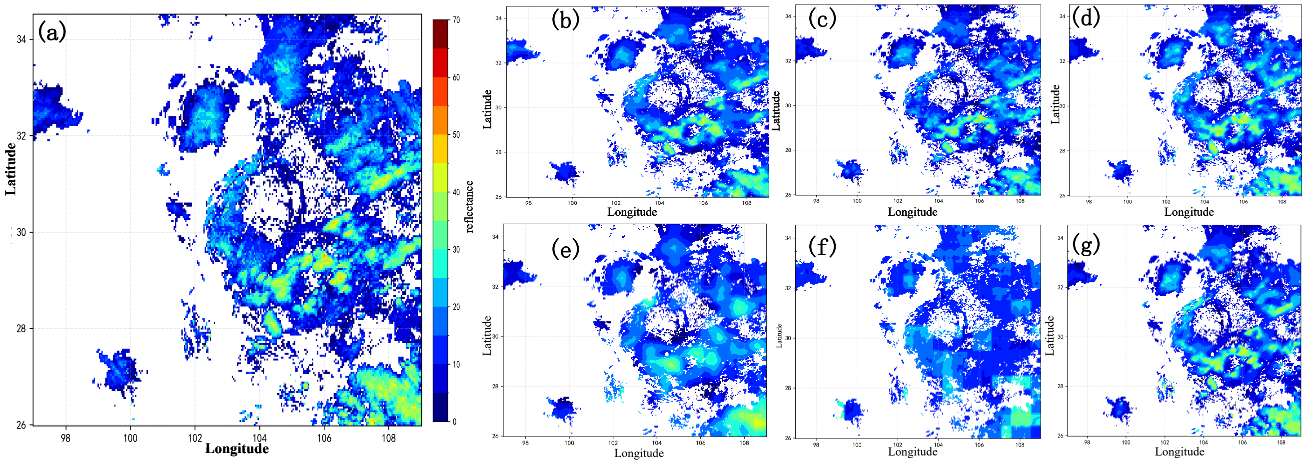

Figure 7. Panoramic overview of radar composite reflectivity at 20:19:17 on May 10, 2023: (a) Ground truth (GT) radar composite reflectivity; (b) U-Net; (c) Attention-U-Net; (d) MCDA-U-Net; (e) SegFormer; (f) SRViT; (g) CBAM-U-Net

https://doi.org/10.5194/egusphere-2026-2395
Preprint. Discussion started: 7 July 2026

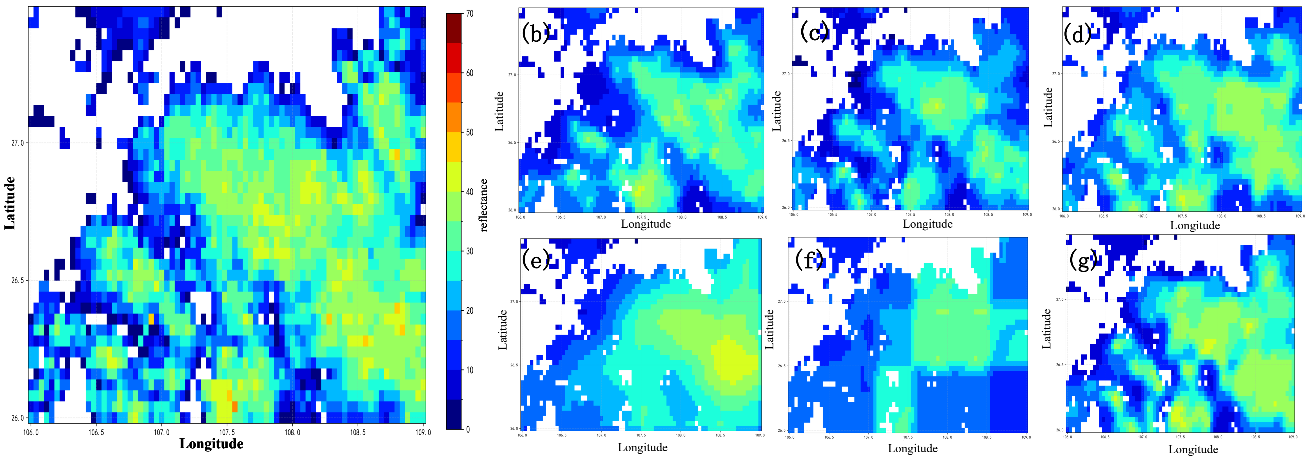

Figure 8. Local magnified view of radar composite reflectivity at 20:19:17 on May 10, 2023: (a) Ground truth (GT) radar composite reflectivity; (b) U-Net; (c) Attention-U-Net; (d) MCDA-U-Net; (e) SegFormer; (f) SRViT; (g) CBAM-U-Net

A joint analysis of the panoramic overview in Figure 7 and the local magnified details in Figure 8 reveals that the ground truth (GT) radar echo field at this timestamp exhibits a distinct local severe convection core, accompanied by scattered weak echo structures in the surrounding areas. In contrast, the inversion results of U-Net and Attention-U-Net are generally blurred: the boundaries of weak echoes are ill-defined, and the intensity of the strong echo core is significantly underestimated, a typical underestimation bias. Although MCDA-Unet slightly outperforms the aforementioned two models in terms of overall echo morphology, it still fails to capture subtle weak echo features and shallow texture details. Transformer-based models, such as SegFormer and SRViT, exhibit suboptimal performance, primarily recovering only the general outline of the echo field. This limitation is further exacerbated by the presence of pronounced blocky artifacts, which significantly compromise the retention of internal fine structural information. In contrast, the inversion results produced by CBAM-U-Net closely approximate ground truth observations: the model not only effectively restores the edge boundaries of scattered light rain echoes but also accurately identifies and enhances high-intensity regions within the core of severe convection. This demonstrates CBAM-U-Net's superior capabilities in detail restoration and structural preservation. In summary, within the context of the high data variability and pronounced dispersion characteristics inherent to the pre-flood onset phase, CBAM-U-Net—utilizing its lightweight and efficient attention mechanism—effectively mitigates the limitations associated with traditional models, which often manifest as "inaccurate predictions of strong echoes and ambiguous reconstructions of weak echoes." As a result, CBAM-U-Net demonstrates a noteworthy enhancement in performance.

