StormMind: Disentangled Layerwise Modeling for Convective Weather Systems

Jun Chen
HKUST (GZ)
Guangzhou, China
jchen512@connect.hkust-gz.edu.cn

Minghui Qiu
HKUST (GZ)
Guangzhou, China
mqiu585@connect.hkust-gz.edu.cn

Lin Chen
HKUST (GZ)
Guangzhou, China
lchen297@connect.hkust-gz.edu.cn

Yan Fang
HKUST (GZ)
Guangzhou, China
yfang243@connect.hkust-gz.edu.cn

Shuxin Zhong*
HKUST (GZ)
Guangzhou, China
shuxinzhong@hkust-gz.edu.cn

Binghong Chen
Guangzhou Meteorological Observatory
Guangzhou, China
chenbinghon1g@163.com

Kaishun Wu*
HKUST (GZ)
Guangzhou, China
wuks@hkust-gz.edu.cn

Abstract

Timely nowcasting is critical for public safety during fast-evolving storms, where even short delays can trigger cascading failures—as in the October 2024 Spain flash flood that claimed over 90 lives within minutes. While radar offers reliable real-time sensing of atmospheric structure, models that collapse 3D volumes into 2D slices inevitably discard vertical information essential for capturing storm growth, phase transitions, and collapse. We introduce `StormMind`, a physically grounded framework that forecasts convective evolution by modeling causal interactions across stratified atmospheric layers. `StormMind` addresses two fundamental challenges: (1) the nonlinear, asynchronous coupling between low-, mid-, and high-level processes; and (2) reflectivity uncertainty, where storms with distinct vertical structures may appear deceptively similar on radar, masking their true phase and intensity. To tackle these issues, `StormMind` designs: i) a *Convection Dynamics Extractor* that models storm evolution from two complementary perspectives—horizontal morphology, capturing the spatial organization of physical processes within individual atmospheric layers, and vertical coupling, modeling energy exchanges across layers; and ii) a *Convection Manifestation Reconstructor* that adaptively fuses intra- and inter-layer signals, conditioned on the evolving storm state, to infer phase transitions (e.g., initiation, intensification, dissipation). Evaluated on the large-scale 3D-NEXRAD dataset (2020–2022, U.S.), `StormMind` outperforms strong baselines, achieving a 14.71% gain in CSI_{40}. In real-world deployment with the Guangzhou Meteorological Bureau (Mar–May 2025), it improves CSI_{40} by 9.39% and boosts early-warning accuracy (98.33%).

CCS Concepts

• **Computing methodologies** → **Computer vision**; • **Applied computing** → **Earth and atmospheric sciences**.

Keywords

Weather System, Precipitation Nowcasting, Deep Learning

ACM Reference Format:
Jun Chen, Minghui Qiu, Lin Chen, Yan Fang, Shuxin Zhong*, Binghong Chen, and Kaishun Wu*. 2026. StormMind: Disentangled Layerwise Modeling for Convective Weather Systems. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.1 (KDD '26), August 09–13, 2026, Jeju Island, Republic of Korea.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3770854.3783926

1 Introduction

Weather nowcasting serves as the first line of defense against atmospheric threats, supporting time-critical decisions in transportation [28] and emergency response [12]—where even short delays can cascade into large-scale disasters. On October 29, 2024, a fast-intensifying rainstorm swept across Spain [15]. Warnings came too late. Cities were unprepared. Within hours, flash floods surged through streets, claiming over 90 lives. This was not an exception—it was a warning we cannot afford to miss again [38].

Existing nowcasting systems are largely dominated by radar echo extrapolation, which frames the task as a visual forecasting problem: predicting future reflectivity frames from past observations using optical flow [32], ConvLSTM [22], Transformer [29], or diffusion-based models [35]. While effective, these methods collapse vertical radar structure through operations like maximum-value projections [22] or Gaussian-smoothed composites [31]. This flattening

*Kaishun Wu and Shuxin Zhong are the corresponding authors.

KDD '26, August 09–13, 2026, Jeju Island, Republic of Korea

ACM ISBN 979-8-4007-2258-5/2026/08
https://doi.org/10.1145/3770854.3783926

erases critical distinctions: *an intensifying storm with deepening updrafts and a dissipating cell with collapsing cores may appear equally bright in projection, yet follow opposite physical trajectories.*

This assumption disregards the vertically stratified nature of atmospheric convection, where processes in the lower, middle, and high troposphere interact asynchronously to shape storm evolution [8]. Each layer contributes distinct but complementary physical cues to this evolution: Near-surface layers capture moisture convergence and gust-front initiation [1]; mid-levels govern updraft dynamics and storm organization [19]; high levels reveal outflows, divergence, and dissipation [4]. Together, these layers encode how storms breathe, intensify, and eventually collapse.

Despite sounding conceptually straightforward, modeling atmospheric convection across vertical layers introduces two fundamental challenges. First, convective processes unfold through **nonlinear, asynchronous interactions across the troposphere (C1).** Moisture convergence in the boundary layer may initiate lifting [39], triggering mid-level updrafts and latent heat release [11], which in turn feed back into high-altitude divergence that either reinforces or suppresses development depending on stratification [13]. Second, even when the patterns of vertical interaction are known, **their impact on reflectivity remain highly uncertain (C2).** For instance, strong low-level convergence may fail to initiate convection when capped by dry mid-level air [17]; transient elevated storms and sustained surface-based systems may exhibit similar reflectivity signatures [9]; and cold pool lifting may or may not succeed depending on environmental shear and inhibition [26].

To address these challenges, we introduce `StormMind`, a physically grounded framework that forecasts convective evolution by explicitly modeling causal interactions across vertically stratified atmospheric layers. To address **C1**, `StormMind` designs *Convection Dynamics Extractor*, which models storm dynamics from two complementary angles: (i) *Horizontal Morphology Encoder* captures the development of convective patterns within each layer, modeling phenomena such as storm expansion, advection, and splitting; and (ii) *Vertical Coupling Encoder* captures cross-layer interactions, including low-level convergence, mid-level latent heating, and high-level divergence. To address **C2**, `StormMind` implements *Convection Manifestation Reconstructor*, which adaptively fuses multi-level signals to infer reflectivity forecasting: (i) *Conditional Experts Integrator* integrates intra-layer and inter-layer cues—ranging from gravity wave triggers and mid-level destabilization to radiative cooling and vertical feedbacks; (ii) *Physical Reasoning Decoder* mimics mesoscale causal chains (e.g., inflow triggering, lifting, ventilation), enabling phase-aware reasoning under varying atmospheric regimes.

Our main contributions are listed as follows:

- We introduce a paradigm shift: from 2D reflectivity extrapolation to 3D storm-structure reasoning—viewing radar not just as a snapshot of reflectivity, but as a vertically layered sensor capturing how storms grow, organize, and collapse.
- `StormMind` designs two physically grounded components: i) *Convection Dynamics Extractor* disentangles convective evolution into horizontal morphology and vertical thermodynamic exchange, enabling structured reasoning aligned with mesoscale dynamics; and ii) *Convection Manifestation Reconstructor* adaptively integrates intra- and inter-layer signals—conditioned on evolving storm states—to infer phase transitions (e.g., initiation, intensification, dissipation) under varying regimes.
- We conduct extensive evaluations on the large-scale 3D-NEXRAD dataset (2020–2022, U.S.),where `StormMind` achieves a 14.71% improvements in CSI_{40} over strong baselines. Furthermore, `StormMind` is deployed in collaboration with the Guangzhou Meteorological Bureau, yielding a 9.393% gain in CSI_{40} and achieving an early-warning accuracy of 98.33% during Mar–May 2025.

