WaveC2R: Wavelet-Driven Coarse-to-Refined Hierarchical Learning for Radar Retrieval

Chunlei Shi[1], Han Xu[1], Yinghao Li[2], Yi-Lin Wei[2], Yongchao Feng[3], Yecheng Zhang[4], Dan Niu[1,5,6,*]

[1]School of Automation, Southeast University
[2]School of Computer Science and Engineering, Sun Yat-sen University
[3]State Key Laboratory of Virtual Reality Technology and Systems, Beihang University
[4]School of Architecture, Tsinghua University
[5]Heavy Rainfall Research Center of China
[6]China Meteorological Administration Xiong'an Atmospheric Boundary Layer KeyLaboratory
danniu1@163.com, 230238514@seu.edu.cn

Abstract

Satellite-based radar retrieval methods are widely employed to fill coverage gaps in ground-based radar systems, especially in remote areas affected by terrain blockage and limited detection range. Existing methods predominantly rely on overly simplistic spatial-domain architectures constructed from a single data source, limiting their ability to accurately capture complex precipitation patterns and sharply defined meteorological boundaries. To address these limitations, we propose WaveC2R, a novel wavelet-driven coarse-to-refined framework for radar retrieval. WaveC2R integrates complementary multi-source data and leverages frequency-domain decomposition to separately model low-frequency components for capturing precipitation patterns and high-frequency components for delineating sharply defined meteorological boundaries. Specifically, WaveC2R consists of two stages *(i) Intensity-Boundary Decoupled Learning*, which leverages wavelet decomposition and frequency-specific loss functions to separately optimize low-frequency intensity and high-frequency boundaries; and *(ii) Detail-Enhanced Diffusion Refinement*, which employs frequency-aware conditional priors and multi-source data to progressively enhance fine-scale precipitation structures while preserving coarse-scale meteorological consistency. Experimental results on the publicly available SEVIR dataset demonstrate that WaveC2R achieves state-of-the-art performance in satellite-based radar retrieval, particularly excelling at preserving high-intensity precipitation features and sharply defined meteorological boundaries.

1 Introduction

Satellite-based radar retrieval plays a critical role in filling coverage gaps of ground-based radar systems, providing essential precipitation monitoring in data-sparse regions (Park et al. 2025; Wen et al. 2024). Geostationary satellites such as Himawari-8, GOES-R, and Fengyun-4A provide high-resolution meteorological observations (Liu et al. 2020; Jing et al. 2022) for radar composite reflectivity reconstruction. These satellites capture visible, infrared, and lightning data, which collectively enhance characterization of complex precipitation patterns in regions lacking ground-based radar coverage (Li et al. 2023; Yang et al. 2023a).

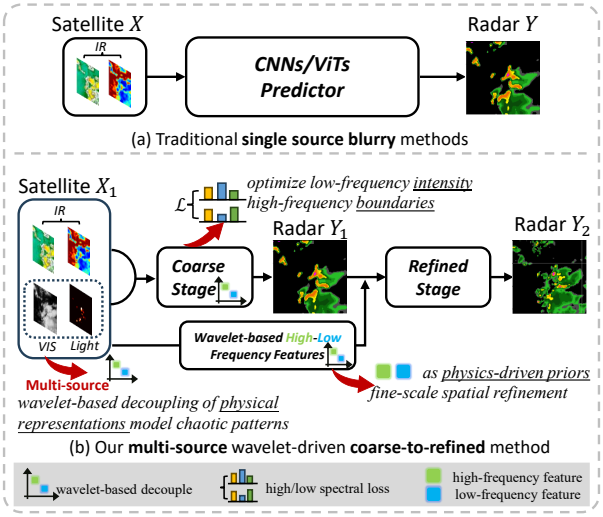

Figure 1: Comparison between traditional and proposed radar retrieval approaches. (a) Traditional methods employ spatial domain architectures with single-source infrared data (IR). (b) Our WaveC2R integrates multi-source satellite observations (visible, infrared, and lightning) through coarse-to-refined stages: initial estimation followed by conditional diffusion refinement with physics-driven priors.

Traditional retrieval methods, such as radiative transfer models (RTMs) (Di et al. 2018) and variational methods (Yin et al. 2021), rely heavily on physical models and statistical techniques. These methods are computationally intensive and often constrained in accuracy due to various factors. In recent years, deep learning methods have become a hot topic in radar retrieval research due to their powerful data processing capabilities and efficient predictive performance (Jiang et al. 2024; Park et al. 2025). However, despite the notable progress in deep learning-based satellite radar retrieval, current methods still suffer from several critical limitations that restrict their performance (Fig.1).

*Corresponding author.

Single-source data limitations. Most current approaches rely solely on single-source satellite data, typically infrared channels. This neglects complementary information from alternative data sources. Visible observations provide crucial cloud-top morphological details during daylight hours that enhance structural identification. Similarly, lightning data serves as a direct indicator of convective intensity and enables precise storm core localization (Yu et al. 2025; Zheng et al. 2024; Zhao et al. 2024b; Chen et al. 2024). Without these multi-source inputs, current methods struggle with incomplete scene representation, limiting their ability to accurately characterize severe convective cells (Niu et al. 2024).

Frequency domain optimization challenges. Accurate characterization of convective-scale structures requires effective modeling of meteorological systems' multi-frequency characteristics (Jin, Lin, and Xu 2023). However, existing satellite-based radar retrieval methods predominantly employ spatial domain optimization with simplistic architectures and L2 loss functions (Zhao et al. 2024a). This approach fails to capture the inherent frequency characteristics of weather phenomena, resulting in oversmoothed gradients, blurred precipitation boundaries, and underestimated high-intensity convective cores (Yang et al. 2023b).

Limited diffusion model exploration in radar retrival. Diffusion models offer significant potential for satellite-based radar retrieval through their ability to model complex probability distributions and progressively refine meteorological structures. Despite proven success in various image generation tasks, diffusion models remain unexplored in satellite-based radar retrieval. Traditional approaches struggle with recovering high-frequency meteorological details, particularly in convective regions with steep gradients (Gong et al. 2024; Wen et al. 2024; Li et al. 2024).