4.2.2 The primary flood season (July-August)

During the primary flood season, the basin is predominantly influenced by the subtropical high-pressure system in July, resulting in elevated temperatures and minimal precipitation. Concurrently, the southwest monsoon transports moisture into the western Sichuan region, leading to localized rainstorm events. As the season progresses, the subtropical high begins to retreat, allowing for the infiltration of cold air, which subsequently triggers the onset of autumnal rainfall. Furthermore, the orographic lift associated with the basin's topography contributes to an extended duration of precipitation.

(1) Quantitative indicator analysis

The regression errors and hierarchical event verification results of each model in this stage are presented in detail in Tables 8 and9, respectively.

Table 8. Evaluation values of comparative models in the primary flood season

Model	RMSE	MAE	R^2	PSNR	SSIM
U-Net	6.9065	4.9072	0.6662	20.2435	0.7753
Attention-U-Net	6.8564	4.8041	0.6711	20.3076	0.7853
MCDA-U-Net	8.0035	5.9494	0.5384	20.8363	0.5384
SegFormer	9.1515	6.9567	0.3965	17.6722	0.6494
SRViT	11.5459	9.1874	0.2394	15.6534	0.5463
CBAM-U-Net	**6.6761**	**4.6490**	**0.6691**	**25.1057**	**0.7979**

As can be seen from Table 8, CBAM-U-Net demonstrates superior performance across all evaluated metrics in the comparative experiments conducted during the primary flood season, indicating its enhanced capabilities in inversion accuracy, error management, structural similarity, and image quality restoration relative to other models. U-Net and Attention-U-Net occupy the second tier in terms of overall performance, while MCDA-U-Net and SegFormer exhibit deficiencies in inversion accuracy and structural restoration capabilities. Among all models assessed, SRViT ranks the lowest, highlighting the challenges associated with achieving stable and high-precision radar echo inversion during the primary flood season, characterized by frequent severe convection events. In summary, the incorporation of the CBAM channel-spatial attention mechanism significantly mitigates inversion errors and improves feature fitting and structural preservation capabilities, rendering it a more dependable choice for practical applications under the complex meteorological conditions prevalent during the primary flood season.

Table 9. Threshold-based evaluation of comparative models during the primary flood season

Model	**POD**			**FAR**			**CSI**		
	0–25 dBZ	25–45 dBZ	45–70 dBZ	0–25 dBZ	25–45 dBZ	45–70 dBZ	0–25 dBZ	25–45 dBZ	45–70 dBZ
U-Net	0.9193	0.6858	0.4085	0.1100	0.2539	0.3172	0.8336	0.5560	0.3394
Attention-U-Net	0.9231	0.6668	0.4147	0.1116	0.2497	0.4002	0.8364	0.5510	0.3376
MCDA-U-Net	0.8899	0.6604	0.3915	0.1240	0.3446	0.3814	0.7904	0.4902	0.2471
SegFormer	0.8261	0.6885	0.1928	0.1209	0.4333	0.5358	0.7418	0.4511	0.1577
SRViT	0.7448	0.5703	0.0614	0.1820	0.5651	0.7115	0.6390	0.6390	0.0533
CBAM-U-Net	**0.9342**	**0.6914**	**0.4585**	**0.0973**	**0.2434**	**0.3347**	**0.8405**	**0.5646**	**0.3845**

https://doi.org/10.5194/egusphere-2026-2395
Preprint. Discussion started: 7 July 2026

As can be seen from the POD, FAR, and CSI metrics across different precipitation intensities in Table 9, CBAM-U-Net attains optimal performance across low, moderate, and heavy precipitation classifications, with the most pronounced enhancement observed within the heavy precipitation range (45–70 dBZ), which is particularly critical for early warning services. This model exhibits a higher detection rate and a lower false alarm rate, thereby surpassing other models in terms of its applicability and value for forecasting purposes.

(2) Visual analysis of typical cases

Meanwhile, to visually demonstrate the inversion effect of the model during the active period, a comprehensive evaluation based on test set samples during the primary flood season is conducted, Figures 9 and 10 present a visual image analysis of a typical case at 18:42:53 (Beijing Time) on August 6, 2023.

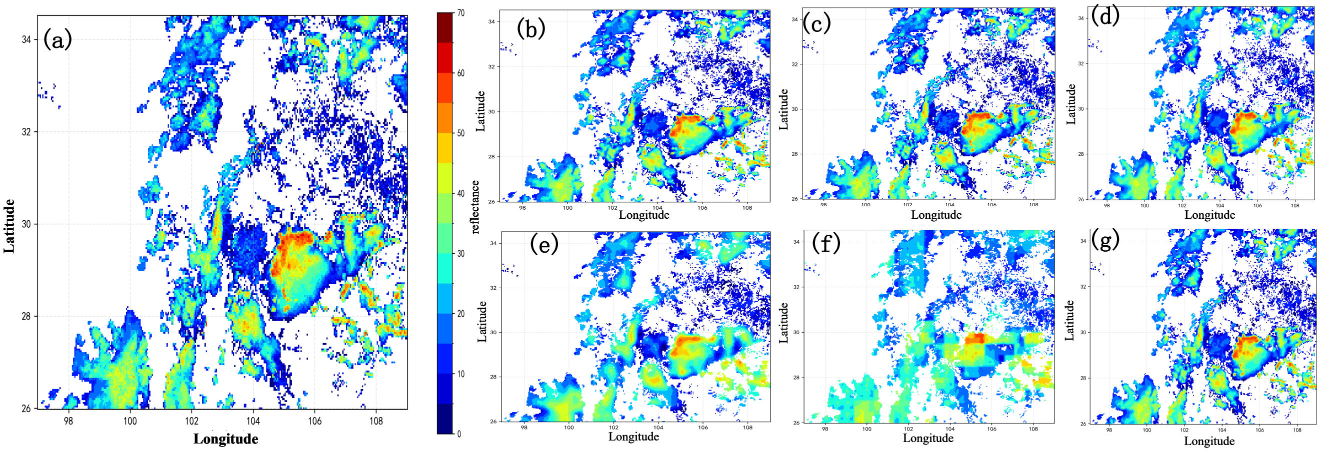

Figure 9. Radar combined reflectivity overview at 18:42:53 on August 6, 2023: (a) Real radar combined reflectivity; (b) U-Net; (c) Attention-U-Net; (d) MCDA-U-Net; (e) SegFormer; (f) SRViT; (g) CBAM-U-Net

https://doi.org/10.5194/egusphere-2026-2395
Preprint. Discussion started: 7 July 2026