2 Related Work

2.1 Video-based Echo Extrapolation

Radar echo extrapolation is commonly formulated as a video prediction task—forecasting reflectivity maps frame-by-frame. Early approaches adopt recurrent architectures such as ConvLSTM [23], which fuses convolution with LSTM units to model local spatiotemporal features, and PhyDNet [7], which adds physics-based constraint to capture long-term storm evolution. To improve scalability and spatial fidelity, recurrent-free models have gained prominence. SimVP [6] separates spatial and temporal learning through pure convolutions, while Earthformer [23] and EarthFarseer [33] utilize transformers for long-range context integration. AlphaPre [14] further enhances structure by disentangling spatial features in the frequency domain, while Diffcast [36] reframes forecasting as a diffusion-based generative process. Despite effectiveness, most methods operate on vertically collapsed radar inputs, e.g., maximum-intensity projections [22]. This erases critical physical distinctions: *an intensifying storm with deepening updrafts and a dissipating cell with collapsing cores may appear equally bright in projection, yet follow opposite physical trajectories.*

2.2 Vertical Structure of Convection System

Convective systems exhibit strong vertical heterogeneity, typically organized into three layers: a lower layer driven by boundary-layer processes such as cold pool dynamics and surface convergence, a mid-layer dominated by deep convection and vertical updrafts, and an high layer characterized by cirrus outflow and anvil spreading [20, 24]. Each layer hosts distinct physical mechanisms, operating at different spatial and temporal scales [10]. Importantly, these layers are dynamically coupled: high-altitude divergence can reinforce mid-level ascent, while cold pool outflows at the surface often trigger new convective cells [2]. These interactions are nonlinear, delayed, and modulated by mesoscale feedbacks and synoptic forcing—together forming a vertically integrated system that governs convective initiation, maintenance, and decay [8]. Despite decades of meteorological understanding, existing ML-based approaches still overlook this structure: they either focus on surface or mid-layer signals, or ingest vertically stacked radar volumes without accounting for layer-specific roles and inter-layer causality [38].

3 Methodology

3.1 Problem Formulation

We re-formulate radar echo extrapolation as a physically grounded forecasting task that models causal interactions across vertically stratified atmospheric layers. Given radar reflectivity sequences $\mathcal{X} = \{\mathcal{X}^L, \mathcal{X}^M, \mathcal{X}^H\}$, where each $\mathcal{X}^i = \{\mathbf{X}_1^i, \ldots, \mathbf{X}_T^i\}$ corresponds

to a different atmospheric layer $i \in \{L, M, H\}$—representing low (0.5–2 km), mid (2–5 km), and high (5+ km) altitudes—our objective is to learn a function $\mathcal{F}_\theta$ that captures cross-layer interactions to forecast the future K-step reflectivity in a target layer. Formally,

$$\hat{\mathcal{Y}}^j = \mathcal{F}_\theta(\mathcal{X}), \quad \text{where } \hat{\mathcal{Y}}^j = \{\mathbf{Y}^j_{T+1}, \ldots, \mathbf{Y}^j_{T+K}\}, \quad j \in \{L, M, H\}. \tag{1}$$

3.2 Overall Architecture

To model causal interactions across vertically stratified convection, we introduce `StormMind` a physically grounded framework composed of two complementary components (see Figure 1):

- **Convection Dynamics Extractor** models convective evolution by disentangling intra-layer dynamics and inter-layer feedbacks—yielding structured representations of horizontal storm organization and vertical thermodynamic exchanges, with an inductive bias grounded in physical convection principles.
- **Convection Manifestation Reconstructor** adaptively fuses intra-layer and inter-layer signals—conditioned on evolving convective states—to infer reflectivity forecasting.

3.3 Convection Dynamics Extractor

Convection Dynamics Extractor models convective evolution through two complementary modules: (i) *Horizontal Morphology Encoder* captures *intra-layer dynamics* within each atmospheric level, modeling storm expansion, advection, and splitting; (ii) *Vertical Coupling Encoder* captures *cross-layer thermodynamic exchanges*, including moisture ascent, latent heat release, downdraft-driven cooling.

3.3.1 Horizontal Morphology Encoder. To capture both localized convection and large-scale stratiform structures, we first apply a *multi-scale inception module* (MSIM) $\phi_{\text{MSIM}}(\cdot)$ [25] to each radar slice $\mathbf{X}^i_t \in \mathbb{R}^{1\times V\times W}$ from the l-th layer. MSIM uses strided convolutions with varying kernel sizes to jointly perform patchification and multi-scale feature extraction, while an average pooling branch preserves low-frequency background signals. This design mirrors the multi-scale nature of convective morphology and enables the Horizontal Morphology Encoder to better represent heterogeneous storm systems. The output $\mathbf{F}^{(i,p)}_t \in \mathbb{R}^{C_{\text{out}}}$ encodes the patch-level representation for spatial region p:

$$\begin{aligned}\mathbf{F}^{(i,p)}_t &= \phi_{\text{MSIM}}(\mathbf{X}^i_t) \\ &= \sum_{k\in\{3,\ldots,n\}} \text{Conv2D}_{k\times k}(\text{Conv2D}_{1\times 1}(X^i_t)) + \text{AvgPool}(X^i_t).\end{aligned} \tag{2}$$

To model temporal dynamics, we apply a localized cross-attention mechanism [27] across each patch token's time series:

$$\mathbf{Z}^{(i,p)}_{T:T+K} = E_{i\to i}(\mathbf{F}^{(i,p)}_{1:T}) = \text{CrossAttn}(Q^{(i,p)}_{T:T+K}; K^{(i,p)}_{1:T}, V^{(i,p)}_{1:T}), \tag{3}$$

where keys and values are generated via a MLP, i.e.,$(K^{(i,p)}_{1:T}, V^{(i,p)}_{1:T}) = \text{MLP}(\mathbf{F}^{(i,p)}_{1:T})$, and $Q^{(i,p)}_{T:T+K}$ denote future-oriented query composed of time positional embeddings.

3.3.2 Vertical Coupling Encoder. Motivated by mesoscale meteorology [2], *Vertical Coupling Encoder* models canonical cross-layer pathways observed in convective development:

- $E_{M\to H}$: *Mid-to-High ascent reinforcement*—Moisture convergence and latent heat release in mid-levels intensify updrafts, enhancing buoyancy and accelerating parcel ascent into higher layers.
- $E_{H\to M}$: High-to-Mid feedback—Cloud-top processes (e.g., anvil spreading and radiative cooling) may stabilize mid-levels, or gravity waves triggered aloft can perturb local thermodynamics and, under favorable conditions, trigger secondary convection.
- $E_{M\to L}$: Mid-level modulation of near-surface inflows—Descending air and intrusions disrupt boundary-layer thermodynamics and reshape near-surface convergence structures.
- $E_{L\to M}$: Surface-driven mid-level initiation—Boundary-layer heating and moistening reduce convective inhibition, lifting parcels into mid-levels to trigger deep convection.

