

RainODE: Continuous-Time Precipitation Forecasting with Latent Neural ODEs

Yeeun Seong^{*}, Doyi Kim^{*}, Minseok Seo, and Changick Kim[†]

Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea
{yeeunseong,doyi.kim,minseok.seo,changick}@kaist.ac.kr

Abstract. In precipitation forecasting, not only accuracy but also temporal resolution is critical. However, increasing temporal resolution is constrained by observational limitations and the computational cost of dense discrete modeling. To overcome this limitation, we reformulate precipitation forecasting as a continuous-time dynamical system and propose RainODE, a framework that models precipitation evolution in latent space using a Neural ODE. This formulation enables derivative-consistent temporal dynamics and captures the dominant large-scale advective motion of precipitation systems. Nevertheless, a purely deterministic ODE struggles to represent non-advective intensity changes such as localized growth, decay, and sub-grid variability, often leading to over-smoothed predictions. To address this issue, we introduce a stochastic source modeling module based on a Brownian Bridge formulation, which refines residual intensity variations and restores fine-grained structures while preserving advective consistency. By combining deterministic continuous dynamics with stochastic refinement, RainODE enables arbitrary-time inference while maintaining sharp predictions. Experiments on SEVIR and the newly introduced Radar-based Precipitation Integrated Dataset (RAPID) demonstrate consistent improvements across multiple temporal intervals and precipitation regimes. The code is available at <https://github.com/SeongYE/RainODE>.

Keywords: Weather Forecasting · Continuous-time Modeling

1 Introduction

Data-driven short-term precipitation forecasting has rapidly advanced with the emergence of large-scale benchmark datasets and diverse deep learning architectures [1, 10–12, 18, 26, 29, 30, 43, 44]. In particular, the SEVIR [34] dataset has become a widely adopted benchmark, where predicting 12 future timestamps at 5-minute intervals, corresponding to one hour ahead, serves as a standard evaluation protocol. Despite strong performance under such discrete settings, precipitation in the Earth system evolves continuously over time. The atmosphere

^{*}Equal contribution.

[†]Corresponding author.

does not change in isolated 5-minute increments; rather, intensity, structure, and motion evolve as a coupled dynamical process [22, 41].

However, practical forecasting models are inevitably trained on discretely sampled snapshots due to observational and modeling constraints. This mismatch between continuous physical processes and discrete prediction frameworks introduces structural limitations. While discrete models can achieve high accuracy and temporal consistency at predefined intervals, such consistency is typically tied to the specific training grid.

Existing forecasting paradigms have to densify discrete predictions to address higher temporal resolution. As shown in Fig. 1, autoregressive models suffer from error accumulation as rollouts become denser [13, 26, 37–40, 45], whereas many-to-many models jointly predict dense spatio-temporal blocks at the cost of rapidly increasing computational complexity [9, 21, 31, 32]. Implicit models [2, 7, 15, 28] provide flexibility in representing arbitrary time points, but they do not explicitly enforce dynamical consistency across time. Instead of merely increasing output frequency within fixed temporal grids, we reformulate precipitation forecasting as a continuous-time dynamical process.

We propose RainODE, a framework for continuous-time precipitation forecasting with Latent Neural ODEs. RainODE models temporal evolution in latent space through a Neural ODE, learning derivative-consistent dynamics rather than relying on discrete frame-to-frame mappings. A complementary Stochastic Source Modeling (SSM) module further captures fine-scale spatial variability while preserving large-scale dynamical consistency. Furthermore, to evaluate generalization beyond trained temporal intervals and to assess long-horizon forecasting performance up to six hours, we introduce RAPID, a new benchmark dataset supporting +6 hour prediction. Experimental results demonstrate that RainODE not only achieves state-of-the-art performance on both SEVIR and RAPID, but also enables time-axis continuous forecasting across varying temporal intervals.

2 Related Works

2.1 Data-Driven Approaches

Data-driven precipitation forecasting is commonly formulated as a spatio-temporal sequence modeling problem. Early autoregressive models [26, 27, 39, 46] preserve temporal dependency by recursively feeding back predictions, but inevitably accumulate errors over time. To mitigate this issue, many-to-many (M2M) frameworks [3, 4, 8, 9, 11, 18, 23, 25, 28, 29, 44] jointly predict the entire target horizon in a single forward pass, effectively capturing correlations within a fixed temporal window. However, extending prediction horizons typically increases computational and memory costs. Implicit forecasting models take the desired lead time as an input query and directly predict the corresponding future state, enabling flexible lead-time forecasting [2, 7, 15, 28]. While this design avoids recursive rollout, predictions at different timestamps are typically generated independently,

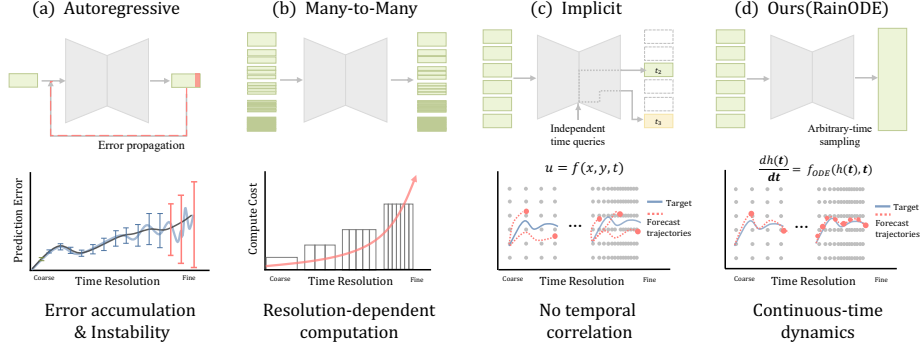


Fig. 1: Task Description. Increasing temporal resolution causes (a) error accumulation, (b) computational bottlenecks, and (c) temporally uncorrelated predictions. In contrast, (d) **RainODE** models shared continuous-time latent dynamics via a Neural ODE, enabling stable long-horizon forecasting and arbitrary-time inference.

and consistency across outputs is not explicitly modeled. Nevertheless, most existing approaches are designed and evaluated on fixed discrete temporal intervals, with both training and inference tied to predefined time steps.

2.2 Neural ODEs in Forecasting

Neural ODEs [5] were originally proposed to integrate ordinary differential equations into deep neural networks by modeling hidden state dynamics in continuous-time. Since then, they have been widely adopted in fluid dynamics and dynamical system modeling, where continuous-time formulations are naturally aligned with physical processes. In meteorology, recent works such as ClimODE [35] and WeatherODE [20] apply Neural ODEs to large-scale atmospheric forecasting. In this context, Neural ODEs are attractive because their continuous-time formulation can, in principle, support temporally flexible forecasting beyond fixed discrete transitions. These approaches demonstrate strong performance in global, spatially closed system settings. However, quantitative analysis under increased temporal resolution, particularly across unseen temporal intervals, has not been explicitly investigated. Moreover, these methods have not been extended to open system, high-resolution precipitation forecasting settings.