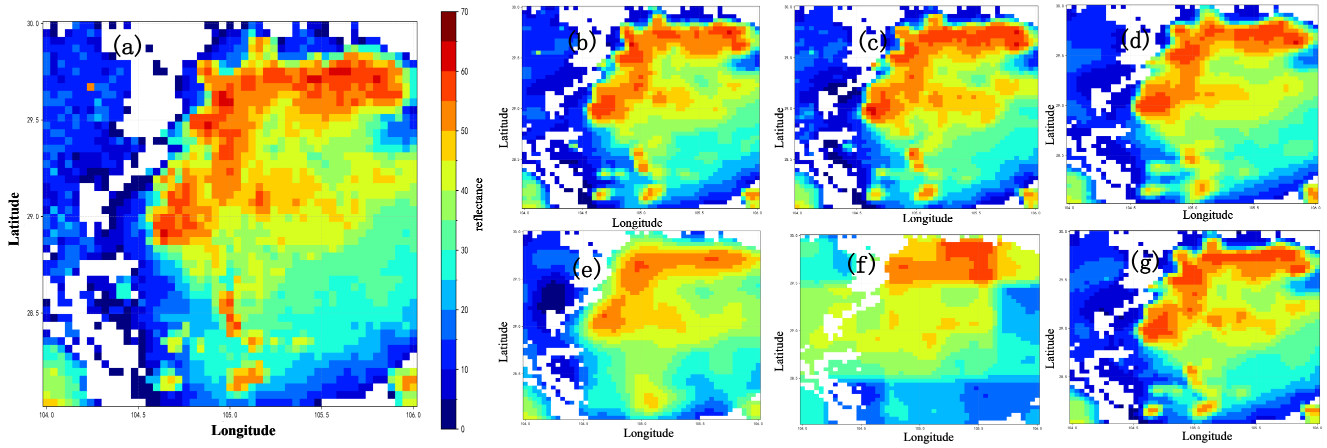

Figure 10. Radar combined reflectivity detailed view at 18:42:53 on August 6, 2023: (a) Real radar combined reflectivity; (b) U-Net; (c) Attention-U-Net; (d) MCDA-U-Net; (e) SegFormer; (f) SRViT; (g) CBAM-U-Net

From the perspective of macroscopic spatial distribution (Figure 9), The ground truth radar echo field reveals aggregated regions characterized by multi-core strong echoes. The CBAM-U-Net model effectively reproduces this macroscopic spatial pattern, demonstrating a high degree of alignment in both the orientation and extent of the echo band when compared to the ground truth. In contrast, while Attention-U-Net and U-Net are capable of capturing the primary distribution of echoes, they exhibit diminished continuity in the echo band, accompanied by fragmentation in localized areas. MCDA-U-Net and SegFormer demonstrate varying degrees of contraction in the distribution range of echoes, with the coverage of weak echo regions significantly lower than that of the ground truth. Notably, SRViT experiences the most pronounced macroscopic structural distortion, as evidenced by deviations not only in the orientation of the echo band but also in the spatial positioning of strong echo cores. From the perspective of detailed characteristics of local severe convection cores (Figure 10), the strong echo core (reflectivity $\geq$ 45 dBZ) marked by the red ellipse in the ground truth field features clear boundaries and concentrated intensity gradients. CBAM-U-Net proficiently restores both the intensity levels and spatial contours of the core region, with the aggregation characteristics of high-reflectivity pixels closely mirroring those of the ground truth field and exhibiting a natural gradient of intensity transitions. Attention-U-Net is capable of identifying the core location; however, it underestimates the intensity within the core region, resulting in significant boundary blurring. U-Net manifests an enlarged core region, leading to an increased proportion of false echo pixels. SegFormer and SRViT demonstrate markedly limited efficacy in accurately characterizing strong echo cores: SRViT nearly entirely loses the high-reflectivity features of the core region, only managing to restore the weak echo background, while MCDA-U-Net exhibits considerable intensity attenuation in the core region, accompanied by a substantial amount of spatial noise. In summary, CBAM-U-Net, which incorporates a dual channel-spatial attention mechanism, effectively addresses the limitations of the baseline U-Net in capturing local features while circumventing the structural distortion issues commonly associated with Transformer-based models in the inversion of small-scale meteorological targets. In the context of radar echo inversion for severe convective weather during the post-flood period, CBAM-U-Net achieves optimal restoration of both macroscopic distributions and localized details, thereby demonstrating its dual efficacy in enhancing the accuracy of meteorological analyses.

4.2.3 The post-flood transition phase (September to November)

During the post-flood transition phase in Western China, precipitation is predominantly characterized by stratiform cloud formations, exhibiting stable intensity and spatial extent. Subsequently, with only a minimal residual presence of warm and humid airflow, precipitation is primarily driven by "cold air-triggered" short-duration heavy rainfall events. The rapid oscillation between cold and warm air masses results in data instability. Nevertheless, the inversion capabilities of CBAM-U-Net demonstrate significant advantages over other models in terms of image accuracy and detection performance.

(1) Quantitative indicator analysis

To evaluate the model performance at this stage, Table 10 presents the regression errors of each model, while Table 11 illustrates the results of hierarchical event verification.

Table 10. Evaluation metrics of comparative models during the post-flood transition phase

Model	RMSE	MAE	R^2	PSNR	SSIM
U-Net	6.2935	4.6386	0.6291	20.9241	0.7618
Attention-U-Net	6.1694	4.5344	0.6482	21.1536	0.7769
MCDA-U-Net	6.7651	5.0422	0.5714	20.2964	0.7403
SegFormer	7.0974	5.3753	0.5283	19.8800	0.7098
SRViT	9.1217	7.1482	0.2208	17.7004	0.6309
CBAM-U-Net	**6.1294**	**4.4336**	**0.6431**	**26.0172**	**0.7935**

As can be seen from the quantitative indicators in Table 10, the radar echo inversion performance of various models exhibits significant variability during the post-flood transition phase. CBAM-U-Net attains optimal values across three critical metrics—Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Structural Similarity Index Measure (SSIM)—recording values of 6.1294, 4.4336, and 0.7935, respectively. These results indicate that the model demonstrates a superior ability to accurately replicate the actual echo distribution during the post-flood transition phase while preserving a higher degree of structural similarity. In summary, the integration of the joint channel-spatial attention mechanism within CBAM-U-Net renders it particularly well-suited for radar echo inversion tasks during the post-flood transition phase, showcasing distinct advantages in prediction accuracy, error mitigation, and the restoration of image structure.