To address these fundamental limitations, we propose WaveC2R, a novel wavelet-driven coarse-to-refined hierarchical learning framework for satellite-to-radar retrieval. Our approach first integrates complementary multi-source satellite observations (VIS, IR, Light) through hierarchical frequency-domain feature decoupling and fusion, enabling comprehensive capture of complex meteorological patterns across temporal-frequency dual domains. Based on our key insight that low-frequency wavelet components primarily encode precipitation intensity distribution while high-frequency components capture sharp meteorological boundaries (see Fig. 2), WaveC2R decomposes the radar retrieval task into separate optimization processes for intensity prediction and boundary localization. Specifically, WaveC2R consists of two stages: the coarse stage employs a Wavelet-Temporal-Frequency (WTF) module with our Frequency-decomposed Intensity-Boundary Loss (FIBL) that separately constrains low-frequency intensity patterns and high-frequency boundary components, while the refined stage pioneers conditional diffusion models with physics-aware frequency-decomposed priors for progressive detail enhancement. By explicitly modeling the distinct physical roles of wavelet frequency components, WaveC2R provides more accurate radar retrieval than existing methods, particularly excelling in preserving high-intensity precipitation structures and sharp meteorological boundaries.

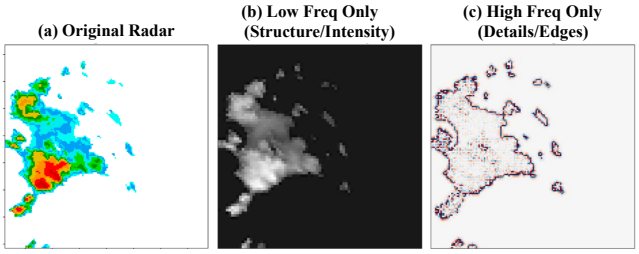

Figure 2: Wavelet-based frequency decomposition of radar reflectivity via DWT and selective IDWT reconstruction reveals distinct physical roles: (a) Original radar data, (b) Low-frequency reconstruction encodes precipitation intensity distribution, (c) High-frequency reconstruction captures boundary structures and meteorological details.

- We propose a temporal-frequency dual-domain fusion framework that integrates complementary multi-source satellite observations and establishes physical interpretability of wavelet decomposition, revealing that low-frequency components encode intensity distribution while high-frequency components capture spatial boundaries for meteorological structure modeling.
- We introduce the wavelet-temporal-frequency module that performs hierarchical frequency-domain decomposition with our frequency-decomposed intensity-boundary loss, which disentangles intensity patterns from boundary components and mitigates energy imbalance inherent in conventional spatial-domain optimization.
- We design a physics-aware conditional diffusion framework for satellite-based radar retrieval that leverages frequency-decomposed priors to progressively refine meteorological details while preserving coarse-scale structure coherence through counterfactual feature enhancement.

2 Related Work

2.1 Radar Retrieval Using Satellite Observations

Deep learning has emerged as a transformative approach for radar composite reflectivity (CREF) reconstruction by leveraging satellite observations. UNet-based models have become predominant due to their encoder-decoder structure and multi-scale feature extraction capabilities, with substantial improvements demonstrated through multi-band satellite integration (Jin, Lin, and Xu 2023; Zhao et al. 2024a). Alternative architectural innovations have emerged to address UNet limitations, including residual deep forest models (Bao et al. 2024) and integrated frameworks combining multi-band infrared data (Wang et al. 2020). Recent investigations have highlighted the importance of integrating diverse satellite observations encompassing visible, infrared, and lightning data to capture spatiotemporal information (Si et al. 2023; Yu et al. 2025; Zheng et al. 2024; Zhao et al. 2024b; Chen et al. 2024). However, existing methodologies continue to face challenges in effectively integrating multi-channel data, often relying on simple concatenation strategies that fail to exploit the complementary strengths of different satellite observations.

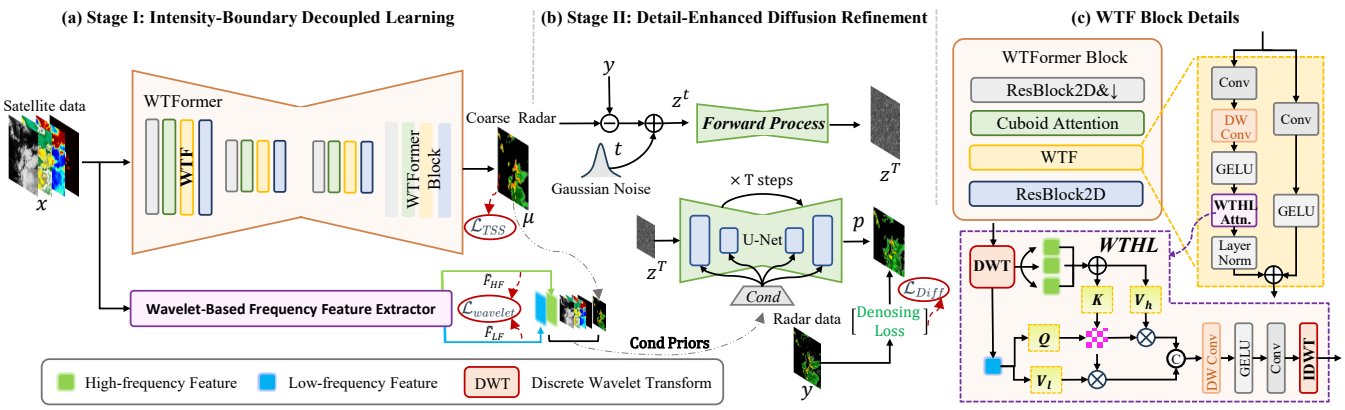

Figure 3: Overall architecture of the proposed WaveC2R framework. (a) *Stage I: Intensity-Boundary Decoupled Learning* employs the Wavelet-Temporal-Frequency (WTF) module to extract hierarchical frequency-domain features from multi-source satellite observations and generates coarse radar estimates through frequency-decomposed optimization. (b) *Stage II: Detail-Enhanced Diffusion Refinement* progressively refines coarse estimates via conditional diffusion with physics-aware frequency-decomposed priors. (c) *WTF Block* demonstrates the core wavelet-based attention mechanism that performs temporal-frequency feature fusion across multiple scales.

2.2 Frequency-Domain and Wavelet Transform in Meteorological Applications

Wavelet transform has established itself as a powerful technique for multi-scale meteorological analysis, offering distinctive advantages over traditional Fourier methods when processing non-stationary weather signals (Saoud, Al-Marzouqi, and Deriche 2021). The integration of frequency-domain techniques with deep learning has yielded promising results, with improvements achieved through Fast Fourier Transform integration in weather forecasting frameworks (Li et al. 2025) and specialized frequency-domain learning methods (Wang et al. 2024). The multi-resolution capabilities of wavelet transforms have proven particularly valuable for precipitation systems, with enhanced predictions resulting from combining wavelet transforms with time series models (Ahn and Hur 2023; Mammedov, Olugu, and Farah 2022) and frequency domain feature extraction fusion mechanisms (Liao, Gao, and Li 2024). However, existing wavelet-based approaches in meteorology primarily focus on mathematical decomposition without exploring the distinct physical roles of different frequency components in precipitation systems.