2.3 Temporal Downscaling

Temporal downscaling methods aim to reconstruct intermediate frames between discretely sampled observations. Motion-based extrapolation and deep learning interpolation approaches [6, 33, 36] increase temporal density by learning mappings between adjacent timestamps. In practice, these methods generate intermediate states conditioned on observed frames or outputs from a separate

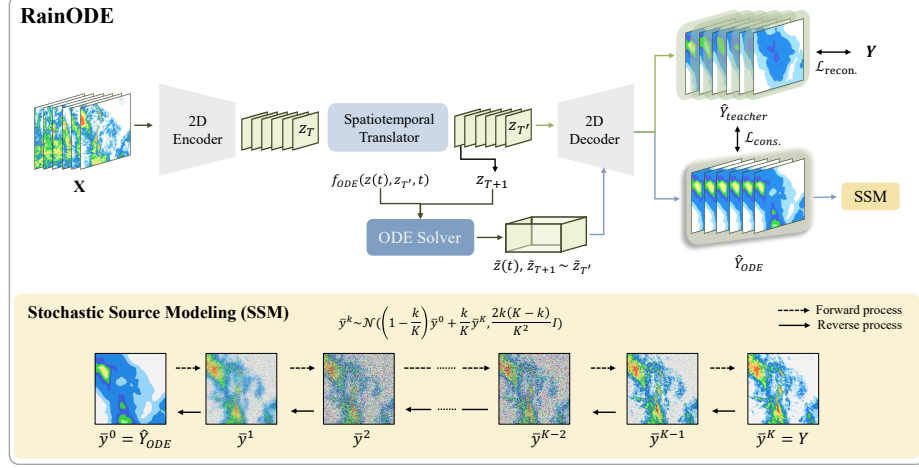


Fig. 2: Overview of RainODE. Input radar sequences are encoded into latent features, which evolve continuously via a Neural ODE conditioned on the target forecasting interval. Consistency loss aligns ODE forecasts $\hat{Y}_{ODE}$ with teacher predictions $\hat{Y}_{teacher}$, while reconstruction loss supervises $\hat{Y}_{teacher}$. The stochastic source modeling (SSM) module further refines predictions using a Brownian Bridge diffusion process.

forecasting model within predefined temporal gaps. As a result, temporal refinement is inherently conditioned on discrete predictions, and overall performance remains strongly dependent on the underlying forecasting model. In contrast, our approach models precipitation evolution itself as a continuous-time dynamical process, intrinsically extending the temporal axis during prediction. This enables time-axis continuous forecasting that is not tied to specific discrete intervals, while promoting spatio-temporal consistency and improving predictive performance.

3 RainODE

3.1 Discrete-Time Forecasting Perspective

Radar-based precipitation forecasting aims to predict future precipitation states $Y \in \mathbb{R}^{T' \times H \times W}$ given past radar observation sequences $X \in \mathbb{R}^{T \times H \times W}$, where T' is the prediction horizon, H and W are the spatial height and width, and T is the number of input frames.

Although existing approaches such as autoregressive, many-to-many, implicit models, and our method differ in their architectural designs, they can be abstracted under a unified learning objective:

$$\min_{\theta} \mathcal{L}(f_{\theta}(X), Y), \quad (1)$$

where f_{θ} denotes a precipitation forecasting model parameterized by θ , and $\mathcal{L}(\cdot)$ measures the discrepancy between predictions and ground truth.

Prior studies mainly improve forecasting performance from two directions: enhancing the model architecture f_θ , such as SimVP [9], MetaVP [31], and Earthformer [11], or designing task-specific loss functions $\mathcal{L}$ [10, 42]. MSE reduces numerical error but often produces overly smooth outputs [9, 11], whereas diffusion-based losses yield sharper fields at the expense of higher pixel-wise errors [10, 12, 44]. Frequency-domain losses such as Fourier Correlation Loss (FCL) emphasize high-frequency components but may introduce intensity biases [42].

Although these approaches improve accuracy under discrete temporal supervision, they do not change the underlying discrete-time formulation.

3.2 Continuous-Time Reformulation

The spatiotemporal evolution of precipitation can be described by a differential equation of the form:

$$\frac{\partial u}{\partial t} = \underbrace{-\mathbf{v} \cdot \nabla u}_{\text{Advection}} + \underbrace{D \nabla^2 u}_{\text{Diffusion / Mixing}} + \underbrace{S(u, t)}_{\text{Source / Sink}} \quad (2)$$

This formulation decomposes precipitation dynamics into three components: large-scale advective transport, diffusion or mixing effects, and local source or sink processes such as growth and decay. Here, u , $\mathbf{v}$, D , and $S(u, t)$ denote precipitation intensity, horizontal transport velocity, diffusion/mixing, and local growth or decay, respectively. Guided by Eq. (2), we focus on the coherent transport-like behavior suggested by the advection term. This transport-like component often evolves smoothly over time, making it compatible with a continuous-time latent dynamics formulation. We therefore use a Neural ODE to parameterize latent temporal evolution in precipitation dynamics.

However, precipitation evolution also involves unresolved processes such as turbulence, local growth, and decay, related to the diffusion and source/sink terms in Eq. (2). These effects introduce stochastic and multi-modal intensity variations that cannot be captured by deterministic latent dynamics alone, often leading to over-smoothed predictions. To account for this uncertainty, we introduce a Brownian Bridge [17, 19]–based conditional refinement that probabilistically models the residual evolution between deterministic forecasts and observed states, restoring high-frequency structures and localized variability (Fig. 2).

Latent ODE Dynamics. We model the deterministic motion component via a Neural ODE in latent space. Given discrete input frames $X_{0:T}$, a 2D spatial encoder $e(\cdot)$ extracts latent representations $Z_{0:T} = \{z_0, \dots, z_T\}$. A Translator network f_{trans} predicts future discrete latent states $Z_{T+1:T'}$.

The ODE is initialized with the first translated latent state: $z(0) = z_{T+1}$. The continuous latent trajectory is parameterized over a normalized temporal variable $s \in [0, S]$, which represents the forecasting interval. The evolution of the latent state is governed by

$$\frac{dz(s)}{ds} = f_{\text{ODE}}(z(s), z(S), s), \quad (3)$$

where s denotes the continuous-time variable. The trajectory is obtained by numerically integrating the ODE using a fourth-order Runge–Kutta (RK4) [24] solver, where τ represents the continuous integration variable used by the solver:

$$\tilde{z}(s) = z(0) + \int_0^s f_{\text{ODE}}(z(\tau), z(S), \tau) d\tau. \quad (4)$$

This formulation enables temporally arbitrary predictions while enforcing derivative-consistent evolution. At inference, querying the trajectory at arbitrary s enables dense forecasts without retraining, and the decoded prediction is

$$\hat{y} = d(\tilde{z}(s)).$$

Consistency-Guided Training. To enforce trajectory-level consistency and stabilize training, we adopt a dual-objective training strategy.