Table 11. Threshold Evaluation of Comparative Models during the post-flood transition phase

Model	**POD**			**FAR**			**CSI**		
	0–25 dBZ	25–45 dBZ	45–70 dBZ	0–25 dBZ	25–45 dBZ	45–70 dBZ	0–25 dBZ	25–45 dBZ	45–70 dBZ
U-Net	0.9465	0.6388	0.2556	0.1070	0.2290	0.3718	0.8501	0.5369	0.2220
Attention-U-Net	0.9505	0.6473	0.3234	0.1035	0.2722	0.4358	0.8565	0.5512	0.2588
MCDA-U-Net	0.9365	0.6496	0.2189	0.1060	0.2815	0.4188	0.8348	0.5179	0.1891
SegFormer	0.9106	0.6587	0.1417	0.1051	0.3180	0.5458	0.8226	0.5039	0.1211
SRViT	0.8677	0.5345	0.0499	0.1453	0.4537	0.6780	0.7562	0.3702	0.0451
CBAM-U-Net	**0.9539**	**0.6797**	**0.3807**	**0.0865**	**0.2660**	**0.3825**	**0.8580**	**0.5618**	**0.2992**

As can be seen from Table 11, in weak echo regions (0–25 dBZ), All models demonstrate commendable simulation performance for the stratiform cloud background field, maintaining overall high Probability of Detection (POD) and low False Alarm Rate (FAR), thereby effectively mitigating weak signal noise. In contrast, the SRViT model exhibits inadequate sensitivity to weak signals, as evidenced by its low POD and Critical Success Index (CSI) values. Within the 25–45 dBZ transition zone, CBAM-U-Net continues to maintain a leading position, accurately representing the convection gradient variation with a POD of 0.6797 and a CSI of 0.5618. Although both SegFormer and MCDA-U-Net achieve comparable POD values, their elevated FAR adversely affects their CSI performance; SRViT experiences a significant decline in performance (POD = 0.5345, CSI = 0.3702), failing to adequately restore the convection transition process. The differences among the models are most pronounced in the 45–70 dBZ strong echo region, where CBAM-U-Net comprehensively leads with a POD of 0.3807, a FAR of 0.3825, and a CSI of 0.2992, thereby validating the enhanced feature extraction capability of the dual channel-spatial attention mechanism for severe convection cores. In contrast, SRViT's POD plummets to a mere 0.0499, with its CSI at 0.0451; both SegFormer and MCDA-U-Net exhibit POD values below 0.22, while the CSI for baseline U-Net and Attention-U-Net does not exceed 0.26. These findings underscore the critical role of the CBAM attention mechanism in overcoming the challenges associated with "intensity underestimation and insufficient recognition."

(2) Visual analysis of typical cases

For the comprehensive evaluation of model performance during the post-flood recession period based on test set samples, Figures 11 and 12 present a visual image analysis of a typical case at 14:23:35 (Beijing Time) on September 20, 2023.

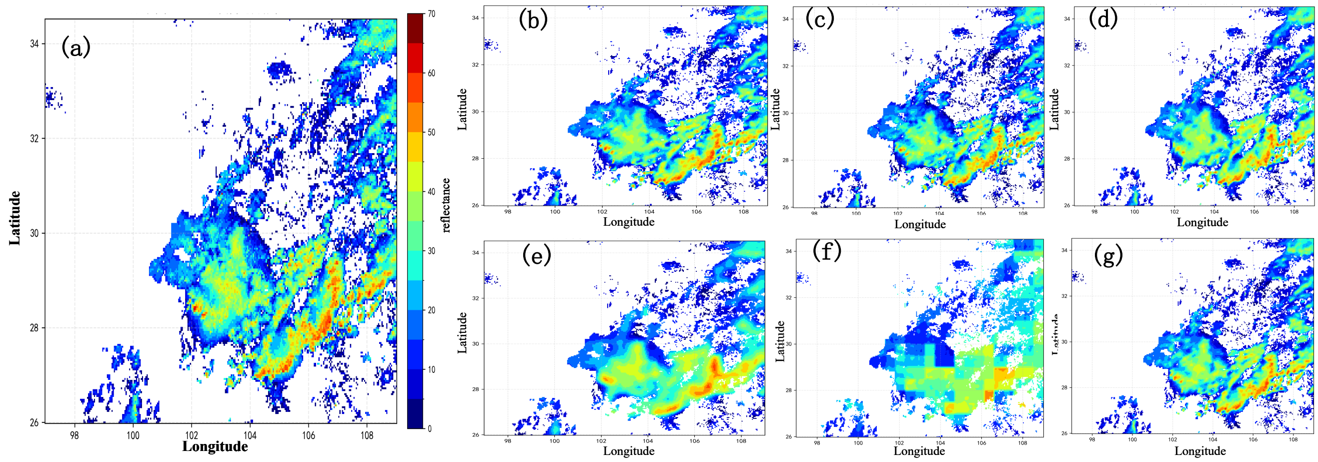

Figure 11. Panoramic view of radar echo inversion results at 14:23:35 (Beijing Time) on September 20, 2023: (a) Ground truth (GT); (b) U-Net; (c) Attention-U-Net; (d) MCDA-U-Net; (e) SegFormer; (f) SRViT; (g) CBAM-U-Net