In this work, we approach radar retrieval from two disentangled frequency-domain perspectives: the intensity distribution (low-frequency) and spatial boundaries (high-frequency) of precipitation systems, distinguishing it from previous approaches that treat frequency components as unified mathematical decompositions.

3 Method

3.1 Problem Definition

We formulate satellite-based radar retrieval as a cross-modal reconstruction problem. Given synchronous multi-source satellite observations $\mathbf{X} = \{X_{\text{vis}}, X_{\text{ir}}, X_{\text{light}}\}$ comprising visible, infrared (dual-channel), and lightning data with total channels $C = 4$, our objective is to retrieve the corresponding radar reflectivity field $\hat{\mathbf{Y}} \in \mathbb{R}^{B\times 1\times H\times W}$ that accurately approximates ground truth radar observations $\mathbf{Y} \in \mathbb{R}^{B\times 1\times H\times W}$.

This retrieval problem is formulated as:

$$\hat{\mathbf{Y}} = f_\theta(\mathbf{X}), \quad \theta^* = \arg\min_\theta \mathbb{E}[\mathcal{L}(f_\theta(\mathbf{X}), \mathbf{Y})], \tag{1}$$

where f_θ represents a learnable retrieval mapping function, and $\mathcal{L}$ is a physics-informed loss incorporating meteorological constraints.

3.2 Preliminaries: Wavelet Transform

Given a 2D feature $\mathbf{F} \in \mathbb{R}^{H\times W}$, the discrete wavelet transform (DWT) decomposes it into four frequency sub-bands through separable 1D transforms applied along rows and columns:

$$\mathbf{F}_{ll}, \mathbf{F}_{lh}, \mathbf{F}_{hl}, \mathbf{F}_{hh} = \text{DWT}(\mathbf{F}), \tag{2}$$

where $\mathbf{F}_{ll} \in \mathbb{R}^{H/2\times W/2}$ represents the low-frequency approximation preserving primary structural features, while $\{\mathbf{F}_{lh}, \mathbf{F}_{hl}, \mathbf{F}_{hh}\}$ capture horizontal, vertical, and diagonal high-frequency details including boundaries and fine-grained meteorological structures.

The DWT operation is implemented through convolution with wavelet basis functions $f_l = \phi$ (low-pass) and $f_h = \psi$ (high-pass):

$$\mathbf{F}_{ab}[h, w] = \sum_{i,j} \mathbf{F}[i, j] \cdot f_a[i - 2h] \cdot f_b[j - 2w], \tag{3}$$

where $a, b \in \{l, h\}$, $h \in [0, H/2 - 1]$ and $w \in [0, W/2 - 1]$ represent the spatial coordinates in the downsampled wavelet domain. The inverse transform IDWT reconstructs the original feature from all four sub-bands, providing the foundation for frequency-domain feature manipulation in meteorological applications.

3.3 WaveC2R Framework

We propose WaveC2R, a two-stage coarse-to-refined framework for satellite-based radar retrieval that addresses the coupled intensity-boundary optimization challenge through physics-aware frequency-domain decomposition. The architecture, illustrated in Fig. 3, follows a sequential mapping: $\mathbf{X} \xrightarrow{\text{Stage I}} \boldsymbol{\mu} \xrightarrow{\text{Stage II}} \hat{\mathbf{Y}}$. Our key insight is that precipitation intensity manifests primarily in low-frequency wavelet components, while precipitation boundaries are captured in high-frequency details, enabling separate optimization objectives for intensity prediction and boundary localization. This framework operates through two complementary stages: coarse structure establishment via frequency-domain decomposition, followed by diffusion-based detail refinement.

Intensity-Boundary Decoupled Learning. We establish intensity-boundary decoupling through wavelet decomposition, recognizing that precipitation intensity and boundary localization exhibit distinct frequency characteristics. This is realized via two key components: the Wavelet-Temporal-Frequency (WTF) module for hierarchical multi-source feature fusion, and the Frequency-decomposed Intensity-Boundary Loss (FIBL) that separately constrains low-frequency intensity patterns and high-frequency boundary components. The WTF module extracts frequency-specific features from satellite inputs, then separate decoder branches independently reconstruct intensity and boundary components under FIBL supervision, generating the coarse radar estimation $\boldsymbol{\mu}$ with physically-consistent modeling (see Fig. 3(a)).

Detail-Enhanced Diffusion Refinement. To further enhance fine-scale meteorological structures while maintaining physical consistency, this refinement stage takes the coarse radar estimation $\boldsymbol{\mu}$ from Stage I and progressively refines it into the final high-fidelity prediction $\hat{\mathbf{Y}}$ through conditional diffusion modeling. The diffusion model employs frequency-aware conditional priors to progressively enhance fine-scale precipitation patterns while preserving coarse-scale meteorological consistency through iterative denoising guided by wavelet-decomposed features. This stage produces the final radar estimate with enhanced boundary precision and preserved meteorological structure (see Fig. 3(b)).

3.4 Stage I: Intensity-Boundary Decoupled Learning

Existing methods jointly optimize precipitation intensity and boundary characteristics, leading to oversmoothed results in spatial domain. We discover that intensity distribution and boundary localization exhibit distinct frequency signatures: intensity manifests in low-frequency wavelet components while boundaries concentrate in high-frequency details (Fig. 2). Based on this insight, we establish frequency-domain decoupling that separates these optimization objectives, enabling specialized learning strategies for accurate precipitation structure reconstruction.

WTFormer Architecture. In the coarse retrieval stage, we employ a WTFormer architecture to generate the coarse radar estimate $\boldsymbol{\mu}$ from multi-source satellite observations $\mathbf{X} \in \mathbb{R}^{B\times C\times H\times W}$. Drawing inspiration from successful encoder-decoder architectures in atmospheric science applications, WTFormer adopts a hierarchical U-Net structure with two encoding levels and corresponding decoding levels. Each level strategically integrates three key components: ResBlock2D for multi-scale down/upsampling, Cuboid Attention mechanism from Earthformer for capturing spatiotemporal dependencies, and our proposed WTF module for frequency-aware enhancement. The encoder $\mathcal{E}(\cdot)$ progressively extracts multi-scale features through these integrated components, while the symmetric decoder $\mathcal{D}(\cdot)$ reconstructs the coarse estimate via skip connections. The overall architecture can be formulated as:

$$\boldsymbol{\mu} = \text{WTFormer}(\mathbf{X}) = \mathcal{D}(\mathcal{E}(\mathbf{X})) \tag{4}$$

To capture both intensity distribution and boundary characteristics, WTFormer generates a coarse radar estimate $\boldsymbol{\mu}$ through progressive optimization with FIBL loss. Detailed architectural specifications are provided in Appendix A.1.