During training, both the discrete latent sequence predicted by the Translator and the continuous latent trajectory obtained by integrating the ODE are projected back to the observation space through the shared 2D spatial decoder $d(\cdot)$. This generates two distinct spatial predictions: the discrete teacher prediction $\hat{Y}_{teacher} = d(Z_{T+1:T'})$ and the continuous ODE prediction $\hat{Y}_{\text{ODE}} = d(\{\tilde{z}(s(t))\}_{t=T+1}^{T'})$.

The total training objective $\mathcal{L}_{\text{Stage I}}$ is defined by two MSE terms. The first component is a task reconstruction loss that directly minimizes the discrepancy between the Translator’s discrete output and the actual ground truth sequence Y . The second component acts as a consistency regularization, enforcing the continuous predictions generated by the ODE solver to align smoothly with the discrete trajectory. The total loss is mathematically formulated as:

$$\mathcal{L}_{\text{Stage I}} = \underbrace{\frac{1}{N} \sum \|\hat{Y}_{teacher} - Y\|_2^2}_{\text{Reconstruction Loss}} + \alpha \underbrace{\frac{1}{N} \sum \|\hat{Y}_{\text{ODE}} - \hat{Y}_{teacher}\|_2^2}_{\text{Consistency Loss}} \quad (5)$$

Stochastic Source Modeling. While Stage I enforces derivative-consistent temporal evolution, its deterministic formulation produces over-smoothed predictions and fails to recover fine-grained variability. To model unresolved components associated with diffusion and source/sink effects in Eq. (2), we introduce stochastic source modeling (SSM) based on a Brownian Bridge diffusion process.

Let the ODE prediction be $\bar{y}^0 = \hat{Y}_{\text{ODE}}$ and the ground truth be $\bar{y}^K = Y$. The forward diffusion process is defined as:

$$\bar{y}^k \sim \mathcal{N}\left(\left(1 - \frac{k}{K}\right)\bar{y}^0 + \frac{k}{K}\bar{y}^K, \frac{2k(K-k)}{K^2}I\right). \quad (6)$$

Except for this endpoint-conditioned formulation, we follow the implementation details of [17]. The mean trajectory gradually bridges the deterministic prediction toward the target state, while the time-dependent variance injects controlled stochastic perturbations. This provides a practical approximation of unresolved precipitation variability without explicitly solving the governing PDE.

Table 1: Precipitation intensity statistics in physical units. Both datasets exhibit highly imbalanced, heavy-tailed distributions, with RAPID showing a wider dynamic range at high intensities.

Dataset	Res.	Zero	50th	90th	99th	Non-zero 99th	99.9th	Max
SEVIR(kg/m ²)	1 km	44%	0.00	0.84	6.38	13.11	26.93	80
RAPID(mm/hr)	2 km	72%	0.00	1.15	8.88	21.28	39.17	180

*SEVIR values are converted to physical units using the decoding formulation provided in the original paper [34].

Through this stochastic refinement, Stage II restores high-frequency spatial structures and localized variability while preserving the large-scale advective consistency learned in Stage I.

4 RAPID Dataset

4.1 Dataset Construction

Precipitation nowcasting research has widely adopted the SEVIR dataset [34], which provides storm-centered multimodal observations under a fixed temporal configuration. To complement existing benchmarks and enable evaluation across multiple temporal scales, we introduce RAPID (Radar-based Precipitation Integrated Dataset), a radar-only benchmark constructed from operational Hybrid Surface Precipitation (HSP) products over the Korean Peninsula¹.

RAPID spans 2022–2025 at 5-minute temporal and 0.5 km spatial resolution. For computational consistency, data are cropped to a fixed domain and resampled to 2 km resolution. Candidate frames are selected if (1) precipitation exceeding 0.1 mm/h occupies more than 10% of the spatial domain, or (2) the 99th percentile intensity exceeds 10 mm/h. This preserves diverse precipitation structures while excluding non-precipitating scenes. Forecasting sequences are constructed using a sliding window with six input and six target frames.

4.2 Dataset Statistics

Intensity Distribution. All statistics are computed in physical units (training split) after inverse decoding. In Tab. 1, RAPID exhibits a heavier upper tail than SEVIR, particularly in extreme rainfall regimes. Although 72% of pixels correspond to zero rainfall, high-percentile intensities are substantially larger. Conditioned on non-zero pixels, the 99th percentile reaches 21.28 mm/h in RAPID compared to 13.11 mm/h in SEVIR, indicating stronger separation among intense precipitation events.

To preserve sensitivity across the full dynamic range during training, rainfall intensity is encoded into $[0, 255]$ using a piecewise transformation.

¹ Those could be found in <https://apihub.kma.go.kr>

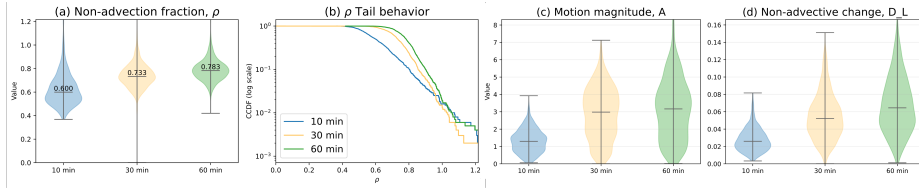


Fig. 3: Temporal dynamics across RAPID. As the time interval increases (10-min, 30-min, and 60-min), the *non-advection fraction* ρ shifts upward, structural evolution plays a progressively larger role relative to displacement at longer horizons. The broader spread at 30 minutes suggests a transition between motion-dominant and evolution-driven regimes.

Temporal Dynamics Across Intervals. RAPID provides forecasting configurations at 10-, 30-, and 60-minute intervals, enabling systematic evaluation of interval-dependent precipitation dynamics.

To quantify how evolution varies with temporal scale, we adopt a motion–evolution decomposition inspired by [14]. Optical flow estimates advective displacement, and residual intensity change after warping measures non-advective evolution. We report motion magnitude (A), non-advective change (D_L), total change (D_E), and the non-advective fraction $\rho = D_L/(D_E + \epsilon)$. As shown in Fig. 3, the non-advective fraction increases consistently with temporal interval. This indicates that rigid displacement explains a decreasing proportion of observed change as lead time extends, while structural evolution becomes increasingly dominant. By providing temporally distinct configurations with systematically varying dynamic characteristics, RAPID establishes a controlled benchmark for evaluating forecasting models under heterogeneous temporal regimes. In particular, it enables assessment of models designed to represent precipitation as continuous temporal evolution rather than fixed-interval frame prediction.

5 Experiments

5.1 Experimental Setting

Dataset. We evaluate our method on SEVIR [34] and RAPID. SEVIR is one of the most widely adopted benchmarks for precipitation nowcasting, enabling direct comparison with prior work. Other datasets [16, 27] share similar lead-time configurations with SEVIR, and therefore do not provide additional variation in temporal structure. Instead, to evaluate forecasting behavior across heterogeneous temporal intervals, we introduce RAPID, which supports 10-, 30-, and 60-minute configurations within a unified benchmark.