To encode these directional pathways, we design an expert-driven Mixture-of-Experts (MoE) component (see Figure 2). The process begins by encoding each tropospheric layer using MSIM: $\mathbf{F}^i_t = \phi_{\text{MSIM}}(\mathbf{X}^i_t), i \in \{L, M, H\}$. Each expert $E_{i\to j}(\cdot)$, built upon a Transformer backbone, specializes in learning how signals from source layer i modulate the future evolution of target layer j:

$$\mathbf{Z}^{i\to j}_{T:T+K} = E_{i\to j}(\mathbf{F}^i_{1:T}), \quad i \neq j \tag{4}$$

To capture evolving local thermodynamic context, we apply condition-aware normalization via AdaLN [18] on target-layer features. For each feature sequence $\{\mathbf{F}^i_t\}^T_{t=1}$, AdaLN predicts adaptive scaling and bias parameters:

$$\begin{aligned}\mathbf{C}^{i\to j}_t &= \texttt{AdaLN}(\mathbf{Z}^{i\to j}_t, \mathbf{F}^j_{1:T}) = \gamma^{(j)}_t \cdot \texttt{LayerNorm}(\mathbf{Z}^{i\to j}_t) + \beta^{(j)}_t \\ \gamma^{(j)}_t, \alpha^{(j)}_t, \beta^{(j)}_t &= \texttt{MLP}(\mathbf{F}^j_{1:T}, Q_t)\end{aligned} \tag{5}$$

The final prediction is decoded by a Swish Transformer module (SwiTM) [21], fusing directional cues and contextual background, which combines a Transformer backbone with Swish activation to fuse directional cues and contextual background, while stably modeling long-range dependencies and fine-grained spatial details:

$$\hat{\mathbf{X}}^j_t = \alpha^{(j)}_t \cdot \texttt{SwiTM}(\mathbf{C}^{i\to j}_t) \tag{6}$$

Each expert is trained with a directional reconstruction loss:

$$\mathcal{L}_{i\to j} = \frac{1}{K}\sum_{t=T}^{T+K} \left\|\hat{\mathbf{X}}^j_t - \mathbf{X}^j_t\right\|^2_2, t \in [T, T+K] \tag{7}$$

3.4 Convection Manifestation Reconstructor

Convection Manifestation Reconstructor adaptively integrates multi-scale convection signal—combining intra-layer triggers (e.g., gravity wave forcing, mid-level destabilization) and inter-layer feedbacks (e.g., radiative cooling, vertical coupling)—to reason about the initiation, evolution, and decay of convection.

3.4.1 Conditional Experts Integrator.

$$\begin{aligned}\mathbf{Z}^{\text{inter}}_t &= \left\{\mathbf{Z}^{M\to H}_t, \mathbf{Z}^{H\to M}_t, \mathbf{Z}^{M\to L}_t, \mathbf{Z}^{L\to M}_t\right\} \\ \mathcal{E}^{(i,\text{inter})}_t &= \sigma\left(\mathbf{W}^{(i)}_1 \cdot \mathbf{Z}^{\text{inter}}_t + \mathrm{b}^{(i)}_1\right) \circ \mathbf{Z}^{\text{inter}}_t\end{aligned} \tag{8}$$

The inter-layer expertise suggestion is then concatenated with intra-layer expert latent feature $\mathcal{M}^{(i)}_t = \left\{\mathcal{E}^{(i,\text{expert})}_t, \mathbf{Z}^{(i)}_t\right\}, i \in \{L, H, M\}$, and is modulated as the refined clues.

$$\mathcal{E}^{(i)}_t = \sigma\left(\mathbf{W}^{(i)}_2 \cdot \mathcal{M}^{(i)}_t + \mathrm{b}^{(i)}_2\right) \circ \mathcal{M}^{(i)}_t. \tag{9}$$

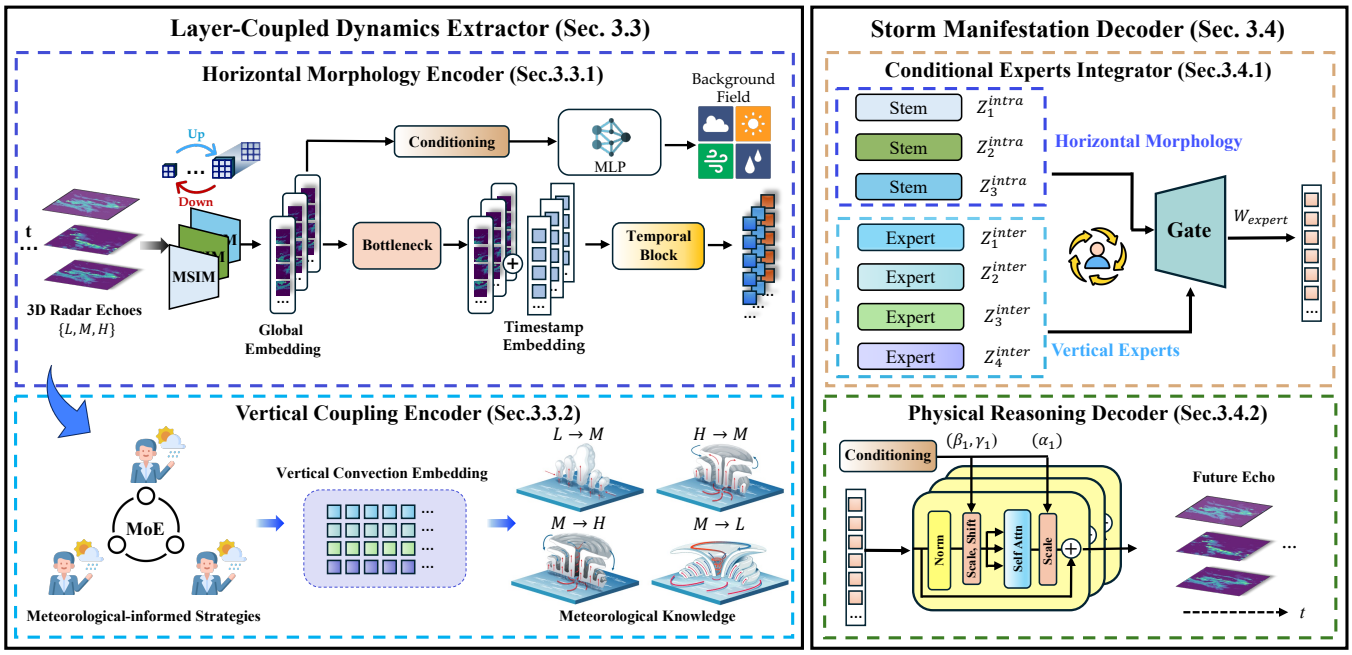

Figure 1: The Framework of StormMind.

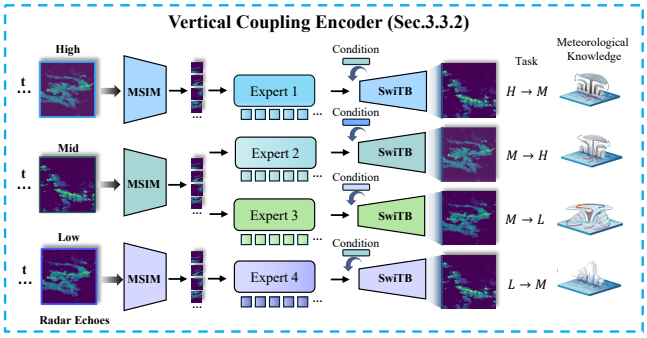

Figure 2: Illustration of Vertical Coupling Encoder.