WTF Block. Given an input feature map $\mathbf{F}_{\text{in}}$, the WTF Block follows a dual-branch structure (see Fig. 3(c)): (1) a convolutional pathway with channel expansion, depthwise convolutions, and nonlinear activations, and (2) a wavelet-based pathway that processes features through the Wavelet Transform High-Low (WTHL) attention module.

In the WTHL module, discrete wavelet transform (DWT) decomposes input features into low-frequency approximation ($\mathbf{F}_{\text{LF}}$) and aggregated high-frequency details ($\mathbf{F}_{\text{HF}}$). These components are processed through a cross-frequency attention mechanism where low-frequency features serve as queries ($\mathbf{Q} = \mathbf{F}_{\text{LF}}\mathbf{W}_Q$) to attend to high-frequency keys ($\mathbf{K} = \mathbf{F}_{\text{HF}}\mathbf{W}_K$), computing attention weights via $\text{softmax}(\mathbf{Q}\mathbf{K}^T/\sqrt{d})$ and generating weighted outputs for both frequency domains. The refined features are then reconstructed using IDWT:

$$\hat{\mathbf{F}}_{\text{WTHL}} = \text{IDWT}(\text{Conv}(\text{GELU}(\mathbf{A}_h + \mathbf{A}_l))) \tag{5}$$

The dual pathways are combined through residual connections, enabling the WTF Block to capture both conventional spatial features and frequency-domain interactions for enhanced multi-source satellite data fusion.

FIBL and Temperature Scheduling Strategy Traditional spatial-domain losses treat all frequency components equally, leading to energy imbalance where low-frequency intensity patterns dominate optimization while high-frequency boundary details are under-represented. Although Fourier Global Loss (FGL)(Yan et al. 2024) effectively captures global motion trends, it struggles with local positioning accuracy due to its global frequency spectrum operation, making it insufficient for capturing localized high-frequency structures that define precipitation boundaries.

To address these limitations, we propose FIBL, leveraging wavelet decomposition to preserve both frequency details and spatial localization. Unlike FGL's global constraints,

FIBL uses localized wavelet coefficients to maintain spatial structures essential for boundary preservation. FIBL decomposes the optimization objective into two complementary components:

$$\mathcal{L}_{\text{FIBL}} = \mathcal{L}_{\text{low}} + \alpha \cdot \mathcal{L}_{\text{high}}, \tag{6}$$

where, $\mathcal{L}_{\text{low}} = \text{MSE}(\hat{F}_{ll}, F_{ll})$, $\mathcal{L}_{\text{high}} = \sum_{d\in\{lh,hl,hh\}} w_d \cdot \text{MSE}(\hat{F}_d, F_d)$. $\hat{F}_{ll}, F_{ll}$ represent the low-frequency approximation components of prediction $\boldsymbol{\mu}$ and ground truth $\mathbf{Y}$, while $\hat{F}_d, F_d$ denote high-frequency details in horizontal (lh), vertical (hl), and diagonal (hh) directions. The parameter α is determined by energy distribution statistics between frequency components (Appendix B.2), and w_d weights account for varying importance of directional boundary characteristics in meteorological structures.

To ensure proper optimization progression that aligns with physical storm development, we employ a temperature scheduling strategy that progressively transitions from FGL (computation details in Appendix B.2) to FIBL during training. This reflects the natural meteorological development where intensity patterns establish before boundary refinement. Using cosine annealing $P(t) = \cos(\pi t/(2T))$ and random sampling $p \sim \mathcal{U}(0, 1)$, the training loss becomes:

$$\mathcal{L}_{\text{TSS}} = \begin{cases} \mathcal{L}_{\text{FGL}}(\boldsymbol{\mu}, \mathbf{Y}), & p > P(t) \\ \mathcal{L}_{\text{FIBL}}(\boldsymbol{\mu}, \mathbf{Y}). & \text{otherwise} \end{cases} \tag{7}$$

3.5 Stage II: Detail-Enhanced Diffusion Refinement

While Stage I generates a coarse radar estimate $\boldsymbol{\mu}$ with proper frequency domain separation, this initial prediction lacks fine-grained precipitation boundaries and detailed intensity variations critical for accurate severe weather characterization. To address this limitation, we introduce a Detail-Enhanced Diffusion Refinement stage that leverages physics-aware frequency-decomposed priors through conditional diffusion modeling.

The conditioning framework integrates multiple information sources to capture complex precipitation dynamics:

$$\mathbf{C} = [\boldsymbol{\mu}, \mathbf{X}, \tilde{\mathbf{F}}_{\text{LF}}, \tilde{\mathbf{F}}_{\text{HF}}], \tag{8}$$

where $\tilde{\mathbf{F}}_{\text{LF}}$ and $\tilde{\mathbf{F}}_{\text{HF}}$ are frequency-decomposed features from our Wavelet-Based Frequency Feature Extractor that provides specialized low-frequency intensity and high-frequency boundary information for targeted refinement.

The diffusion model then progressively refines precipitation details through iterative denoising:

$$p_\theta(\mathbf{z}_{t-1}|\mathbf{z}_t, \mathbf{C}) = \mathcal{N}(\mathbf{z}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{z}_t, \mathbf{C}), \boldsymbol{\sigma}_\theta^2\mathbf{I}). \tag{9}$$

Wavelet-Based Frequency Feature Extractor To provide effective frequency-aware conditioning, we employ a feature extractor that applies DWT to satellite observations $\mathbf{X}$, decomposing them into frequency components $\{F_{LF}, F_{HF}^{(d)}, F_{HF}^{(v)}, F_{HF}^{(h)}\}$ (see Fig. 4). The high-frequency components are aggregated as $F_{HF} = F_{HF}^{(d)} + F_{HF}^{(v)} + F_{HF}^{(h)}$,

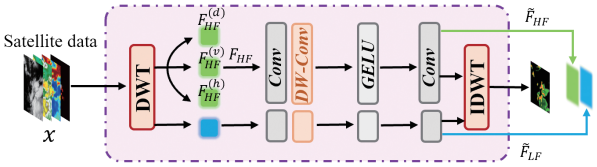

Figure 4: Wavelet-Based Frequency Feature Extractor.

and both frequency domains are processed through convolutional transformations:

$$\begin{aligned} \tilde{\mathbf{F}}_{\text{LF}} &= \text{Conv}(\text{DWConv}(\text{GELU}(\text{Conv}(F_{LH})))), \\ \tilde{\mathbf{F}}_{\text{HF}} &= \text{Conv}(\text{DWConv}(\text{GELU}(\text{Conv}(F_{HF})))). \end{aligned} \tag{10}$$

This extractor enables the diffusion model to leverage both intensity coherence from low-frequency components and boundary precision from high-frequency components for progressive detail enhancement.