- **SEVIR** predicts 60 minutes (12 frames) of future VIL from 70 minutes (13 frames) of context at 384×384 resolution. We follow the standard temporal split, using samples before June 1, 2019 for training and those after for testing.

Table 2: Quantitative results on **SEVIR** and **RAPID** (10-, 30-, 60-minute intervals). All metrics are reported on a 0–1 scale. CSI-th is evaluated at intensity thresholds {16, 160, 219}/255. $\uparrow$ indicates higher is better and $\downarrow$ indicates lower is better. The best score in each column is highlighted in green.

Data	Method	CSI-M $\uparrow$			CSI-th $\uparrow$			RMSE $\downarrow$	FSS $\uparrow$	HSS $\uparrow$	FAR $\downarrow$
		P1	P4	P16	16	160	219				
SEVIR	3D U-Net (MICCAI '16)	0.332	0.334	0.361	0.746	0.188	0.038	0.064	0.461	0.431	0.238
	SimVP (CVPR '22)	0.415	0.419	0.438	0.761	0.290	0.131	0.060	0.554	0.539	0.308
	Earthformer (NeurIPS '22)	0.426	0.425	0.435	0.768	0.300	0.134	0.057	0.565	0.550	0.284
	PreDiff (NeurIPS '23)	0.338	0.365	0.454	0.699	0.200	0.086	0.078	0.486	0.448	0.505
	CasCast (ICML '24)	0.415	0.450	0.540	0.742	0.292	0.160	0.068	0.442	0.542	0.437
	exPreCast (ICLR '26)	0.420	0.452	0.538	0.755	0.296	0.145	0.062	0.557	0.545	0.404
	Latent ODE	0.425	0.426	0.437	0.762	0.302	0.148	0.059	0.558	0.551	0.284
	RainODE	0.430	0.464	0.544	0.766	0.304	0.177	0.060	0.564	0.558	0.399
RAPID-10min	3D U-Net (MICCAI '16)	0.395	0.413	0.470	0.682	0.340	0.130	0.072	0.458	0.529	0.257
	SimVP (CVPR '22)	0.399	0.404	0.444	0.673	0.338	0.178	0.073	0.426	0.537	0.289
	Earthformer (NeurIPS '22)	0.406	0.382	0.374	0.675	0.330	0.159	0.067	0.411	0.542	0.224
	PreDiff (NeurIPS '23)	0.375	0.425	0.521	0.621	0.337	0.140	0.084	0.449	0.509	0.411
	CasCast (ICML '24)	0.399	0.470	0.585	0.640	0.384	0.137	0.084	0.509	0.534	0.462
	exPreCast (ICLR '26)	0.389	0.455	0.568	0.661	0.332	0.165	0.076	0.483	0.526	0.334
	Latent ODE	0.443	0.427	0.417	0.676	0.399	0.197	0.067	0.451	0.583	0.279
	RainODE	0.420	0.491	0.597	0.618	0.416	0.187	0.082	0.514	0.561	0.418
RAPID-30min	3D U-Net (MICCAI '16)	0.197	0.216	0.280	0.485	0.124	0.035	0.099	0.185	0.285	0.469
	SimVP (CVPR '22)	0.224	0.225	0.238	0.512	0.154	0.026	0.101	0.196	0.320	0.458
	Earthformer (NeurIPS '22)	0.161	0.165	0.165	0.462	0.040	0.000	0.097	0.109	0.221	0.452
	PreDiff (NeurIPS '23)	0.138	0.180	0.305	0.339	0.091	0.023	0.128	0.135	0.200	0.702
	CasCast (ICML '24)	0.234	0.290	0.396	0.445	0.215	0.043	0.123	0.234	0.334	0.671
	exPreCast (ICLR '26)	0.234	0.290	0.437	0.485	0.174	0.055	0.101	0.267	0.336	0.561
	Latent ODE	0.234	0.237	0.256	0.517	0.166	0.026	0.093	0.206	0.333	0.429
	RainODE	0.273	0.340	0.458	0.491	0.244	0.080	0.105	0.288	0.391	0.556
RAPID-60min	3D U-Net (MICCAI '16)	0.132	0.146	0.195	0.351	0.080	0.022	0.109	0.131	0.197	0.635
	SimVP (CVPR '22)	0.141	0.145	0.155	0.400	0.061	0.011	0.107	0.130	0.203	0.626
	Earthformer (NeurIPS '22)	0.000	0.000	0.000	0.000	0.000	0.000	0.118	0.068	0.000	0.000
	PreDiff (NeurIPS '23)	0.064	0.085	0.144	0.177	0.036	0.009	0.121	0.092	0.099	0.729
	CasCast (ICML '24)	0.033	0.053	0.112	0.107	0.017	0.003	0.228	0.035	0.006	0.953
	exPreCast (ICLR '26)	0.140	0.172	0.305	0.363	0.079	0.021	0.112	0.128	0.205	0.710
	Latent ODE	0.152	0.155	0.156	0.404	0.072	0.002	0.101	0.137	0.217	0.560
	RainODE	0.220	0.270	0.358	0.400	0.201	0.056	0.113	0.223	0.328	0.596

- **RAPID** evaluates forecasting at 10-, 30-, and 60-minute intervals, corresponding to 1-, 3-, and 6-hour lead times. Each sample consists of six input and six future frames at 224×224 resolution. Data from 2022–2024 are used for training and 2025 for testing.

Evaluation Metrics. We report both categorical and pixel-wise metrics. Critical Success Index (CSI) evaluates event-level detection performance. We report mean CSI (CSI-M) and threshold-based CSI (CSI-th) at 16, 160, and 219, including evaluations under 4×4 and 16×16 pooling to assess robustness to spatial displacement. We further report Heidke Skill Score (HSS) and Fractions Skill Score (FSS) for spatially tolerant evaluation, along with RMSE and FAR to measure pixel-wise intensity error and false alarms.

Implementation Details. All experiments are conducted on a single NVIDIA H200 GPU (140 GB memory). We train all models using the AdamW optimizer

with an initial learning rate of 1×10^{-3} . A cosine annealing scheduler is applied with $T_{\max} = 100$ and $\eta_{\min} = 1 \times 10^{-6}$. The training objective is the mean squared error (MSE) loss and the weighting coefficient for the consistency loss is set to $\alpha = 0.1$. Models are trained for 100 epochs with a batch size of 12 for RAPID and 4 for SEVIR. The checkpoint with the highest CSI-M on the validation set is selected for evaluation. The SSM module is implemented strictly following [17].

All baseline models are trained from scratch in our H200 environment using their official implementations. No pretrained weights are used unless explicitly stated. 3D U-Net, SimVP, and Earthformer are implemented based on the FACL repository², while PreDiff³, CasCast⁴, and exPreCast⁵ strictly follow the official code released by their authors. Implicit forecasting models such as MetNet [28] are excluded due to the absence of official implementations and differences in input modalities.