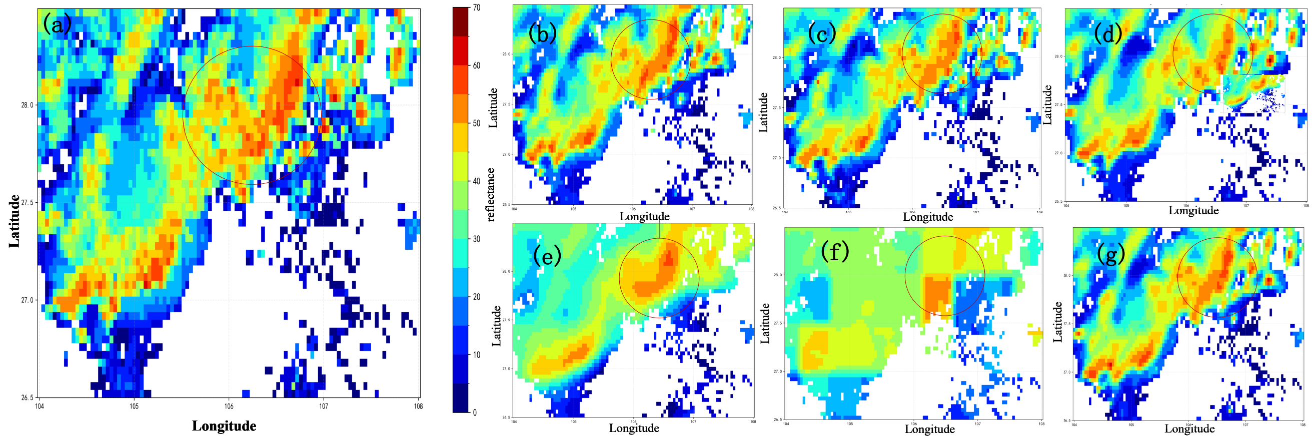

Figure 12. Local detailed view of radar echo inversion results at 14:23:35 (Beijing Time) on September 20, 2023: (a) Ground truth (GT); (b) U-Net; (c) Attention-U-Net; (d) MCDA-U-Net; (e) SegFormer; (f) SRViT; (g) CBAM-U-Net

A joint analysis of Figures 11 and 12 shows that significant disparities in model performance are observed concerning the spatial continuity of the weak echo background field (reflectivity < 10 dBZ). CBAM-U-Net demonstrates a high degree of consistency with the ground truth in its representation of weak echo distribution, exhibiting neither pronounced "hollowing effects" nor "excessive dispersion." In contrast, U-Net and Attention-U-Net display considerable signal loss within the background field, indicative of insufficient sensitivity to weak convective signals. Furthermore, MCDA-U-Net, SegFormer, and SRViT experience excessive dispersion in the background field, suggesting that their threshold control capabilities require further enhancement. These results confirm that the attention mechanism of the CBAM-U-Net model can effectively enhance the weight of severe convection features, suppress background noise, and improve the simulation accuracy of microphysical processes.

4.3 Model generalization test across satellite platforms

To objectively assess the model's robustness outside the distribution of training samples and its generalization capability in practical business scenarios, this study performed a cross-domain test utilizing FY-4B satellite data.

4.3.1 Experimental setup and data consistency

Both the FY-4A and FY-4B satellites are equipped with the Advanced Geostationary Radiation Imager (AGRI). While FY-4B integrates an additional low-level water vapor detection channel not present in FY-4A, the observation channels common to both satellites exhibit a high degree of congruence in terms of central wavelength, spectral response range, and spatial resolution, thereby demonstrating homologous physical detection characteristics. This alignment establishes a solid physical basis for validating the feasibility of transferring the model from FY-4A to FY-4B. For the experimental analysis, multi-channel observation data from FY-4B AGRI collected in July 2024 (a year distinct from that of the training dataset) and concurrent radar composite reflectivity (RCRF) were selected to construct an independent test set. To rigorously adhere to the standards of machine learning generalization testing and to prevent data leakage within the test set, several stringent control measures were implemented in this study. Firstly, to ensure consistency in input features, observation channels (Ch08-09, Ch11-13) and brightness temperature difference features (BTD1, BTD2) that correspond precisely to the inputs of the FY-4A model were selected. Secondly, standardized parameter freezing was enforced, ensuring that the preprocessing of FY-4B test data strictly followed the mean and variance derived from the statistical analysis of the FY-4A training dataset. Any use of statistical features from FY-4B's own data for calibration was explicitly prohibited. Finally, model parameter freezing was applied, whereby the model weights trained on the FY-4A dataset were directly utilized for inference without any fine-tuning on the FY-4B data, thereby allowing for an accurate assessment of the model's zero-shot transfer capability.

4.3.2 Visual qualitative analysis

Figure 13 presents a comparison of the inversion performance of different models on cross-satellite FY-4B data.

https://doi.org/10.5194/egusphere-2026-2395
Preprint. Discussion started: 7 July 2026

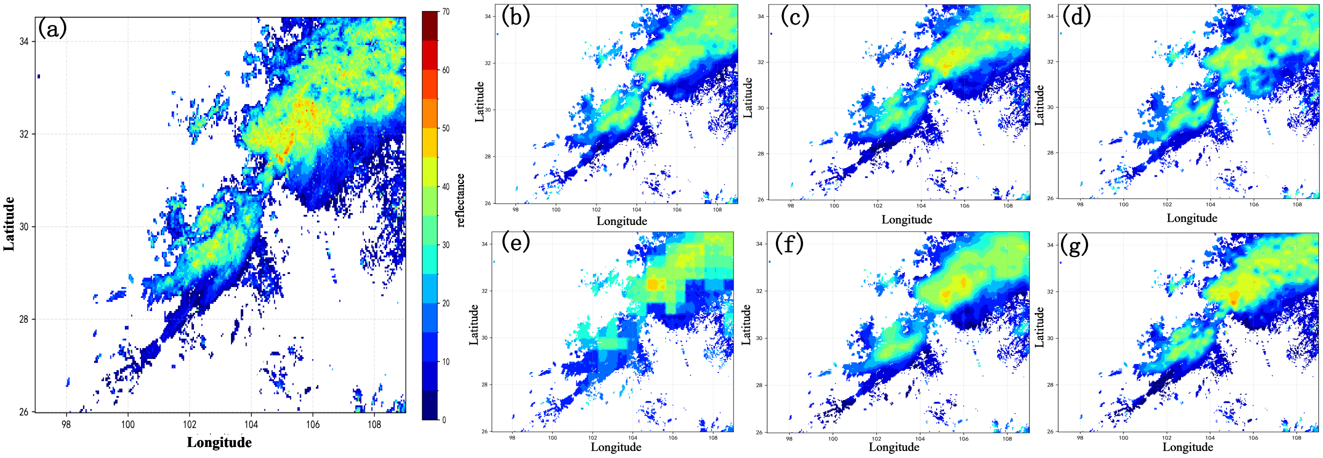

Figure 13. Comparison of inversion performance on FY-4B satellite data: (a) Ground truth radar composite reflectivity; (b) U-Net; (c) Attention-U-Net; (d) MCDA-Unet; (e) SegFormer; (f) SRViT; (g) CBAM-U-Net