Stage II Loss. The second stage optimizes a combined loss function that integrates diffusion denoising with frequency-domain consistency:

$$\mathcal{L}_{\text{Stage II}} = \mathcal{L}_{\text{Diff}} + \lambda_{\text{freq}}\mathcal{L}_{\text{Wavelet}} \tag{11}$$

where $\mathcal{L}_{\text{Diff}} = \mathbb{E}_{t,\boldsymbol{\epsilon}}[|\boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta(\mathbf{z}_t, \mathbf{C}, t)|^2]$ represents the standard diffusion denoising loss that enables progressive refinement through iterative conditioning on $\mathbf{C} = [\boldsymbol{\mu}, \mathbf{X}, \tilde{\mathbf{F}}_{\text{LF}}, \tilde{\mathbf{F}}_{\text{HF}}]$. The frequency consistency term $\mathcal{L}_{\text{Wavelet}}$ ensures proper frequency domain separation in the extracted features $\tilde{\mathbf{F}}_{\text{LF}}$ and $\tilde{\mathbf{F}}_{\text{HF}}$, enabling the diffusion model to leverage physics-aware priors for targeted boundary enhancement while preserving intensity coherence.

4 Experiments

4.1 Implementation Details

Dataset. We utilize the SEVIR dataset (Veillette, Samsi, and Mattioli 2020), which integrates multi-source observations including visible data, infrared bands (IR069, IR107), vertically integrated liquid (VIL), and lightning data. The dataset contains over 20,000 storm events from 2017-2019, each with 4-hour temporal coverage at 1 km spatial resolution. We evaluate performance at VIL thresholds of 74, 133, 160, 181, and 219 kg/m². Training details are provided in Appendix B.1.

Evaluation Metrics. To ensure fair comparison with existing methods, we adopt the same evaluation metrics as FACL (Yan et al. 2024). Meteorological performance is assessed using Critical Success Index (CSI) and Heidke Skill Score (HSS) at multiple reflectivity thresholds, which are standard metrics in weather radar evaluation (Feng et al. 2025). Image quality evaluation uses Structural Similarity Index (SSIM). For perceptual assessment, we employ Learned Perceptual Image Patch Similarity (LPIPS).

Baseline Comparisons. We compare WaveC2R against CNN-Transformer hybrid models (AA-TransUnet (Yang 2022)), meteorology-oriented architectures (Earthformer (Gao et al. 2022)), CNN-based networks (Smaat-Unet (Kevin 2021)), and generative models including

Type	Model	Loss	Skill				Perceptual
			Avg.CSI↑	Avg.HSS↑	CSI_POOL4↑	CSI_POOL16↑	LPIPS↓
Pred.	AA-TransUnet (Yang 2022)	MSE	0.302	0.437	0.328	0.407	0.318
		FACL	0.313	0.452	0.365	0.488	0.281
		FIBL	0.314	0.455	0.366	0.489	0.306
	Earthformer (Gao et al. 2022)	MSE	0.293	0.422	0.310	0.376	0.360
		FACL	0.302	0.436	0.323	0.384	0.426
		FIBL	0.311	0.449	0.355	0.483	0.278
	Smaat-Unet (Kevin 2021)	MSE	0.293	0.421	0.309	0.364	0.362
		FACL	0.312	0.451	0.355	0.440	0.375
		FIBL	0.315	0.455	0.359	0.448	0.361
Gen.	Diffcast (Yu et al. 2024)	–	0.310	0.448	0.391	0.579	0.236
	UVCGAN (Torbunov et al. 2023)	–	0.089	0.131	0.125	0.254	0.362
	Pix2pix (Akter 2024)	–	0.240	0.355	0.266	0.337	0.349
	MeanFlow (Geng et al. 2025)	–	0.200	0.300	0.386	0.583	0.300
	Ours vs (AA-TransUnet_$\mathcal{L}_{mse}$)	–	**0.327**(+8.2%)	**0.469**(+7.3%)	**0.405**(+23.4%)	**0.592**(+45.4%)	**0.219**(+31.1%)

Table 1: Quantitative comparison on the SEVIR dataset. Gray rows highlight our FIBL loss applied to deterministic models and our WaveC2R method. Best results are in bold, second-best underlined.

Type	Models	Loss	CSI (↑)					HSS (↑)				
			$\gamma \geq 74$	$\gamma \geq 133$	$\gamma \geq 160$	$\gamma \geq 181$	$\gamma \geq 219$	$\gamma \geq 74$	$\gamma \geq 133$	$\gamma \geq 160$	$\gamma \geq 181$	$\gamma \geq 219$
Pred.	AA-TransUnet (Yang 2022)	MSE	0.506	0.329	0.317	0.262	0.095	0.638	0.483	0.476	0.412	0.174
		FACL	0.512	0.330	0.317	0.270	0.134	0.642	0.484	0.475	0.422	0.235
		FIBL	0.503	0.328	0.321	0.276	0.143	0.635	0.482	0.479	0.429	0.249
	Earthformer (Gao et al. 2022)	MSE	0.524	0.321	0.304	0.249	0.069	**0.653**	0.474	0.461	0.395	0.129
		FACL	0.520	0.302	0.288	0.253	0.147	0.642	0.441	0.441	0.400	0.256
		FIBL	0.514	0.307	0.315	0.270	0.149	0.644	0.458	0.459	0.422	0.260
	Smaat-Unet (Kevin 2021)	MSE	0.516	0.324	0.313	0.256	0.055	0.646	0.478	0.471	0.404	0.105
		FACL	0.508	0.319	0.317	0.280	0.138	0.637	0.470	0.475	0.434	0.241
		FIBL	0.510	0.314	0.319	0.289	0.144	0.638	0.465	0.477	0.445	0.251
Gen.	Diffcast (Yu et al. 2024)	–	0.501	0.322	0.321	0.282	0.126	0.629	0.473	0.479	0.436	0.223
	UVCGAN (Torbunov et al. 2023)	–	0.228	0.094	0.063	0.043	0.015	0.306	0.144	0.104	0.074	0.028
	Pix2pix (Akter 2024)	–	0.418	0.259	0.269	0.241	0.014	0.547	0.397	0.417	0.384	0.028
	MeanFlow (Geng et al. 2025)	–	0.425	0.215	0.177	0.141	0.044	0.550	0.333	0.290	0.242	0.084
	Ours	–	**0.525**	**0.337**	**0.331**	**0.292**	**0.152**	0.652	**0.490**	**0.491**	**0.448**	**0.263**
	vs (AA-TransUnet_$\mathcal{L}_{mse}$)		(+3.7%)	(+2.4%)	(+4.4%)	(+11.4%)	(+60%)		(+1.4%)	(+3.1%)	(+8.7%)	(+51%)