3.4.2 Physical Reasoning Decoder. Convective systems do not evolve randomly but follow structured, physically-governed trajectories involving thermodynamic, kinematic, and microphysical interactions. To mirror this structure, we first extract the evolving environmental context, followed by adaptive fusion of expert modules specialized for different convective states:

$$\mathbf{E}_t = \texttt{AdaLN}(\mathcal{E}_t^{(i)}, \mathbf{F}_{1:T}^i), \quad i \in \{L, M, H\}. \tag{10}$$

The fused representation is passed through a SwiTM for spatial reasoning, followed by a shallow projection network (i.e., Ψ) to produce the final radar reflectivity:

$$\begin{aligned} \hat{\mathbf{E}}_t^{\text{out}} &= \texttt{SwiTM}(\mathbf{E}_t), \\ \hat{\mathbf{Y}}_t^i &= \Psi(\hat{\mathbf{E}}_t^{\text{out}}). \end{aligned} \tag{11}$$

3.5 Training Objective

We design a dual-level loss function to jointly model vertically stratified dynamics and preserve the overall coherence of the convective system: i) at the expert level, directional supervision encourages each expert $E_{i\to j}(\cdot)$ to faithfully capture cross-layer influences:

$$\mathcal{L}_{\text{expert}} = \frac{1}{N} \sum_{i \neq j} \mathcal{L}_{i \to j}. \tag{12}$$

ii) At the system level, a global loss enforces coherence across the predicted reflectivity fields in all three layers:

$$\mathcal{L}_{\text{global}} = \sum_{i \in \{L,M,H\}} \sum_{t=T}^{T+K} \|\hat{\mathbf{Y}}_t^i - \mathbf{Y}_t^i\|_2^2. \tag{13}$$

The final objective combines both terms, modulated by a tunable expert weight λ_{expert}:

$$\mathcal{L} = \mathcal{L}_{\text{global}} + \lambda_{\text{expert}} \mathcal{L}_{\text{expert}}. \tag{14}$$

4 Evaluation

This section presents a comprehensive evaluation of our StormMind by addressing the following key questions:

- **RQ1**: What is StormMind's overall performance?
- **RQ2**: How does each component affect performance?
- **RQ3**: What is the effect of the major hyperparameters?
- **RQ4**: How does StormMind perform during storm initiation and dissipation phases?
- **RQ5**: What is the real-world impact of deploying StormMind?
- **RQ6**: What are the computational requirements of StormMind?

Table 1: Performance comparison across different levels. Best results are in bold, second best are underlined. Final column shows the relative improvement of the best model over the second best (%).

Level	Metric	ConvLSTM	PredRNNv2	PhyDNet	SimVP	Earthformer	PastNet	Earthfarseer	AlphaPre	StormMind	Improv.
High	MAE↓	2.316	2.224	2.333	2.301	2.296	2.397	2.367	2.475	**2.134**	4.045%
	mCSI↑	0.291	0.301	0.324	0.294	0.318	0.274	0.328	0.303	**0.348**	6.098%
	CSI_{20}↑	0.520	0.525	0.549	0.520	0.550	0.517	0.563	0.531	**0.578**	2.664%
	CSI_{30}↑	0.253	0.262	0.293	0.255	0.286	0.239	0.299	0.270	**0.316**	5.685%
	CSI_{40}↑	0.101	0.115	0.131	0.105	0.120	0.067	0.123	0.106	**0.152**	16.031%
	SSIM↑	0.761	0.769	0.743	0.758	0.751	0.723	0.710	0.718	**0.781**	1.562%
	PSNR↑	22.918	23.192	23.693	22.879	23.660	23.067	23.660	23.546	**23.959**	1.122%
Mid	MAE↓	2.983	2.872	2.897	3.030	2.817	2.999	2.957	2.985	**2.767**	1.773%
	mCSI↑	0.405	0.414	0.431	0.414	0.414	0.401	0.433	0.402	**0.463**	6.933%
	CSI_{20}↑	0.566	0.573	0.598	0.568	0.587	0.580	0.601	0.584	**0.612**	1.831%
	CSI_{30}↑	0.451	0.460	0.490	0.458	0.477	0.451	0.484	0.451	**0.507**	3.469%
	CSI_{40}↑	0.198	0.209	0.205	0.216	0.179	0.171	0.214	0.172	**0.269**	24.537%
	SSIM↑	0.714	0.725	0.699	0.703	0.722	0.681	0.669	0.695	**0.736**	1.517%
	PSNR↑	21.452	21.666	22.323	21.375	22.268	21.710	22.277	22.221	**22.472**	0.667%
Low	MAE↓	2.732	2.622	2.610	2.713	2.653	2.714	2.560	2.771	**2.480**	3.125%
	mCSI↑	0.374	0.388	0.400	0.391	0.379	0.364	0.399	0.370	**0.427**	6.750%
	CSI_{20}↑	0.528	0.540	0.564	0.536	0.550	0.540	0.566	0.546	**0.575**	1.590%
	CSI_{30}↑	0.418	0.431	0.451	0.434	0.431	0.406	0.441	0.404	**0.473**	4.880%
	CSI_{40}↑	0.177	0.194	0.187	0.204	0.158	0.145	0.190	0.161	**0.234**	14.706%
	SSIM↑	0.723	0.734	0.723	0.724	0.715	0.691	0.718	0.700	**0.754**	2.722%
	PSNR↑	21.745	21.980	22.605	21.741	22.578	22.020	22.583	22.523	**22.804**	0.880%

4.1 Evaluation Settings.

4.1.1 Datasets. We evaluate `StormMind` on the **3D-NEXRDA** dataset, a large-scale benchmark of high-resolution volumetric radar scans capturing severe convective storms across the continental United States [3]. Spanning January 2020 to December 2022, each event in the dataset is represented as a 25-frame sequences sampled at 5-minute intervals. Each volume offers a horizontal resolution of 512×512 and 28 vertical levels from 0.5 km to 22 km altitude.

4.1.2 Baselines. We compare `StormMind` with eight baselines:

- **ConvLSTM** [22] models radar evolution as a seq-to-seq video prediction task to capture local spatiotemporal dependencies.
- **PredRNN-V2** [30] extends ConvLSTM with spatiotemporal memory cells to better model long-term dependencies.
- **PhyDNet** [7] implements a physics-inspired architecture that disentangles motion and content via dedicated encoders.
- **SimVP** [6] adopts a minimalist two-stage architecture that separates spatial encoding and temporal forecasting.
- **Earthformer** [5] employs cuboid self-attention to efficiently capture short-range spatiotemporal dependencies.
- **PastNet** [34] employs a multi-path encoder infused with physics-informed inductive biases to strike a balance between prediction accuracy and computational cost.
- **Earthfarseer** [33] advances Earthformer with causality-aware modules and hierarchical fusion for long-range modeling.
- **AlphaPre** [14] improves spatial detail preservation by disentangling amplitude and phase components in the frequency domain.

4.2 Implementation Details.

`StormMind` is trained on a single NVIDIA A800 GPU for 200 epochs with a batch size of 8, utilising the PyTorch framework. During training, we incorporate a pretext task of predicting the input frames. A warmup phase with 20 epochs is introduced to ensure the effective model learning, during which the weight of the input prediction loss is set to 1 and is fixed at 1×10^{-4} thereafter.

4.2.1 Metrics. To assess the quantitative accuracy and physical realism of predicted radar fields, we adopt a comprehensive set of metrics: (1) **Mean Absolute Error (MAE)** measures pointwise deviation; lower values indicate higher numerical precision. (2) **Critical Success Index (CSI)**, computed at 20, 30, and 40 dBZ thresholds (and averaged), evaluates event-based skill in detecting convective structures; higher values reflect more reliable storm localization. (3) **Structural Similarity Index (SSIM)** and **Peak Signal-to-Noise Ratio (PSNR)** assess spatial and perceptual consistency; higher values indicate sharper, more physically plausible forecasts.

4.3 Overall Performance (RQ1)

Table 1 shows our layer-wise evaluation and reveals three insights:

- **CSI-optimal models vary by both layer and intensity threshold.** Recurrent-based models (e.g., PhyDNet) consistently achieve top CSI_{30} in the mid and low layers, where storm organization benefits from temporally coherent modeling. These models also lead on CSI_{40} in the upper layer, suggesting that physically constrained recurrence generalizes better to sparse, vertically driven

Table 2: Ablation results grouped by level (high, mid, low).