5.2 Main Results

Quantitative Result. Table 2 presents the quantitative comparison on SEVIR and RAPID (10-, 30-, and 60-minute intervals), evaluating whether RainODE generalizes across both a widely used public benchmark and our proposed multi-interval dataset. As shown in Tab. 2, RainODE achieves strong performance across most evaluation metrics on both datasets, consistently outperforming existing spatiotemporal and generative forecasting baselines across multiple temporal intervals. The performance gap becomes more pronounced on RAPID-60min. RainODE maintains stable categorical and spatial skill scores, whereas most discrete-time baselines degrade significantly. This result suggests that modeling precipitation dynamics in continuous-time improves robustness under longer forecasting horizons.

Notably, Earthformer exhibits near-zero CSI scores at the 60-minute setting, indicating severe performance degradation under extended lead times. All methods are trained under identical settings in our H200 environment using official implementations without pretrained weights, ensuring a controlled and reproducible comparison.

Qualitative Result. These trends are also reflected in the qualitative comparisons shown in Fig. 4, which presents results on SEVIR (10-minute interval) and RAPID-60min. On SEVIR, models trained solely with MSE loss (e.g., 3D U-Net, Earthformer, and SimVP) capture the spatial location of precipitation reasonably well, but often underestimate rainfall intensity. In contrast, generative or diffusion-based models such as PreDiff and CasCast better preserve intensity patterns, yet exhibit noticeable spatial misalignment. Our method combines Neural ODE-based trajectory modeling for spatial evolution with SSM for intensity refinement, resulting in improved alignment of both precipitation location and intensity.

² <https://github.com/argenycw/FACL>

³ <https://github.com/gaozhihan/PreDiff>

⁴ <https://github.com/OpenEarthLab/CasCast>

⁵ <https://github.com/tony890048/exPreCast>

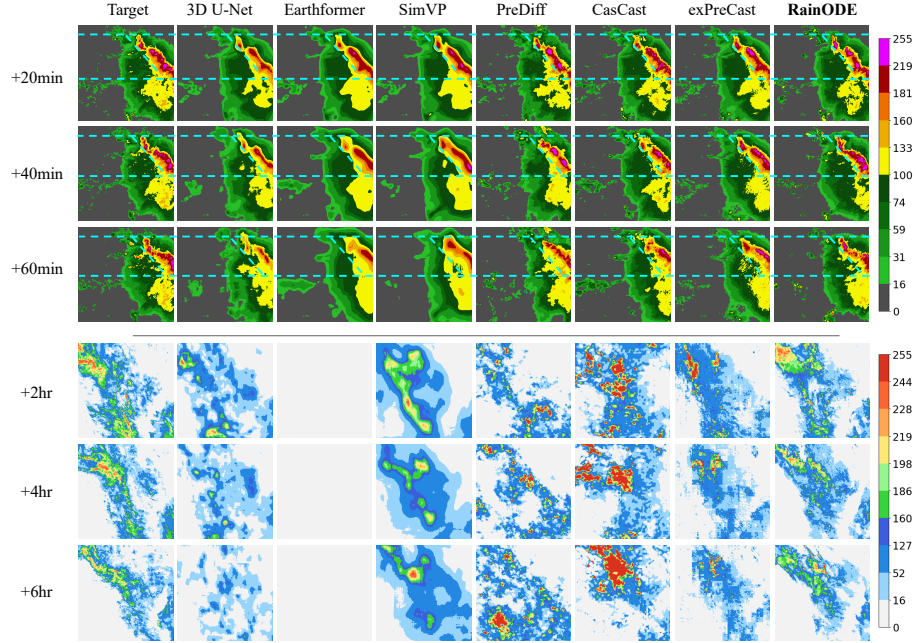


Fig. 4: Qualitative results on **SEVIR** and **RAPID-60min**. (top) SEVIR forecasts at +20, +40, and +60 minutes. (bottom) RAPID-60min forecasts at +2, +4, and +6 hours. Red dashed lines are overlaid to visually assess spatial alignment of predicted precipitation structures. Pixel values are encoded in $[0, 255]$, with higher values corresponding to extreme rainfall.

Table 3: CSI-M performance of **RainODE** across lead times. T denotes the last observed timestep, and $T + 1$ corresponds to the first forecasted frame. Performance gradually degrades as the lead time increases.

Dataset	$T + 1$	$T + 2$	$T + 3$	$T + 4$	$T + 5$	$T + 6$	Lead time
RAPID-10min	0.489	0.459	0.430	0.405	0.382	0.360	1-hour
RAPID-30min	0.360	0.313	0.278	0.250	0.226	0.202	3-hour
RAPID-60min	0.298	0.249	0.221	0.200	0.170	0.144	6-hour

The differences become more pronounced on RAPID-60min. Pure MSE-based models (3D U-Net and SimVP) maintain relatively accurate spatial displacement but struggle with intensity calibration. Earthformer predicts near-uniform low-intensity rainfall across wide regions, leading to visually diminished precipitation structures. PreDiff and CasCast exhibit larger spatial deviations at extended lead times, exPreCast shows visible grid-like artifacts, likely caused by the combination of trilinear interpolation and 3D pixel-shuffle operations within its CDU block.

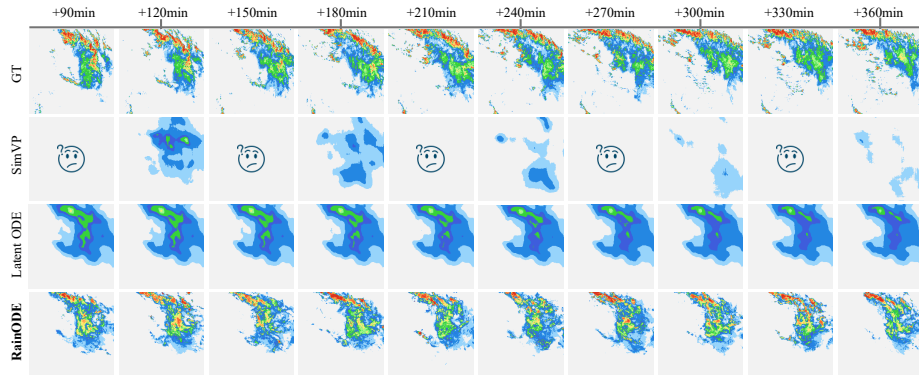


Fig. 5: Continuous-time forecasting up to 6 hours. Discrete forecasting models such as SimVP produce predictions only at predefined timesteps and progressively dissipate precipitation intensity over extended horizons. In contrast, Latent ODE (Stage I of RainODE) and the full RainODE model continuous latent dynamics, enabling predictions at arbitrary time points while maintaining coherent precipitation structures.