Table 2: CSI and HSS performance comparison. **Note:** To ensure readability, we present the performance gains (in red) in the row immediately following the Ours model results.

diffusion-based (Diffcast (Yu et al. 2024)), GAN-based methods (UVCGAN (Torbunov et al. 2023), Pix2Pix (Akter 2024)) and Flow-based methods (MeanFlow (Geng et al. 2025)). MeanFlow details are provided in Appendix A.2.

4.2 Experimental Results

Metric Comparison. Table 1 demonstrates WaveC2R's superiority over state-of-the-art methods. Compared to the best deterministic baseline (AA-TransUnet-MSE), WaveC2R achieves significant improvements: 8.2% in CSI, 7.3% in HSS, and 31.1% in LPIPS, indicating both better skill scores and visual fidelity.

Extreme Weather Performance. Table 2 presents WaveC2R's strength in severe precipitation detection. At the highest threshold ($\gamma \geq 219$), it achieves 60% improvement over AA-TransUnet-MSE (0.152 vs 0.095 CSI), while generative models like UVCGAN and Pix2pix fail with near-zero performance (0.015 and 0.014 CSI). This demonstrates WaveC2R's practical value for operational severe weather applications.

Qualitative Performance. Fig. 5 further illustrates qualitative comparisons for heavy precipitation events. CNN-based models (AA-TransUnet, Smaat-Unet) tend to over-smooth high-reflectivity cores, while Earthformer overestimates moderate-intensity regions. Generative models (Diffcast, UVCGAN, Pix2pix) introduce spurious artifacts. In contrast, WaveC2R accurately preserves both intensity distribution and sharp precipitation boundaries, demonstrating superior reconstruction quality for severe weather systems.

4.3 Ablation Studies

To validate the effectiveness of each component, we conduct controlled experiments on the SEVIR dataset. Our method consists of four key components: Wavelet-Temporal-Frequency (WTF) module, visible (VIS) channel, Detail-Enhanced Diffusion Refinement (DEDR), and

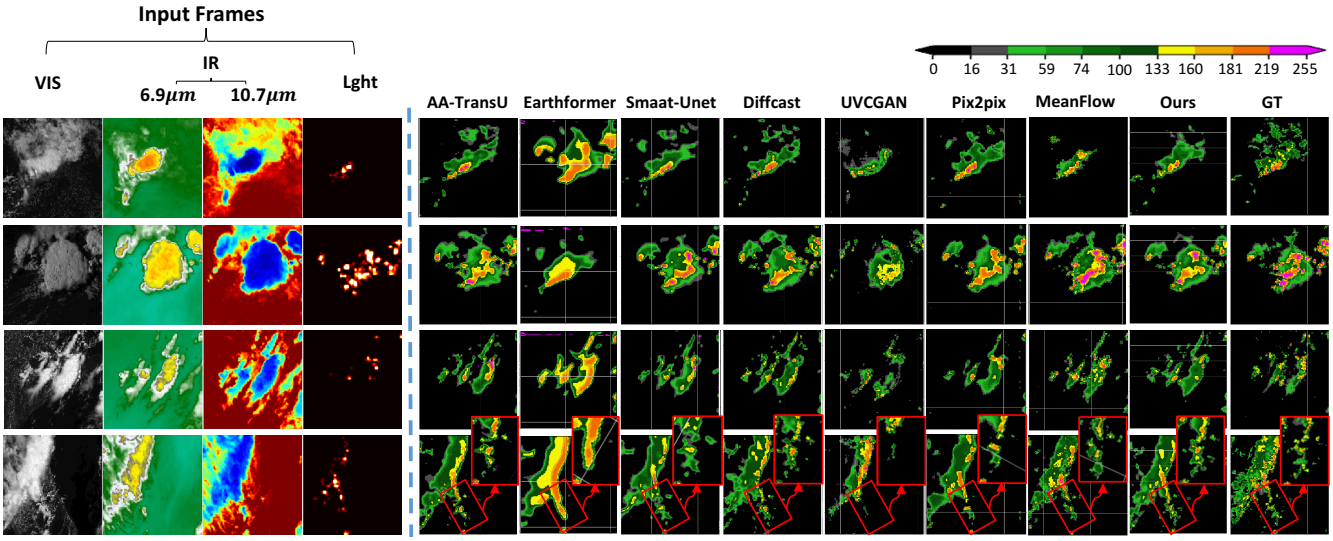

Figure 5: Qualitative comparison of radar reflectivity predictions for four heavy precipitation events from the SEVIR dataset. Red boxes highlight magnified convective regions where our WaveC2R demonstrates superior intensity accuracy and boundary sharpness compared to competing methods. Input modalities include visible, dual-channel infrared, and lightning observations.

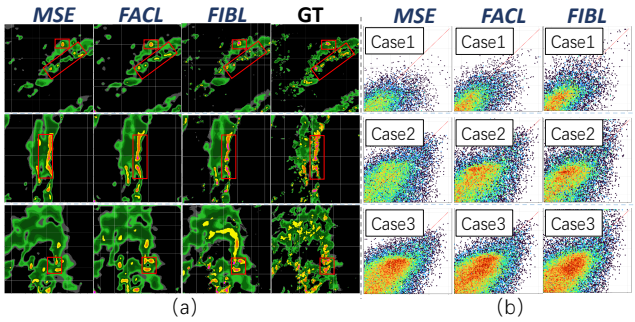

Figure 6: Ablation study of loss functions on Earthformer. (a) Three convective cases shows FIBL achieves superior intensity accuracy and sharper boundaries. (b) Density scatter plots demonstrate better ground truth alignment, particularly in high-intensity regions.