Level	Method	MAE ↓	mCSI ↑	CSI_{20} ↑	CSI_{30} ↑	CSI_{40} ↑	SSIM ↑	PSNR ↑	SNR ↑
High	**Ours**	**2.134**	**0.348**	**0.578**	**0.316**	**0.152**	**0.781**	**23.959**	**6.273**
	StormMind w/o VCE	2.187	0.330	0.560	0.298	0.133	0.775	23.765	6.079
	StormMind w/o Merge	2.276	0.326	0.558	0.298	0.122	0.764	23.632	5.946
	StormMind w/ E_i ($E_{M\to H}$)	2.193	0.331	0.553	0.298	0.141	0.775	23.555	5.868
Mid	**Ours**	**2.767**	**0.463**	**0.612**	**0.507**	**0.269**	**0.736**	**22.472**	**6.338**
	StormMind w/o VCE	2.837	0.444	0.596	0.489	0.247	0.731	22.175	6.041
	StormMind w/o Merge	2.845	0.444	0.601	0.494	0.238	0.729	22.160	6.026
	StormMind w/ E_i ($E_{H\to M}$)	2.931	0.448	0.593	0.490	0.261	0.721	21.991	5.858
	StormMind w/ E_i ($E_{L\to M}$)	3.035	0.432	0.577	0.478	0.243	0.710	21.633	5.499
Low	**Ours**	**2.480**	**0.427**	**0.575**	**0.473**	**0.234**	**0.754**	**22.804**	**5.160**
	StormMind w/o VCE	2.603	0.390	0.542	0.438	0.189	0.744	22.342	4.698
	StormMind w/o Merge	2.690	0.413	0.567	0.463	0.208	0.728	22.317	4.673
	StormMind w/ E_i ($E_{M\to L}$)	2.570	0.422	0.572	0.471	0.223	0.742	22.594	4.950

storm tops. In contrast, recurrent-free models (e.g., SimVP, Earthfarseer) perform best at CSI_{20} across layers—due to their global attention and spatial filters being more suited for broad, low-intensity echoes or localized rain cores.

- **Other metrics favor recurrent-based models.** PhyDNet and PredRNN-V2 achieve lower MAE and higher SSIM/PSNR, particularly in mid and low layers. These layers exhibit richer textures and denser reflectivity, which benefit from the stepwise memory and structural regularization provided by recurrence.
- **StormMind consistently outperforms all baselines across layers.** StormMind yields the best scores across all sub-tasks, improving CSI_{40} by up to 24.537% and reducing MAE by 4.045%. These results underscore *the value of explicitly modeling causal and vertical physical interactions in stratified radar fields.*

4.4 Ablation Study (RQ2)

To evaluate each design, we conduct an ablation study (Table 2):

- StormMind w/o VCE: Removes *Vertical Coupling Encoder* (Sec. 3.3.2), disabling top-down and bottom-up context exchange.
- StormMind w/o Merge: Replaces *Horizontal Morphology Encoder* (Sec. 3.3.1) with a shared encoder to assess stratified processing.
- StormMind w/ E_i: For each layer, one expert path (e.g., $E_{H\to M}$, $E_{L\to M}$, or $E_{M\to H}$) is retained to assess diverse contextualization.

The results yield several key observations:

- **Stratified encoding outperforms monolithic design.** StormMind w/ Merge resulting drop in performance at all levels highlights the necessity of preserving vertical heterogeneity, as atmospheric dynamics differ significantly across altitudes.
- **Explicit interaction modeling is critical.** StormMind w/o VCE shows consistent degradation across all levels—particularly in mid and low layers, which confirms that explicit modeling of vertical interactions is essential for capturing contextual dependencies that cannot be learned implicitly through backbone alone.
- **Bidirectional expert fusion ensures robust vertical reasoning.** Retaining only a single expert (e.g., StormMind w/ $E_{H\to M}$ or StormMind w/ $E_{L\to M}$) reveals nuanced effects: i) at the high layer, performance remains close to StormMind w/o VCE, suggesting that updraft-related cues from mid layers suffice under relatively stable high-atmosphere conditions; ii) at the mid layer, single-directional experts underperform, as mid-level precipitation is influenced by both surface fluxes and high-level dynamics—requiring dual-sided contextualization; and iii) at the low layer, surprisingly, $E_{M\to L}$ alone performs slightly better than StormMind w/o VCE, which may stem from dominant downward influence and relatively simpler low-level structure.

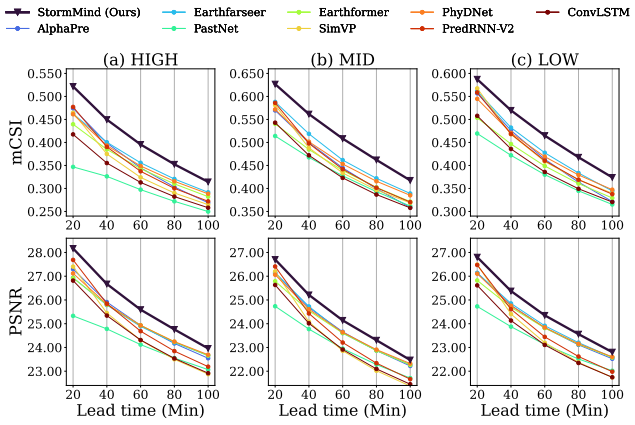

Figure 3: mCSI/MAE over Time across Different Levels.

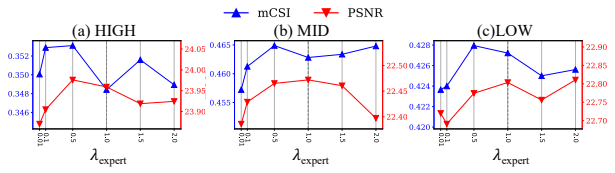

Figure 4: mCSI/PSNR over Different λ_{expert}.

4.5 Sensitivity Analysis (RQ3)

Figure 3 compares performance across lead times (20-100 mins) and vertical levels using mCSI and PSNR. As forecast horizon increases, performance gradually degrades—reflecting increased uncertainty in long-range prediction. Nevertheless, `StormMind` consistently outperforms all baselines at every lead time, maintaining both superior value-level accuracy and structural coherence.

To assess the contribution of cross-layer knowledge, we conduct a sensitivity study on the expert weighting parameter λ_{expert} (Figure 4). As λ_{expert} increases, performance first improves—highlighting the benefit of incorporating information from other layers—but then declines as the expert signal overwhelms local features. The optimal setting at $\lambda_{\text{expert}} = 1$ suggests that a balanced 1:1 fusion of cross-layer cues and layer-specific features yields the best results.

4.6 Case Study: Capturing Convective Development and Dissipation (RQ4)

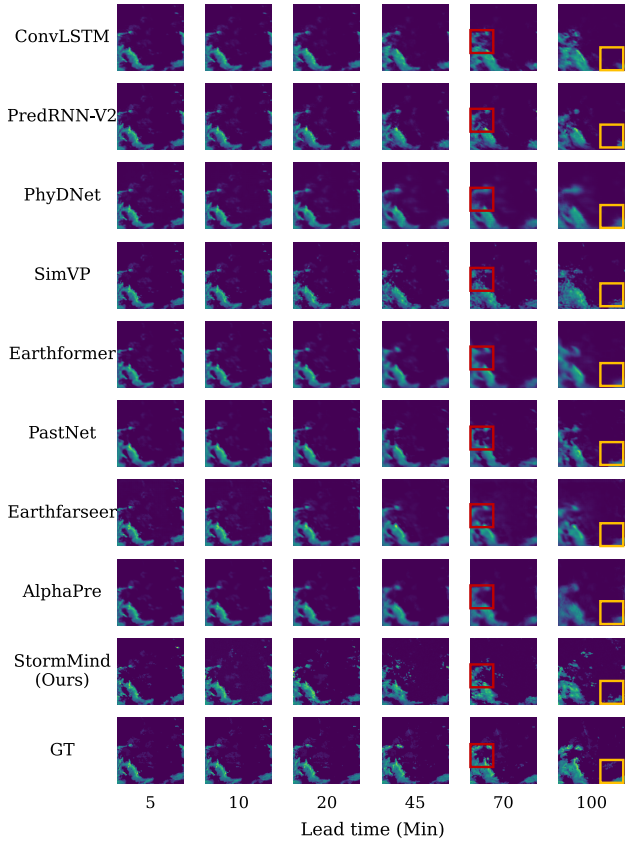

Figure 5: The Convective Development Case.