In cross-domain scenarios, CNN-based models consistently demonstrate superior performance compared to Transformer-based baseline models. Among these, U-Net and its enhanced variants can approximately recover the primary morphological characteristics of radar echoes; however, they exhibit a pronounced tendency toward smoothing in response to minor shifts in data distribution resulting from changes in sensor observation sources (specifically, transitioning from FY-4A to FY-4B). This is particularly evident in the blurring of weak echo boundaries and a significant loss of information pertaining to the core of strong echoes. In contrast, CBAM-U-Net exhibits optimal structural preservation capabilities in cross-satellite evaluations. Leveraging the synergistic effects of channel response recalibration and spatial location aggregation facilitated by the CBAM module, this model is more adept at extracting robust features that are less sensitive to sensor variations. Consequently, it ensures that the inversion results maintain a high degree of consistency with radar ground truth in terms of shape, texture, and the structural integrity of severe convection cores.

4.3.3 Quantitative indicator analysis

Tables 12 and 13 present the overall regression metrics and threshold-based event verification metrics of the cross-satellite test, respectively.

Table 12. Cross-Satellite Test: Overall Regression Metrics of CBAM-U-Net and Comparative Models

Model	RMSE	MAE	R^2	PSNR	SSIM
U-Net	8.8405	6.7437	0.4273	17.9724	0.6900
Attention-U-Net	8.5707	6.4643	0.4617	18.2416	0.6990
MCDA-U-Net	8.6095	6.4725	0.4568	16.4717	0.6589
SegFormer	9.0401	7.0029	0.4011	17.7785	0.6748
SRViT	11.0072	8.8237	0.3122	16.0685	0.5989
CBAM-U-Net	**8.2141**	**6.0147**	**0.4918**	**23.0558**	**0.7562**

Table 13. Threshold-based event verification of CBAM-U-Net and comparative models in the cross-satellite FY-4B test

Model	**POD**			**FAR**			**CSI**		
	0–25 dBZ	25–45 dBZ	45–70 dBZ	0–25 dBZ	25–45 dBZ	45–70 dBZ	0–25 dBZ	25–45 dBZ	45–70 dBZ
U-Net	0.8655	0.6292	0.1833	0.1493	0.3772	0.5428	0.7514	0.4555	0.2758
Attention-U-Net	0.9139	0.5287	0.2474	0.1733	0.3233	0.6039	0.7670	0.4221	0.3203
MCDA-U-Net	0.8445	0.5114	0.2815	0.1659	0.4415	0.6478	0.7125	0.4112	0.2815
SegFormer	0.8588	0.6151	0.1978	0.1532	0.3919	0.6558	0.7433	0.4405	0.1824
SRViT	0.7625	0.5646	0.1251	0.1915	0.5279	0.8129	0.6458	0.3461	0.2226
CBAM-U-Net	**0.9258**	**0.6248**	**0.3316**	**0.1369**	**0.3482**	**0.5664**	**0.7690**	**0.4680**	**0.3723**

As shown in Table 12, CBAM-U-Net outperforms all other models across all evaluation dimensions: it achieves the lowest RMSE (8.2141) and MAE (6.0147), the highest coefficient of determination R^2 (0.4918), and significant advantages in image quality metrics (PSNR=23.0558, SSIM=0.7562). This indicates that CBAM-U-Net can maintain high inversion accuracy and physical consistency even when facing minor shifts in input data distribution.

5 Conclusions

Addressing the operational challenge of radar coverage blind spots over complex topography, this study developed an end-to-end deep learning retrieval framework that maps multi-channel radiometric signatures from the FY-4A/AGRI geostationary satellite to high-resolution Radar Composite Reflectivity (RCRF). Specifically, we proposed a lightweight, attention-augmented architecture—designated as CBAM-U-Net—which cascades channel and spatial attention mechanisms within a U-Net backbone to adaptively filter critical spectral responses and precisely localize intense convective cores. Evaluated on an extensive dataset from the 2023 extended flood season in the Sichuan region, CBAM-U-Net demonstrated significant enhancements in both retrieval accuracy and structural fidelity. It achieved the optimal pixel-wise error metrics (RMSE = 6.8290 dBZ, MAE = 4.8579 dBZ, R^2 = 0.6277) and superior spatial structure preservation (SSIM = 0.7894, PSNR = 25.2701 dB) compared to the baseline models. These results confirm that the dual-dimensional focusing mechanism effectively mitigates the boundary-blurring degradation typical of conventional models, striking an optimal balance among numerical accuracy, morphological fidelity, and structural consistency. Beyond overall statistical metrics, the proposed model exhibited a superior capability in delineating severe convective cores, which is critical for meteorological operations. Threshold-based verification revealed that CBAM-U-Net consistently outperformed baselines in the highly challenging strong echo region (45–70 dBZ), achieving the highest Probability of Detection (POD = 0.5296) and Critical Success Index (CSI = 0.4384) alongside the lowest False Alarm Ratio (FAR = 0.4357). This robust performance across the entire intensity spectrum ensures highly accurate localization of heavy precipitation centers, providing crucial diagnostic support for short-term severe weather warnings. Furthermore, the lightweight design of the CBAM module augments the network's representational capacity without incurring substantial computational overhead. Notably, during a cross-sensor, zero-shot generalization test utilizing FY-4B satellite data from July 2024, the model successfully mitigated the domain shift induced by sensor transitions, meticulously maintaining the consistency of echo morphology and confirming its immense potential for real-world operational deployment. Despite these promising results, further explorations are warranted to break through existing accuracy bottlenecks in severe convection retrieval. Future research will focus on multimodal feature integration by incorporating topographical data (e.g., Digital Elevation Models), cloud microphysical parameters, and Numerical Weather Prediction (NWP) background fields to construct a physically constrained fusion framework. Additionally, acknowledging the highly dynamic evolution of convective systems, incorporating recurrent or self-attention temporal modules (such as ConvLSTM or Spatiotemporal Transformers) will enforce temporal consistency and enhance the kinematic tracking of fast-moving storms. Finally, advancing the architectural paradigm by exploring CNN-Transformer hybrid networks with multi-scale cross-attention mechanisms could organically unify global structural context modeling with local fine-grained detail fidelity, further solidifying the operational viability of satellite-based radar echo reconstruction.