Although our method provides the most coherent alignment between location and intensity among the compared models, noticeable errors remain. Forecasting up to +6 hours from radar-only observations remains inherently challenging, highlighting the intrinsic difficulty of long-horizon precipitation prediction.

5.3 Continuous-Time Generalization

Temporal Dynamics Across Lead Times. Table 3 reports the CSI-M performance of our model across increasing lead times. Across all temporal configurations, performance consistently decreases as the lead time increases. This trend is consistent with prior precipitation forecasting studies, where longer lead times inherently introduce greater uncertainty. As the forecast horizon extends, small discrepancies in the predicted motion and intensity evolution accumulate over time. Moreover, precipitation systems exhibit increasingly complex structural changes, including growth, decay, merging, and splitting, which are difficult to model accurately over extended periods.

Consequently, the degradation in CSI-M reflects increasing dynamical uncertainty and structural evolution at longer lead times. Nevertheless, the gradual decline observed in Tab. 3 suggests that our model maintains temporal consistency, which we further examine qualitatively.

Figure 5 qualitatively compares three models trained at 1-hour intervals: SimVP, a variant using only the Latent ODE (without the SSM), and the full RainODE. Since SimVP is trained at fixed 1-hour intervals, it can only produce predictions at those discrete timestamps and does not support intermediate inference. In contrast, RainODE enables continuous-time inference and can generate predictions at arbitrary temporal resolutions between hourly steps.

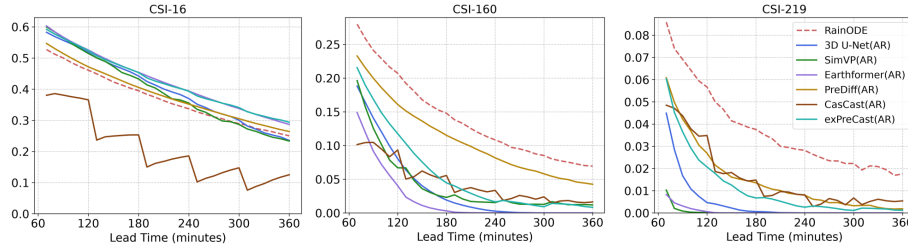


Fig. 6: The robustness of continuous modeling. Performance is measured using CSI-16/160/219 at 10-minute evaluation intervals. RainODE is trained at 1-hour intervals and evaluated at finer temporal resolution without retraining, while baseline models rely on autoregressive (AR) rollout for long-horizon prediction.

The comparison highlights a key difference between discrete and continuous-time forecasting. While SimVP gradually dissipates precipitation structures, RainODE maintains coherent spatial trajectories through continuous latent dynamics. The Latent ODE primarily captures large-scale motion and directional evolution, preserving the overall displacement of precipitation systems. The SSM module further refines high-frequency structures and intensity patterns. As a result, the full RainODE produces forecasts that better preserve precipitation structure and intensity compared to the Latent ODE-only variant.

The Robustness of Continuous Modeling. As illustrated in Fig. 6, we evaluate whether a model trained at coarse temporal intervals can generalize to finer inference resolutions. Specifically, RainODE is trained using 1-hour intervals to predict up to +6 hours, while inference is performed at a 10-minute resolution without any retraining. Discrete baselines are trained at 10-minute intervals and extended to +6 hours through autoregressive (AR) rollout.

As shown in Fig. 6, under the CSI-16 metric, RainODE achieves slightly lower scores compared to AR-based baselines. We attribute this behavior to the tendency of AR models to produce smoother and more spatially diffused predictions, which can artificially increase detection scores at low intensity thresholds. However, this trend reverses at higher thresholds (CSI-160 and CSI-219), and RainODE outperforms the baselines with a noticeable margin.

Importantly, RainODE is trained only with 1-hour interval data, yet it remains competitive when evaluated at 10-minute resolution. This observation suggests that modeling precipitation dynamics in continuous-time represents temporal evolution as a coherent trajectory rather than a sequence of independent predictions. Compared to autoregressive discrete-time models, which accumulate prediction errors through repeated rollouts, RainODE supports both long-horizon forecasting and dense temporal inference within a single framework.

Temporal Consistency in Latent Space. To evaluate the long-horizon, temporally dense, and stable forecasting behavior of Latent ODE, we analyze latent-space trajectories. Figure 7 presents visualizations of latent representations from the same sample used in Fig. 4, drawn from the RAPID-60min setting. All la-

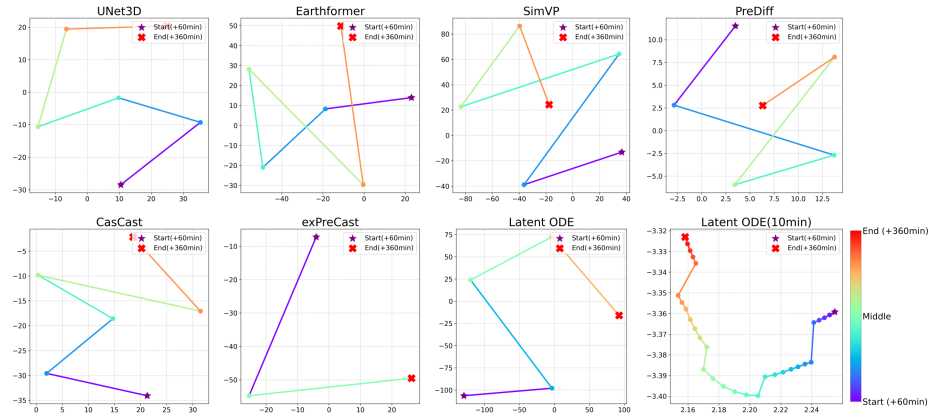


Fig. 7: Temporal consistency of latent trajectories on RAPID-60min. Latent ODE (Stage I of RainODE) is additionally evaluated at 10-minute intervals to visualize the continuous-time evolution. Colors indicate temporal progression, with warmer colors approaching the final prediction time (+360 minutes).

tent states from the compared models are jointly embedded using t-SNE to illustrate their temporal progression in the learned representation space. Among the compared models, Latent ODE shows a more coherent progression from the prediction start point to the final horizon, with fewer abrupt directional changes across forecast steps. In contrast, the latent trajectories of the baseline models are more irregular and often show zigzag patterns, suggesting less coherent temporal progression in latent space. ExPreCast shows a relatively sparse trajectory, which may be attributed to the compression of temporal information in its latent representation. Beyond the original 1-hour prediction setting, Latent ODE also maintains coherent latent transitions when evaluated at 10-minute intervals without additional training. While these visualizations are qualitative, they indicate that Latent ODE preserves temporal consistency in latent evolution more effectively than discrete-time baselines and supports temporally dense forecasting over extended horizons.