To visualize the physical validity of `StormMind` during storm evolution, we conduct two case studies representing distinct lifecycle phases: convective *intensification and dissipation*, selected based on the Convective Area Variation Rate (CAVR) [16], a physically grounded indicator of echo dynamics. Figure 5 and Figure 6 present

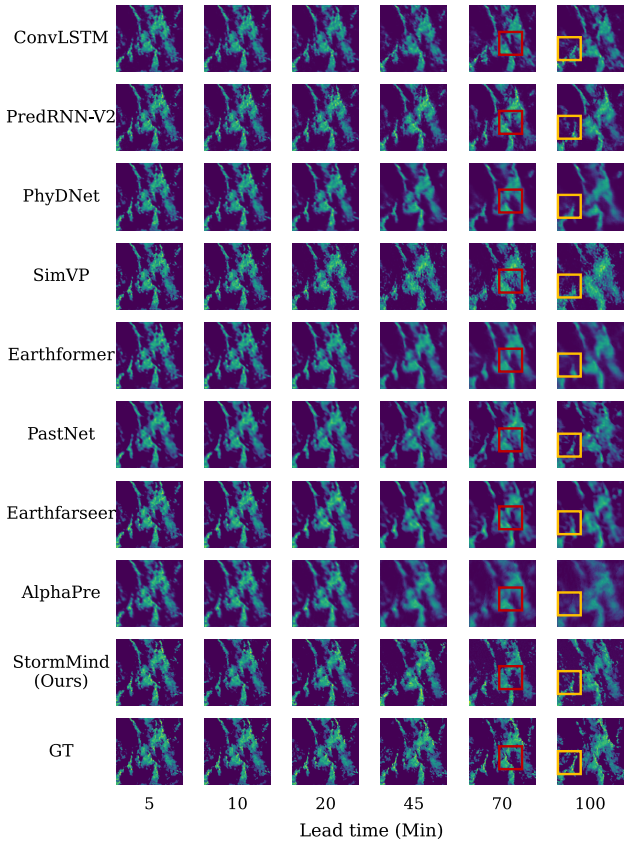

Figure 6: The Convective Dissipation Case.

mid-tropospheric forecasts (the operational layer), where `StormMind` consistently outperforms state-of-the-art baselines. In intensifying cases (positive CAVR), `StormMind` accurately captures the emergence and localization of high-intensity cores—often diffused or missed by competing models. In dissipating cases (negative CAVR), it preserves fine-grained decay structures and avoids the spatial drift. These cases highlight `StormMind`'s ability to track the full convective lifecycle with high spatial and semantic fidelity—addressing a long-standing challenge in operational nowcasting.

4.7 Real-World Deployment (RQ5)

To assess real-world effectiveness, we deployed `StormMind` in collaboration with the Guangzhou Municipal Meteorological Bureau as part of their operational early-warning system (Figure 7). The radar system operates at a horizontal resolution of 4412×4412 ($3600\ m^2$ per grid), with vertical range from 500 to 5500 meters and 1-minute update intervals. Between March to May 2025, official forecasters issued rainfall alerts based on `StormMind`'s reflectivity forecasts. In subsequent quality inspection reports, `StormMind` achieved the highest hit rate (98.33%) across all signal levels and enabled alerts issued over one hour in advance. For further validation, we conducted retrospective experiments using using the same period's data. Table 3 shows that `StormMind` outperforms all

baselines across key metrics, demonstrating superior accuracy and storm structure preservation under real-world conditions.

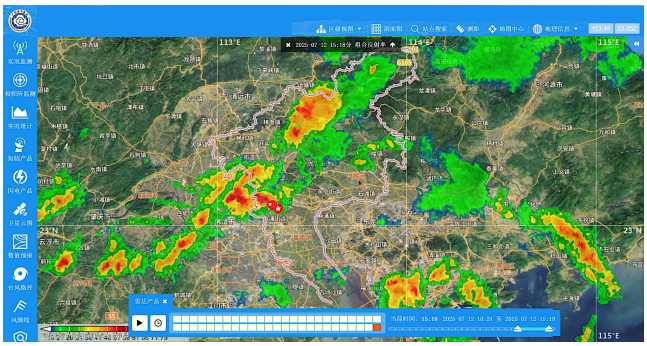

Figure 7: The Deployment of StormMind.

4.8 Real-Time Inference Efficiency (RQ6)

To evaluate the computational efficiency of StormMind , we assess their memory footprint and inference latency on a single NVIDIA A800 GPU using the U.S. and Guangzhou datasets, as shown in Table 4. Despite incorporating multiple experts, StormMind achieves a forward inference latency of approximately 100–116 ms per 20 frames, falling well within practical operational constraints and demonstrating its efficiency and suitability for real-world deployment.

5 Discussion

In this section, we summarize some insights and lessons learned. We also discuss the limitations and future directions, followed by implication and generalization.

5.1 Insights and Lessons Learned

Our comprehensive analysis yields two key insights:

- **Radar signals are vertically heterogeneous—no single model fits all.** Low levels exhibit intense, texture-rich echoes shaped by boundary forcing and terrain; mid levels encode convective cores dominated by updrafts and latent heat; high levels are sparse and noisy, indicating storm decay and ice-phase processes. As confirmed in Table 2, uniform backbones underperform compared to our layer-specific design.
- **Mid-troposphere serves as the dynamical core of convective systems.** We observe the sharpest performance variations across models and metrics (CSI, SSIM) in mid levels (Table 2). This layer concentrates both bottom-up inputs (e.g., surface convergence) and top-down constraints (e.g., shear, divergence)—making it pivotal for storm growth, structure, and transitions.

5.2 Deployment and Operational Experience

StormMind is deployed on a GPU node with 80 GB memory within an internal computing cluster, following a layered architecture that separates data acquisition, model inference, and decision support. Volumetric X-band radar scans are ingested in real time, undergoing decoding, quality control, and format standardization before being forwarded to the StormMind inference engine. Trained on South China data, the model delivers 0–60 min precipitation forecasts with an average inference latency of approximately 100 ms per 20-frame sequence. Forecast results are visualized and disseminated to forecasters through the Guangzhou Meteorological Disaster Intelligent Monitoring Subsystem for operational use.

Real-world deployment exposed several domain-specific challenges. Radar data quality issues were addressed through attenuation correction based on joint Z_H–K_{DP} calibration for X-band phased-array radars, combined with clutter suppression and velocity de-aliasing[37]. Inter-radar calibration inconsistencies were mitigated by adaptive normalization within overlapping network regions, while data latency and missing scans were handled via real-time, temporally gated interpolation. Together, these strategies enabled stable system operation under practical constraints.

Continuous deployment yielded several key findings. First, data quality assurance and consistent preprocessing proved more critical than marginal gains in model accuracy, as robust calibration and artefact suppression are essential for reliable forecasts. Second, layerwise disentanglement of the deployment pipeline improved both stability and interpretability, facilitating generalization across heterogeneous radar configurations. Finally, the design principles of StormMind—physically grounded modularization, expert feedback integration, and adaptive real-time processing—extend beyond precipitation nowcasting and may benefit other environmental AI applications, such as air quality monitoring and typhoon tracking. These findings suggest that long-term trust and operational reliability, rather than accuracy alone, are central to the sustainable deployment of AI systems in scientific practice.