WTF	VIS	DEDR	HLF	CSI ↑	CSI_POOL4 ↑
✗	✓	✗	✗	0.311	0.328
✓	✗	✗	✗	0.344	0.383
✓	✓	✗	✗	0.360	0.413
✓	✓	✓	✗	<u>0.370</u>	<u>0.446</u>
✓	✓	✓	✓	**0.388**	**0.459**

Table 3: Ablation study on the SEVIR dataset based on CSI.

WTF	VIS	DEDR	HLF	LPIPS ↓	SSIM ↑
✗	✓	✗	✗	0.410	0.604
✓	✗	✗	✗	0.331	0.597
✓	✓	✗	✗	0.301	<u>0.612</u>
✓	✓	✓	✗	<u>0.227</u>	0.611
✓	✓	✓	✓	**0.219**	**0.615**

Table 4: Ablation study on the SEVIR dataset based on image quality metrics.

High-Low Frequency (HLF). Tables 3 and 4 demonstrate that the complete WaveC2R achieves superior performance across all metrics, highlighting the synergistic effects between components for accurate radar retrival.

Component Analysis. Each component contributes distinctly to reconstruction quality. Removing WTF causes the most significant degradation (CSI: 0.388→0.311, LPIPS: 0.219→0.410), confirming that frequency-aware processing is essential. Excluding VIS, DEDR, and HLF components results in progressive performance drops (CSI: 0.344, 0.360, 0.370 respectively), demonstrating their cumulative importance for accurate radar reflectivity reconstruction. These results validate the effectiveness of each proposed component. Additional ablation studies are provided in Appendix C.

FIBL Loss Analysis. Comparing FIBL against MSE and FACL losses on Earthformer shows FIBL achieves the highest extreme-threshold performance (see Table 2): CSI ($\gamma \geq 219$) improves from 0.069 (MSE) to 0.149 (+116%), while maintaining superior boundary sharpness without artifacts. Fig. 6 visualizes FIBL's effectiveness in preserving convective cores across diverse cases.

Conclusion

In this work, we address the challenge of coupled intensity-boundary optimization in satellite-based radar retrieval by introducing WaveC2R, a wavelet-driven coarse-to-refined framework. Based on the insight that precipitation intensity manifests in low-frequency components while boundaries are captured in high-frequency details, WaveC2R decomposes the retrieval task into separate optimization processes through frequency-domain decomposition. Our WTF module enables multi-source feature fusion, while our FIBL mitigates energy imbalance in spatial-domain approaches. The diffusion refinement stage leverages physics-aware frequency-decomposed priors for progressive detail enhancement.

Acknowledgments

This work was supported by the Heavy Rainfall Research Foundation of China (No. BYKJ2025M14), China Meteorological Administration Xiong'an Atmospheric Boundary Layer Key Laboratory (No. 2025LABL-B12), Open Project of Key Laboratory of High Impact Weather(special), by the Beijige Foundation of Nanjing Joint Institute for Atmospheric Sciences (NJIAS) under Grant BJG202305, by Key Laboratory of Smart Earth, NO. KF2023YB03-05, and by the National Natural Science Foundation of China (62374031, 62331009), and by NSFC-Jiangsu Province (BK20240173).

References

Ahn, E.; and Hur, J. 2023. A short-term forecasting of wind power outputs using the enhanced wavelet transform and arimax techniques. *Renewable Energy*, 212: 394–402.

Akter, S. 2024. Generative AI: A Pix2pix-GAN-Based Machine Learning Approach for Robust and Efficient Lung Segmentation. *arXiv preprint arXiv:2412.10826*.

Bao, C.; Xing, K.; Zhang, X.; Ma, Z.; Wang, Y.; Wan, J.; Li, L.; Liu, C.; Wen, J.; Zhang, L.; et al. 2024. Infrared Precipitation Retrieval Method Based on Residual Deep Forest. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*

Chen, W.; Hao, X.; Wu, Y.; and Liang, Y. 2024. Terra: A multimodal spatio-temporal dataset spanning the earth. In *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, volume 37, 66329–66356.

Di, D.; Li, J.; Han, W.; Bai, W.; Wu, C.; and Menzel, W. P. 2018. Enhancing the fast radiative transfer model for FengYun-4 GIIRS by using local training profiles. *J. Geophys. Res. Atmos.*, 123(22): 12–583.

Feng, W.; Li, X.; Wu, Z.; Lin, K.; Yu, D.; Ye, Y.; and Wang, Y. 2025. Perceptually Constrained Precipitation Nowcasting Model. In *Proc. Int. Conf. Mach. Learn. (ICML)*.

Gao, Z.; Shi, X.; Wang, H.; Zhu, Y.; Wang, Y. B.; Li, M.; and Yeung, D.-Y. 2022. Earthformer: Exploring space-time transformers for earth system forecasting. In *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, volume 35, 25390–25403.

Geng, Z.; Deng, M.; Bai, X.; Kolter, J. Z.; and He, K. 2025. Mean flows for one-step generative modeling. *arXiv preprint arXiv:2505.13447*.

Gong, J.; Bai, L.; Ye, P.; Xu, W.; Liu, N.; Dai, J.; Yang, X.; and Ouyang, W. 2024. Cascast: Skillful high-resolution precipitation nowcasting via cascaded modelling. *arXiv preprint arXiv:2402.04290*.

Jiang, P.; Liu, R.; Huo, Y.; Wu, Y.; Ye, S.; Wang, S.; Mu, X.; and Zhu, L. 2024. Retrieving the Atmospheric Water Vapor Profile Combining FY-4A/GIIRS and Ground-based GNSS PWV in Hongkong Region. *IEEE Trans. Geosci. Remote Sens.*

Jin, Z.; Lin, H.; and Xu, X. 2023. RDA-Unet: A Retrieval Model for Radar Composite Reflectivity Factor from Himawari-8 Observations. In *Proc. IEEE Int. Conf. High Perform. Comput. Commun. (HPCC)*, 994–1000.

Jing, Y.; Lin, L.; Li, X.; Li, T.; and Shen, H. 2022. Cascaded downscaling–calibration networks for satellite precipitation estimation. *IEEE Geosci. Remote Sens. Lett.*, 19: 1–5.

Kevin, T. 2021. SmaAt-UNet: Precipitation nowcasting using a small attention-UNet architecture. *Pattern Recognit. Lett.*, 145: 178–186.

Li, S.; Yang, W.; Zhang, P.; Xiao, X.; Cao, D.; Qin, Y.; Zhang, X.; Zhao, Y.; and Bogdan, P. 2025. ClimateLLM: Efficient Weather Forecasting via Frequency-Aware Large Language Models. *arXiv preprint arXiv:2502.11059*.