5.4 Ablations

Table 4 evaluates the contribution of each stage in RainODE. When integrating SimVP with the Latent ODE alone, CSI-16 and RMSE improve, indicating that continuous deterministic dynamics effectively capture the overall precipitation motion. However, this produces smoother predictions and does not improve CSI-219, suggesting that it is insufficient for concentrated heavy-rainfall cores. By further incorporating SSM, RainODE substantially improves high-intensity detection (CSI-219) and perceptual quality metrics (LPIPS and SSIM), demonstrating its role in restoring high-frequency and intensity-sensitive structures. Although this refinement leads to a small trade-off in RMSE and CSI-16, this

Table 4: Ablation study on the proposed framework. (a) Importance of each stage integrating SimVP, Latent ODE, and SSM(RainODE). (b) Computational cost in parameters (M) and FLOPs (G). L_{out} denotes the length of output sequence.

(a) Importance of each stage								(b) Computational cost			
Method	CSI-M $\uparrow$	CSI-16 $\uparrow$	CSI-219 $\uparrow$	RMSE $\downarrow$	FSS $\uparrow$	LPIPS $\downarrow$	SSIM $\uparrow$	Model	L_{out}	Params	FLOPs
SimVP	0.141	0.400	0.011	0.107	0.130	0.346	0.580	Latent ODE	6	9.32	68.84
Latent ODE	0.152	0.404	0.002	0.101	0.137	0.376	0.588	Latent ODE	36	9.32	104.97
RainODE	0.220	0.400	0.056	0.113	0.223	0.275	0.690	SimVP	6	13.62	90.91
								SimVP	36	15.25	129.78

is consistent with the role of SSM, which prioritizes sharper and more localized precipitation structures.

From a computational perspective, the proposed framework remains efficient. The Latent ODE model maintains the same number of parameters even when the prediction length increases. Also, FLOPs remain lower than SimVP, although they increase with additional integration steps. In contrast, SimVP requires more parameters as the output horizon expands and generally exhibits higher FLOPs.

6 Limitations and Future Work

We introduced RainODE, which enables inference at unseen temporal intervals through continuous-time modeling. Although it generalizes well across resolutions, our validation is limited to radar-only forecasting. As shown in the RAPID-60min results, long-term prediction from radar alone has intrinsic limitations. With increasing lead time, precipitation evolution depends more on large-scale atmospheric dynamics that cannot be fully captured from radar observations. This fundamentally constrains performance and makes it difficult to predict precipitation that emerges from completely clear conditions. In future work, we aim to extend RainODE to multi-modal settings, incorporating additional atmospheric variables to improve long-horizon stability and structural consistency.

7 Conclusion

We presented RainODE, a continuous-time precipitation forecasting framework that reformulates frame-based prediction as latent trajectory modeling. By parameterizing temporal evolution through a Neural ODE with stochastic refinement, our approach enables flexible and stable forecasting across varying temporal intervals. Experiments on SEVIR and RAPID demonstrate consistent gains at high-intensity thresholds, improved structural coherence, and stronger long-horizon stability compared to discrete recursive methods. These results suggest that continuous-time latent modeling offers a promising paradigm for precipitation forecasting beyond fixed temporal grids.

References

1. An, S., Oh, T.J., Sohn, E., Kim, D.: Deep learning for precipitation nowcasting: A survey from the perspective of time series forecasting. *Expert Systems with Applications* **268**, 126301 (2025)
2. Andrychowicz, M., Espeholt, L., Li, D., Merchant, S., Merose, A., Zyda, F., Agrawal, S., Kalchbrenner, N.: Deep learning for day forecasts from sparse observations. *arXiv preprint arXiv:2306.06079* (2023)
3. Bai, C., Sun, F., Zhang, J., Song, Y., Chen, S.: Rainformer: Features extraction balanced network for radar-based precipitation nowcasting. *IEEE Geoscience and Remote Sensing Letters* **19**, 1–5 (2022)
4. Chen, L., Du, F., Hu, Y., Wang, Z., Wang, F.: Swinrdm: integrate swinrnn with diffusion model towards high-resolution and high-quality weather forecasting. In: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*. vol. 37, pp. 322–330 (2023)
5. Chen, R.T., Rubanova, Y., Bettencourt, J., Duvenaud, D.K.: Neural ordinary differential equations. *Advances in Neural Information Processing Systems (NeurIPS)* **31** (2018)
6. Demiray, B.Z., Sit, M., Demir, I.: Efficienttempnet: Temporal super-resolution of radar rainfall. *International Conference on Learning Representations (ICLR) 2023 Workshop on Tackling Climate Change with Machine Learning* (2023)
7. Espeholt, L., Agrawal, S., Sønderby, C., Kumar, M., Heek, J., Bromberg, C., Gazen, C., Carver, R., Andrychowicz, M., Hickey, J., et al.: Deep learning for twelve hour precipitation forecasts. *Nature communications* **13**(1), 5145 (2022)
8. Feng, W., Li, X., Wu, Z., Lin, K., Yu, D., Ye, Y., Wang, Y.: Perceptually constrained precipitation nowcasting model. In: *Proceedings of the International Conference on Machine Learning (ICML)* (2025)
9. Gao, Z., Tan, C., Wu, L., Li, S.Z.: Simvp: Simpler yet better video prediction. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 3170–3180 (2022)
10. Gao, Z., Shi, X., Han, B., Wang, H., Jin, X., Maddix, D., Zhu, Y., Li, M., Wang, Y.B.: Prediff: Precipitation nowcasting with latent diffusion models. *Advances in Neural Information Processing Systems (NeurIPS)* **36**, 78621–78656 (2023)
11. Gao, Z., Shi, X., Wang, H., Zhu, Y., Wang, Y.B., Li, M., Yeung, D.Y.: Earthformer: Exploring space-time transformers for earth system forecasting. *Advances in Neural Information Processing Systems (NeurIPS)* **35**, 25390–25403 (2022)
12. Gong, J., Bai, L., Ye, P., Xu, W., Liu, N., Dai, J., Yang, X., Ouyang, W.: Cast: Skillful high-resolution precipitation nowcasting via cascaded modelling. In: *Proceedings of the International Conference on Machine Learning (ICML)* (2024)
13. Guen, V.L., Thome, N.: Disentangling physical dynamics from unknown factors for unsupervised video prediction. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 11474–11484 (2020)
14. Kim, D., Seo, M., Choi, Y.: Ace metric: Advection and convection evaluation for accurate weather forecasting. *arXiv preprint arXiv:2406.04678* (2024)
15. Kim, D., Seo, M., Lee, H., Seo, J.: Masked autoregressive model for weather forecasting. *arXiv preprint arXiv:2409.20117* (2024)
16. Larvor, G., Berthomier, L., Chabot, V., Le Pape, B., Pradel, B., Perez, L.: Meteonet, an open reference weather dataset by meteo france. *Technical Report* (2020), <https://meteofrance.github.io/meteonet/>