5.3 Limitations and Future Work

While StormMind demonstrates strong performance across two operational regions (US and China), its current evaluation is limited to fixed geographic domains and radar-only inputs. Our work lies in the hierarchical modeling approach, which emphasizes the interactions within weather systems. And satellite imagery and surface station data provide complementary information—satellites capture upper-atmospheric evolution in radar-sparse regions, while ground stations enhance near-surface observations. Incorporating these signals represents an important future direction to further improve vertical continuity and regional coverage. Future work will explore domain adaptation strategies to improve transferability across regions with differing sensor configurations, climates, and radar calibration schemes. We will also integrate additional modalities, e.g., automatic weather station, to model radar–rainfall relationships and further enhance physical grounding.

5.4 Implication and Generalization

By leveraging physically structured priors—such as layerwise convection and vertical causal pathways— StormMind not only improves accuracy but also enhances interpretability and acceptance by domain experts. This design principle readily generalizes to other meteorological tasks that require cross-field reasoning, including storm lifecycle classification, convective event detection, and precipitation phase estimation. More broadly, it offers a scalable

Table 3: Performance comparison on the Guangzhou Meteorological Bureau dataset across different levels.

Level	Metric	ConvLSTM	PredRNNv2	PhyDNet	SimVP	Earthformer	PastNet	Earthfarseer	AlphaPre	`StormMind`	Improv.
High	**MAE ↓**	1.785	1.598	2.312	1.730	1.662	1.797	2.338	2.038	**1.419**	11.199%
	mCSI ↑	0.419	0.442	0.439	0.449	0.417	0.303	0.216	0.407	**0.504**	12.249%
	CSI_{20} ↑	0.614	0.650	0.636	0.645	0.641	0.552	0.426	0.633	**0.687**	5.692%
	CSI_{30} ↑	0.380	0.413	0.402	0.399	0.380	0.239	0.165	0.382	**0.468**	13.317%
	CSI_{40} ↑	0.264	0.261	0.278	0.305	0.229	0.117	0.058	0.206	**0.357**	17.049%
	SSIM ↑	0.819	0.830	0.619	0.802	0.824	0.820	0.771	0.739	**0.859**	3.494%
	PSNR ↑	24.122	25.342	24.382	25.113	25.187	24.204	22.536	24.445	**26.151**	3.192%
Mid	**MAE ↓**	2.800	2.496	3.318	2.524	2.484	2.692	2.798	3.148	**2.368**	4.670%
	mCSI ↑	0.598	0.643	0.625	0.639	0.625	0.582	0.587	0.575	**0.674**	4.821%
	CSI_{20} ↑	0.710	0.744	0.742	0.749	0.733	0.705	0.724	0.692	**0.754**	0.668%
	CSI_{30} ↑	0.641	0.675	0.666	0.675	0.668	0.615	0.615	0.599	**0.709**	5.037%
	CSI_{40} ↑	0.442	0.511	0.468	0.492	0.474	0.426	0.423	0.434	**0.559**	9.393%
	SSIM ↑	0.764	0.771	0.516	0.763	0.785	0.762	0.733	0.693	**0.792**	0.892%
	PSNR ↑	20.266	21.557	21.249	21.577	21.274	20.689	20.931	20.687	**21.652**	0.348%
Low	**MAE ↓**	2.849	2.528	3.262	2.525	2.662	2.853	2.986	3.031	**2.473**	2.059%
	mCSI ↑	0.590	0.636	0.621	0.639	0.602	0.565	0.573	0.584	**0.664**	3.912%
	CSI_{20} ↑	0.705	0.740	0.732	0.747	0.722	0.701	0.704	0.707	**0.752**	0.669%
	CSI_{30} ↑	0.635	0.675	0.664	0.676	0.639	0.600	0.604	0.622	**0.698**	3.254%
	CSI_{40} ↑	0.432	0.493	0.469	0.495	0.446	0.395	0.411	0.422	**0.542**	9.495%
	SSIM ↑	0.759	0.765	0.545	0.776	0.764	0.746	0.720	0.705	**0.779**	0.387%
	PSNR ↑	20.096	21.337	21.025	21.420	20.970	20.246	20.230	20.649	**21.483**	0.294%

Table 4: Memory footprint and inference latency comparison on different datasets.

Model	Static Memory	Peak Memory	Dynamic Memory	Latency
U.S.				
Earthfarseer	81.32	607.21	525.89	86.89
Ours	177.81	429.48	251.67	103.60
Guangzhou				
Earthfarseer	81.90	670.55	588.65	86.11
Ours	208.51	529.00	320.49	116.29

Note: *Static, peak, and dynamic memory are measured in MB. Inference latency is reported in milliseconds (ms) and corresponds to the average forward inference time for processing 20 consecutive frames.*

paradigm for integrating deep learning into scientific forecasting systems where physical consistency is critical.

6 Conclusion

This study presents a paradigm shift in radar-based forecasting, from 2D reflectivity extrapolation to 3D storm-structure reasoning, treating radar not just as a visual input, but as a dynamic, physically grounded lens into convective phase transitions. We introduce `StormMind`, a vertically structured framework for interpretable nowcasting that explicitly models causal interactions across stratified atmospheric layers. To address the fundamental challenges of nonlinear vertical coupling and reflectivity uncertainty, `StormMind` designs: a *Convection Dynamics Extractor* that disentangles storm evolution into horizontal morphology and vertical thermodynamic exchanges, grounded in mesoscale meteorology; and a *Convection Manifestation Reconstructor* that adaptively fuses intra- and inter-layer cues—conditioned on storm state—to infer critical phase transitions such as initiation, intensification, and dissipation. Experimentally, `StormMind` consistently outperforms strong baselines on the large-scale 3D-NEXRAD dataset (2020–2022, U.S.), achieving a 14.71% gain in CSI_{40}, and further proves its operational value in real-world deployment with the Guangzhou Meteorological Bureau (Mar–May 2025), improving CSI_{40} by 9.39% and enabling 98.33% early-warning accuracy. Together, these results demonstrate that vertical reasoning is not only physically essential but also practically transformative—paving the way for reliable, interpretable, and deployable storm-scale nowcasting.

Acknowledgments

This work is supported partly by Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things (No.2023B1212010007), China NSFC Grant (No.62472366), the Project of DEGP (No.2024GCZX003, 2023KCXTD042), 111 Center (No.D25008), Shenzhen Science and Technology Foundation (ZDSYS20190902092853047), Guangdong-Hong Kong-Macao Greater Bay Area Academy of Meteorological Research (GHMA2025Y03). Guangdong Basic and Applied Basic Research Foundation (2025A1515510015).

References

[1] Lanqiang Bai, Zhiyong Meng, Yipeng Huang, Yunji Zhang, Shuzhen Niu, and Tao Su. 2019. Convection initiation resulting from the interaction between a quasi-stationary dryline and intersecting gust fronts: A case study. *Journal of Geophysical Research: Atmospheres* 124, 5 (2019), 2379–2396.

[2] Sandrine Bony, Bjorn Stevens, Dargan MW Frierson, Christian Jakob, Masa Kageyama, Robert Pincus, Theodore G Shepherd, Steven C Sherwood, A Pier Siebesma, Adam H Sobel, et al. 2015. Clouds, circulation and climate sensitivity. *Nature Geoscience* 8, 4 (2015), 261–268.

[3] Kenneth Bowman and Cameron Homeyer. 2017. Gridrad-Three-dimensional gridded Nexrad WSR-88D radar data. (2017).

[4] Michael C Coniglio, David J Stensrud, and Louis J Wicker. 2006. Effects of upper-level shear on the structure and maintenance of strong quasi-linear mesoscale convective systems. *Journal of the Atmospheric Sciences* 63, 4 (2006), 1231–1252.