Li, Y.; Niu, D.; Li, Y.; Zang, Z.; Wang, H.; and Jiang, M. 2024. SCRD: A Spatiotemporal Cues-Guided Residual Diffusion Model for Precipitation Nowcasting. *IEEE Geosci. Remote Sens. Lett.*

Li, Z.; Wright, D. B.; Hartke, S. H.; Kirschbaum, D. B.; Khan, S.; Maggioni, V.; and Kirstetter, P.-E. 2023. Toward a globally-applicable uncertainty quantification framework for satellite multisensor precipitation products based on GPM DPR. *IEEE Trans. Geosci. Remote Sens.*, 61: 1–15.

Liao, Y.; Gao, Z.; and Li, X. 2024. Wind Farm Meteorological Prediction Model based on Frequency Domain Feature Extraction Fusion Mechanism. *IEEE Access*.

Liu, S.; Grassotti, C.; Liu, Q.; Lee, Y.-K.; Honeyager, R.; Zhou, Y.; and Fang, M. 2020. The NOAA Microwave Integrated Retrieval System (MiRS): Validation of precipitation from multiple polar-orbiting satellites. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 13: 3019–3031.

Mammedov, Y. D.; Olugu, E. U.; and Farah, G. A. 2022. Weather forecasting based on data-driven and physics-informed reservoir computing models. *Environ. Sci. Pollut. Res.*, 1–14.

Niu, D.; Li, Y.; Wang, H.; Zang, Z.; Jiang, M.; Chen, X.; and Huang, Q. 2024. FsrGAN: A Satellite and Radar-Based Fusion Prediction Network for Precipitation Nowcasting. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*

Park, Y.-J.; Kim, D.; Seo, M.; Jeon, H.-G.; and Choi, Y. 2025. Data-driven Precipitation Nowcasting Using Satellite Imagery. In *Proc. AAAI Conf. Artif. Intell. (AAAI)*, volume 39, 28284–28292.

Saoud, L. S.; Al-Marzouqi, H.; and Deriche, M. 2021. Wind speed forecasting using the stationary wavelet transform and quaternion adaptive-gradient methods. *IEEE Access*, 9: 127356–127367.

Si, J.; Li, X.; Chen, H.; and Han, L. 2023. A Novel CNN Based Radar Reflectivity Retrieval Network Using Geostationary Satellite Observations. *IEEE Geosci. Remote Sens. Lett.*

Torbunov, D.; Huang, Y.; Yu, H.; Huang, J.; Yoo, S.; Lin, M.; Viren, B.; and Ren, Y. 2023. Uvcgan: Unet vision transformer cycle-consistent gan for unpaired image-to-image translation. In *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis. (WACV)*, 702–712.

Veillette, M.; Samsi, S.; and Mattioli, C. 2020. Sevir: A storm event imagery dataset for deep learning applications in radar and satellite meteorology. In *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, volume 33, 22009–22019.

Wang, C.; Xu, J.; Tang, G.; Yang, Y.; and Hong, Y. 2020. Infrared precipitation estimation using convolutional neural network. *IEEE Trans. Geosci. Remote Sens.*, 58(12): 8612–8625.

Wang, H.; Pan, L.; Chen, Z.; Yang, D.; Zhang, S.; Yang, Y.; Liu, X.; Li, H.; and Tao, D. 2024. Fredf: Learning to forecast in frequency domain. *arXiv preprint arXiv:2402.02399*.

Wen, P.; Bai, L.; He, M.; Filippi, P.; Zhang, F.; Bishop, T. F.; Wang, Z.; and Hu, K. 2024. DuoCast: Duo-Probabilistic Meteorology-Aware Model for Extended Precipitation Nowcasting. *arXiv preprint arXiv:2412.01091*.

Yan, C.-W.; Foo, S. Q.; Trinh, V. H.; Yeung, D.-Y.; Wong, K.-H.; and Wong, W.-K. 2024. Fourier Amplitude and Correlation Loss: Beyond Using L2 Loss for Skillful Precipitation Nowcasting. In *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, volume 37, 100007–100041.

Yang. 2022. Aa-transunet: Attention augmented transunet for nowcasting tasks. In *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, 01–08.

Yang, W.; Chen, H.; Han, L.; and Ge, Y. 2023a. Multitask learning for precipitation estimation using satellite observations from the GOES-R series. *IEEE Geosci. Remote Sens. Lett.*, 20: 1–5.

Yang, Y.; Chen, H.; Hilburn, K. A.; Kuligowski, R. J.; and Cifelli, R. 2023b. Deep learning for precipitation retrievals using abi and glm measurements on the goes-r series. *IEEE Trans. Geosci. Remote Sens.*, 61: 1–14.

Yin, R.; Han, W.; Gao, Z.; and Li, J. 2021. Impact of high temporal resolution FY-4A Geostationary Interferometric Infrared Sounder (GIIRS) radiance measurements on Typhoon forecasts: Maria (2018) case with GRAPES global 4D-Var assimilation system. *Geophys. Res. Lett.*, 48(15): e2021GL093672.

Yu, D.; Feng, W.; Lin, K.; Li, X.; Ye, Y.; Luo, C.; and Du, W. 2025. Integrating Multi-Source Data for Long Sequence Precipitation Forecasting. In *Proc. AAAI Conf. Artif. Intell. (AAAI)*, volume 39, 28539–28547.

Yu, D.; Li, X.; Ye, Y.; Zhang, B.; Luo, C.; Dai, K.; Wang, R.; and Chen, X. 2024. Diffcast: A unified framework via residual diffusion for precipitation nowcasting. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 27758–27767.

Zhao, J.; Tan, J.; Chen, S.; Huang, Q.; Gao, L.; Li, Y.; and Wei, C. 2024a. Intelligent Reconstruction of Radar Composite Reflectivity Based on Satellite Observations and Deep Learning. *Remote Sens.*, 16(2): 275.

Zhao, X.; Zhou, Z.; Zhang, W.; Liu, Y.; Chen, X.; Gong, J.; Chen, H.; Fei, B.; Chen, S.; Ouyang, W.; et al. 2024b. Weathergfm: Learning a weather generalist foundation model via in-context learning. *arXiv preprint arXiv:2411.05420*.

Zheng, K.; He, L.; Ruan, H.; Yang, S.; Zhang, J.; Luo, C.; Tan, S.; Zhang, J.; Tian, Y.; and Cheng, J. 2024. A Cross-Modal Spatiotemporal Joint Predictive Network for Rainfall Nowcasting. *IEEE Trans. Geosci. Remote Sens.*