17. Li, B., Xue, K., Liu, B., Lai, Y.K.: Bbdm: Image-to-image translation with brownian bridge diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1952–1961 (2023)
18. Lin, K., Zhang, B., Yu, D., Feng, W., Chen, S., Gao, F., Li, X., Ye, Y.: Alphapre: Amplitude-phase disentanglement model for precipitation nowcasting. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 17841–17850 (2025)
19. Liu, G.H., Vahdat, A., Huang, D.A., Theodorou, E.A., Nie, W., Anandkumar, A.: I2sb: Image-to-image schrodinger bridge. In: Proceedings of the International Conference on Machine Learning (ICML) (2023)
20. Liu, P., Zhou, T., Sun, L., Jin, R.: Mitigating time discretization challenges with weatherode: A sandwich physics-driven neural ode for weather forecasting. arXiv preprint arXiv:2410.06560 (2024)
21. Ning, S., Lan, M., Li, Y., Chen, C., Chen, Q., Chen, X., Han, X., Cui, S.: Mimo is all you need: A strong multi-in-multi-out baseline for video prediction. In: Proceedings of the AAAI Conference on Artificial Intelligence (AAAI). vol. 37, pp. 1975–1983 (2023)
22. Otsuka, S., Tuerhong, G., Kikuchi, R., Kitano, Y., Taniguchi, Y., Ruiz, J.J., Satoh, S., Ushio, T., Miyoshi, T.: Precipitation nowcasting with three-dimensional space–time extrapolation of dense and frequent phased-array weather radar observations. *Weather and Forecasting* **31**(1), 329–340 (2016)
23. Ribeiro, B.P., Pucer, J.F.: Flowcast: Advancing precipitation nowcasting with conditional flow matching. In: International Conference on Learning Representations (ICLR) (2026)
24. Runge, C.: Über die numerische auflösung von differentialgleichungen. *Mathematische Annalen* **46**(2), 167–178 (1895)
25. Sarabia, R.P., Nyborg, J., Birk, M., Sjørup, J.L., Vesterholt, A.L., Assent, I.: Rainpro-8: An efficient deep learning model to estimate rainfall probabilities over 8 hours. In: International Conference on Learning Representations (ICLR) (2026)
26. Shi, X., Chen, Z., Wang, H., Yeung, D.Y., Wong, W.K., Woo, W.c.: Convolutional lstm network: A machine learning approach for precipitation nowcasting. *Advances in Neural Information Processing Systems (NeurIPS)* **28** (2015)
27. Shi, X., Gao, Z., Lausen, L., Wang, H., Yeung, D.Y., Wong, W.k., Woo, W.c.: Deep learning for precipitation nowcasting: A benchmark and a new model. *Advances in Neural Information Processing Systems (NeurIPS)* **30** (2017)
28. Sønderby, C.K., Espenholt, L., Heek, J., Dehghani, M., Oliver, A., Salimans, T., Agrawal, S., Hickey, J., Kalchbrenner, N.: Metnet: A neural weather model for precipitation forecasting. arXiv preprint arXiv:2003.12140 (2020)
29. Song, C., Chang, T.Y., Hong, Y.: Extreme weather nowcasting via local precipitation pattern prediction. In: International Conference on Learning Representations (ICLR) (2026)
30. Sun, H., Yang, Y., Han, W., Huang, W., Chen, H., Gao, Z., Li, Z., Huo, Z., Niu, Z.: Stormdit: A generative ai model bridges the 2-6 hour ‘gray zone’ in precipitation nowcasting. arXiv preprint arXiv:2601.20342 (2026)
31. Tan, C., Gao, Z., Li, S., Li, S.Z.: Simvpv2: Towards simple yet powerful spatiotemporal predictive learning. *IEEE Transactions on Multimedia* (2025)
32. Tang, Y., Qi, L., Xie, F., Li, X., Ma, C., Yang, M.H.: Video prediction transformers without recurrence or convolution. *Transactions on Machine Learning Research* (2026)

33. Tatsubori, M., Moriyama, T., Ishikawa, T., Fraccaro, P., Jones, A., Edwards, B., Kuehnert, J., Remy, S.L.: Deep temporal interpolation of radar-based precipitation. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1685–1689. IEEE (2022)
34. Veillette, M., Samsi, S., Mattioli, C.: Sevir: A storm event imagery dataset for deep learning applications in radar and satellite meteorology. *Advances in Neural Information Processing Systems (NeurIPS)* **33**, 22009–22019 (2020)
35. Verma, Y., Heinonen, M., Garg, V.: Climode: Climate and weather forecasting with physics-informed neural odes. In: International Conference on Learning Representations (ICLR) (2024)
36. Wang, H., Zeng, Q., Wang, H., He, J., Yu, T., Liu, G.: Temporal super-resolution reconstruction of weather radar echoes using a deep learning approach. *Expert Systems with Applications* p. 130189 (2025)
37. Wang, Y., Gao, Z., Long, M., Wang, J., Yu, P.S.: Predrnn++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning. In: Proceedings of the International Conference on Machine Learning (ICML). pp. 5123–5132 (2018)
38. Wang, Y., Jiang, L., Yang, M.H., Li, L.J., Long, M., Fei-Fei, L.: Eidetic 3d lstm: A model for video prediction and beyond. In: International Conference on Learning Representations (ICLR) (2019)
39. Wang, Y., Long, M., Wang, J., Gao, Z., Yu, P.S.: Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms. In: *Advances in Neural Information Processing Systems (NeurIPS)*. vol. 30 (2017)
40. Wang, Y., Zhang, J., Zhu, H., Long, M., Wang, J., Yu, P.S.: Memory in memory: A predictive neural network for learning higher-order non-stationarity from spatiotemporal dynamics. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 9154–9162 (2019)
41. Wen, Y., Schuur, T., Vergara, H., Kuster, C.: Effect of precipitation sampling error on flash flood monitoring and prediction: Anticipating operational rapid-update polarimetric weather radars. *Journal of Hydrometeorology* **22**(7), 1913–1929 (2021)
42. Yan, C.W., Foo, S.Q., Trinh, V.H., Yeung, D.Y., Wong, K.H., Wong, W.K.: Fourier amplitude and correlation loss: Beyond using l2 loss for skillful precipitation nowcasting. *Advances in Neural Information Processing Systems (NeurIPS)* **37**, 100007–100041 (2024)
43. Yoon, D., Seo, M., Kim, D., Choi, Y., Cho, D.: Probabilistic weather forecasting with deterministic guidance-based diffusion model. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. pp. 108–124. Springer (2024)
44. Yu, D., Li, X., Ye, Y., Zhang, B., Luo, C., Dai, K., Wang, R., Chen, X.: Diffcast: A unified framework via residual diffusion for precipitation nowcasting. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 27758–27767 (2024)
45. Yu, W., Lu, Y., Easterbrook, S., Fidler, S.: Efficient and information-preserving future frame prediction and beyond. In: International Conference on Learning Representations (ICLR) (2020)
46. Zhao, Z., Wang, Z., Zhao, G., Zhao, J.: A new strong convective precipitation forecasting method based on attention mechanism and spatio-temporal reasoning. *Scientific Reports* **14**(1), 19024 (2024)