[5] Zhihan Gao, Xingjian Shi, Hao Wang, Yi Zhu, Bernie Wang, Mu Li, and Dit-Yan Yeung. 2022. Earthformer: Exploring Space-Time Transformers for Earth System Forecasting. In *Advances in Neural Information Processing Systems*, Alice H. Oh, Alekh Agarwal, Danielle Belgrave, and Kyunghyun Cho (Eds.). https://openreview.net/forum?id=lzZstLVGVGW

[6] Zhangyang Gao, Cheng Tan, Lirong Wu, and Stan Z Li. 2022. Simvp: Simpler yet better video prediction. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 3170–3180.

[7] Vincent Le Guen and Nicolas Thome. 2020. Disentangling physical dynamics from unknown factors for unsupervised video prediction. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 11474–11484.

[8] Robert A Houze Jr. 2004. Mesoscale convective systems. *Reviews of Geophysics* 42, 4 (2004).

[9] Robert A Houze Jr and Alan K Betts. 1981. Convection in GATE. *Reviews of Geophysics* 19, 4 (1981), 541–576.

[10] Robert A Houze Jr and Peter V Hobbs. 1982. Organization and structure of precipitating cloud systems. In *Advances in geophysics*. Vol. 24. Elsevier, 225–315.

[11] Lee Ray Hoxit, Charles F Chappell, and J Michael Fritsch. 1976. Formation of mesolows or pressure troughs in advance of cumulonimbus clouds. *Monthly Weather Review* 104, 11 (1976), 1419–1428.

[12] Kai Jiang, Nian Liu, Kunyu Wang, Yubing Chen, Jianxiao Wang, and Yu Liu. 2025. Spatiotemporal assessment of renewable adequacy during diverse extreme weather events in China. *Nature Communications* 16, 1 (2025), 1–12.

[13] Douglas K Lilly. 1979. The dynamical structure and evolution of thunderstorms and squall lines. *Annual Review of Earth and Planetary Sciences, Vol. 7, p. 117* 7 (1979), 117.

[14] Kenghong Lin, Baoquan Zhang, Demin Yu, Wenzhi Feng, Shidong Chen, Feifan Gao, Xutao Li, and Yunming Ye. 2025. AlphaPre: Amplitude-Phase Disentanglement Model for Precipitation Nowcasting. In *Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR)*. 17841–17850.

[15] NEWS. 2024. 2024 Spain floods. https://en.wikipedia.org/wiki/2024_Spanish_floods. Accessed: 2025-07-04.

[16] Xiang Pan, Yinghui Lu, Kun Zhao, Hao Huang, Mingjun Wang, and Haonan Chen. 2021. Improving nowcasting of convective development by incorporating polarimetric radar variables into a deep-learning model. *Geophysical Research Letters* 48, 21 (2021), e2021GL095302.

[17] Douglas J Parker, Peter Willetts, Cathryn Birch, Andrew G Turner, John H Marsham, Christopher M Taylor, Seshagirirao Kolusu, and Gill M Martin. 2016. The interaction of moist convection and mid-level dry air in the advance of the onset of the Indian monsoon. *Quarterly Journal of the Royal Meteorological Society* 142, 699 (2016), 2256–2272.

[18] William Peebles and Saining Xie. 2023. Scalable diffusion models with transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*. 4195–4205.

[19] Robert F Rogers and J Michael Fritsch. 2001. Surface cyclogenesis from convectively driven amplification of midlevel mesoscale convective vortices. *Monthly weather review* 129, 4 (2001), 605–637.

[20] Russ S Schumacher and Kristen L Rasmussen. 2020. The formation, character and changing nature of mesoscale convective systems. *Nature Reviews Earth & Environment* 1, 6 (2020), 300–314.

[21] Noam Shazeer. 2020. Glu variants improve transformer. *arXiv preprint arXiv:2002.05202* (2020).

[22] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. 2015. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems* 28 (2015).

[23] Xingjian SHI, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-kin Wong, and Wang-chun WOO. 2015. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. In *Advances in Neural Information Processing Systems*, C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett (Eds.), Vol. 28. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2015/file/07563a3fe3bbe7e3ba84431ad9d055af-Paper.pdf

[24] Roland B Stull. 2012. *An introduction to boundary layer meteorology*. Vol. 13. Springer Science & Business Media.

[25] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. 2015. Going Deeper With Convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[26] Adrian M Tompkins. 2001. Organization of tropical convection in low vertical wind shears: The role of cold pools. *Journal of the atmospheric sciences* 58, 13 (2001), 1650–1672.

[27] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems* 30 (2017).

[28] Hong-Wei Wang, Zhong-Ren Peng, Dongsheng Wang, Yuan Meng, Tianlong Wu, Weili Sun, and Qing-Chang Lu. 2020. Evaluation and prediction of transportation resilience under extreme weather events: A diffusion graph convolutional approach. *Transportation research part C: emerging technologies* 115 (2020), 102619.

[29] Kun Wang, Hao Wu, Yifan Duan, Guibin Zhang, Kai Wang, Xiaojiang Peng, Yu Zheng, Yuxuan Liang, and Yang Wang. 2024. NuwaDynamics: Discovering and updating in causal spatio-temporal modeling. In *The Twelfth International Conference on Learning Representations*.

[30] Yunbo Wang, Haixu Wu, Jianjin Zhang, Zhifeng Gao, Jianmin Wang, Philip S Yu, and Mingsheng Long. 2022. Predrnn: A recurrent neural network for spatiotemporal predictive learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 2 (2022), 2208–2225.

[31] Ziye Wang, Yiran Qin, Lin Zeng, and Ruimao Zhang. 2025. High-Dynamic Radar Sequence Prediction for Weather Nowcasting Using Spatiotemporal Coherent Gaussian Representation. In *The Thirteenth International Conference on Learning Representations*. https://openreview.net/forum?id=Cjz9Xhm7sI

[32] Wang-chun Woo and Wai-kin Wong. 2017. Operational application of optical flow techniques to radar-based rainfall nowcasting. *Atmosphere* 8, 3 (2017), 48.

[33] Hao Wu, Yuxuan Liang, Wei Xiong, Zhengyang Zhou, Wei Huang, Shilong Wang, and Kun Wang. 2024. Earthfarsser: Versatile Spatio-Temporal Dynamical Systems Modeling in One Model. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 38. 15906–15914.

[34] Hao Wu, Fan Xu, Chong Chen, Xian-Sheng Hua, Xiao Luo, and Haixin Wang. 2024. Pastnet: Introducing physical inductive biases for spatio-temporal video prediction. In *Proceedings of the 32nd ACM International Conference on Multimedia*. 2917–2926.

[35] Demin Yu, Xutao Li, Yunming Ye, Baoquan Zhang, Chuyao Luo, Kuai Dai, Rui Wang, and Xunlai Chen. 2024. Diffcast: A unified framework via residual diffusion for precipitation nowcasting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 27758–27767.

[36] Demin Yu, Xutao Li, Yunming Ye, Baoquan Zhang, Chuyao Luo, Kuai Dai, Rui Wang, and Xunlai Chen. 2024. DiffCast: A Unified Framework via Residual Diffusion for Precipitation Nowcasting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 27758–27767.

[37] Zhang Yu, Liu Xian-Tong, Chen Bing-Hong, Feng Jia-Bao, Zeng Lin, and Tian Cong-Cong. 2023. Application of X-band polarimetric phased-array radars in quantitative precipitation estimation. *Journal of Tropical Meteorology* 29, 1 (2023), 142–152.

[38] Yuchen Zhang, Mingsheng Long, Kaiyuan Chen, Lanxiang Xing, Ronghua Jin, Michael I Jordan, and Jianmin Wang. 2023. Skilful nowcasting of extreme precipitation with NowcastNet. *Nature* 619, 7970 (2023), 526–532.

[39] Conrad L Ziegler and Erik N Rasmussen. 1998. The initiation of moist convection at the dryline: forecasting issues from acase study perspective. *Weather and forecasting* 13, 4 (1998), 1106–1131.