Data availability. The data used in this study are available upon reasonable request from the corresponding author.

Author contributions. Wen Kang conducted the experiments, analysed the results, and prepared the original manuscript. Hao Wang supervised the research and revised the manuscript. Qiangyu Zeng, Tiantian Yu, Jiafeng Zheng, and Zhi Li contributed to data processing, discussion of the results, and manuscript review.

Competing interests. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements. This work was sponsored by the National Natural Science Foundation of China (42575154, U2342216), the National Key R&D Program of China (2023YFC3007501), the Project of Sichuan Provincial Department of Science and Technology (2026NSF-SCZY0088), and the Key Laboratory of Atmosphere Sounding, China Meteorological Administration (2024KLAS06M).(Corresponding author: Hao Wang.))

References

Biswas, S. K. and Chandrasekar, V.: Cross validation of observations between the GPM dual-frequency precipitation radar and ground-based dual-polarization radars, Remote Sensing, 10, 1773, https://doi.org/10.3390/rs10111773, 2018.

Chen, P., Chen, L., Wang, G., Wu, Q., Wang, H., and Zhang, P.: Comparison of reflectivity consistency between spaceborne precipitation radar and ground-based weather radar in China and the United States, Advances in Atmospheric Sciences, 42, 1376–1394, 2025.

Chen, W., Hua, W., Ge, M., Su, F., Liu, N., Liu, Y., and Xiong, A.: Severe precipitation recognition using attention-UNet of multichannel doppler radar, Remote Sensing, 15, 1111, 2023.

Cheng, Y., Dai, T., Goto, D., Chen, L., Si, Y., Murakami, H., Yoshida, M., Zhang, P., Cao, J., Nakajima, T., et al.: Improved hourly estimate of aerosol optical thickness over Asian land by fusing geostationary satellites Fengyun-4B and Himawari-9, Science of The Total Environment, 923, 171 541, 2024.

Duan, M., Xia, J., Yan, Z., Han, L., Zhang, L., Xia, H., and Yu, S.: Reconstruction of the radar reflectivity of convective storms based on deep learning and Himawari-8 observations, Remote Sensing, 13, 3330, 2021a.

Duan, M., Xia, J., Yan, Z., Han, L., Zhang, L., Xia, H., and Yu, S.: Reconstruction of the Radar Reflectivity of Convective Storms Based on Deep Learning and Himawari-8 Observations, Remote Sensing, 13, 3330, https://doi.org/10.3390/rs13163330, 2021b.

Gao, Y., Wang, X., and Mao, D.: Performance of FY-4B GIIRS temperature products under cloudy skies and their enhancement of surface precipitation type forecasting, Atmospheric Research, 302, 107 305, 2024.

Gao, Y., Zeng, Q., Liu, Y., Zhang, F., Wang, H., and Ren, Z.: WTC-MobResNet: A Deep Learning Approach for Detecting Wind Turbine Clutter in Weather Radar Data, Remote Sensing, 17, 2763, 2025.

He, R., Li, H., Luo, J., Huang, H., and Zhu, Y.: Comparison of the reflectivities from precipitation measurement radar onboard the FY-3G satellite and ground-based S-Band dual-polarization radars, Remote Sensing, 17, 1117, 2025.

He, W., Xiong, T., Wang, H., He, J., Ren, X., Yan, Y., and Tan, L.: Radar echo spatiotemporal sequence prediction using an improved ConvGRU deep learning model, Atmosphere, 13, 88, 2022.

Hilburn, K. A., Ebert-Uphoff, I., and Miller, S. D.: Development and Interpretation of a Neural-Network-Based Synthetic Radar Reflectivity Estimator Using GOES-R Satellite Observations, Journal of Applied Meteorology and Climatology, 60, 3–21, https://doi.org/10.1175/JAMC-D-20-0084.1, 2021.

Ioffe, S. and Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift, in: International conference on machine learning, pp. 448–456, 2015.

Kou, L., Mao, Y., Wang, Z., Chen, Y., Chu, Z., and Chen, A.: Comparison of rainfall estimates from GPM dual-frequency precipitation radar and ground dual-polarization radar, Authorea Preprints, 2022.

Li, Q., Bai, X., Hu, L., Li, L., Bao, Y., Geng, X., and Yan, X.-H.: Semantic Segmentation of Typical Oceanic and Atmospheric Phenomena in SAR Images Based on Modified Segformer, Remote Sensing, 18, 113, 2025.

Liu, Y., Wen, J., Zheng, J., and Wang, H.: Vertical Structures and Macro-Microphysical Characteristics of Southwest Vortex Precipitation over Sichuan, China, Remote Sensing, 18, 533, 2026.

Roberts, R. D. and Rutledge, S.: Nowcasting Storm Initiation and Growth Using GOES-8 and WSR-88D Data, Weather and Forecasting, 18, 562–584, https://doi.org/10.1175/1520-0434(2003)018<0562:NSIAGU>2.0.CO;2, 2003.

Ronneberger, O., Fischer, P., and Brox, T.: U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, pp. 234–241, 2015.

Shen, F., Shu, A., Liu, Z., Li, H., Jiang, L., Zhang, T., and Xu, D.: Assimilating FY-4A AGRI radiances with a channel-sensitive cloud detection scheme for the analysis and forecasting of multiple typhoons, Advances in Atmospheric Sciences, 41, 937–958, 2024.

Stock, J., Hilburn, K., Ebert-Uphoff, I., and Anderson, C.: SRViT: Vision transformers for estimating radar reflectivity from satellite observations at scale, arXiv preprint arXiv:2406.16955, 2024.

Sun, F., Li, B., Min, M., and Qin, D.: Deep Learning-Based Radar Composite Reflectivity Factor Estimations from Fengyun-4A Geostationary Satellite Observations, Remote Sensing, 13, 2229, https://doi.org/10.3390/rs13112229, 2021a.

Sun, F., Li, B., Min, M., and Qin, D.: Deep learning-based radar composite reflectivity factor estimations from Fengyun-4A geostationary satellite observations, Remote Sensing, 13, 2229, https://doi.org/10.3390/rs13112229, 2021b.

Veillette, M., Samsi, S., and Mattioli, C.: SEVIR: A storm event imagery dataset for deep learning applications in radar and satellite meteorology, Advances in neural information processing systems, 33, 22 009–22 019, 2020.

Veillette, M. S., Hassey, E. P., Mattioli, C. J., Iskenderian, H., and Lamey, P. M.: Creating Synthetic Radar Imagery Using Convolutional Neural Networks, Journal of Atmospheric and Oceanic Technology, 35, 2323–2338, https://doi.org/10.1175/JTECH-D-18-0010.1, 2018.

Wang, C., Xu, J., Tang, G., Yang, Y., and Hong, Y.: Infrared Precipitation Estimation Using Convolutional Neural Network, IEEE Transactions on Geoscience and Remote Sensing, 58, 8612–8625, https://doi.org/10.1109/TGRS.2020.2989183, 2020a.

Wang, H., Wei, M., Li, G., Zhou, S., and Zeng, Q.: Analysis of precipitable water vapor from GPS measurements in Chengdu region: Distribution and evolution characteristics in autumn, Advances in Space Research, 52, 656–667, 2013.

Wang, H., Wang, L., He, J., Ge, F., Chen, Q., Tang, S., and Yao, S.: Can the GPM IMERG hourly products replicate the variation in precipitation during the wet season over the Sichuan Basin, China?, Earth and Space Science, 7, e2020EA001 090, 2020b.

Wang, H., Tan, L., Zhang, F., Zheng, J., Liu, Y., Zeng, Q., Yan, Y., Ren, X., and Xiang, J.: Three-dimensional structure analysis and droplet spectrum characteristics of Southwest Vortex precipitation system based on GPM-DPR, Remote Sensing, 14, 4063, 2022a.

Wang, H., Yan, Y., Long, K., Chen, Q., Fan, X., Zhang, F., and Tan, L.: Relationships Between Rapid Urbanization and Extreme Summer Precipitation Over the Sichuan–Chongqing Area of China, Frontiers in Earth Science, 10, https://doi.org/10.3389/feart.2022.909547, 2022b.

Wang, H., Li, Z., Zhang, T., Chen, Q., Guo, X., Zeng, Q., and Xiang, J.: Downscaling of GPM satellite precipitation products based on machine learning method in complex terrain and limited observation area, Advances in Space Research, 72, 2226–2244, 2023.

Wang, H., Yang, R., He, J., Zeng, Q., Xiong, T., Liu, Z., and Jin, H.: Enhancing Precipitation Nowcasting Through Dual-Attention RNN: Integrating Satellite Infrared and Radar VIL Data, Remote Sensing, 17, 238, https://doi.org/10.3390/rs17020238, 2025.

Wang, H., Liu, Z., Zhang, F., Li, Z., Zeng, Q., Yu, T., and Yang, R.: A CBAM-Integrated Deep Convolutional Neural Network for Weather Radar Echo Correction to Overcome Terrain Occlusion Challenge, Advances in Space Research, 2026.

Woo, S., Park, J., Lee, J.-Y., and Kweon, I. S.: CBAM: Convolutional block attention module, in: Proceedings of the European conference on computer vision (ECCV), pp. 3–19, 2018.

Yan, Y., Wang, H., Li, G., Xia, J., Ge, F., Zeng, Q., Ren, X., and Tan, L.: Projection of future extreme precipitation in China based on the CMIP6 from a machine learning perspective, Remote Sensing, 14, 4033, 2022.

Yang, L., Zhao, Q., Xue, Y., Sun, F., Li, J., Zhen, X., and Lu, T.: Radar Composite Reflectivity Reconstruction Based on FY-4A Using Deep Learning, Sensors, 23, 81, https://doi.org/10.3390/s23010081, 2023.

Yang, R., Wang, H., Zhang, F., Zeng, Q., Xiong, T., Liu, Z., and Jin, H.: Enhancing severe weather predictions with the I-ConvGRU model: An iterative approach for radar echo time series through ConvGRU and RainNet integration, Journal of Hydroinformatics, 26, 2197–2215, 2024.

Yongguang, Z., Kanghui, Z., Jie, S., Yinjing, L., Fuyou, T., Wenyuan, T., Yu, L., and Wenjian, Z.: Advances in techniques of monitoring, forecasting and warning of severe convective weather, Journal of Applied Meteorological Science, 26, 641–657, 2015.

Yu, Q., Zhu, M., Zeng, Q., Wang, H., Chen, Q., Fu, X., and Qing, Z.: Weather radar super-resolution reconstruction based on residual attention back-projection network, Remote Sensing, 15, 1999, 2023a.

Yu, X., Lou, X., Yan, Y., Yan, Z., Cheng, W., Wang, Z., Zhao, D., and Xia, J.: Radar echo reconstruction in oceanic area via deep learning of satellite data, Remote Sensing, 15, 3065, 2023b.

Zeng, Q., Li, L., Wang, H., He, J., Wang, H., and Gao, Y.: MCDA-UNet: A satellite data-based model for radar composite reflectivity retrieval, Atmospheric Research, p. 108619, 2025a.

Zeng, Q., Zhang, G., Huang, S., Song, W., He, J., Wang, H., and Liu, Y.: A novel tornado detection algorithm based on XGBoost, Remote Sensing, 17, 167, 2025b.

Zhang, P., Zhu, L., Tang, S., Gao, L., Chen, L., Zheng, W., Han, X., Chen, J., and Shao, J.: General comparison of FY-4A/AGRI with other GEO/LEO instruments and its potential and challenges in non-meteorological applications, Frontiers in Earth Science, 6, 224, 2019.