SDRNET: COUPLED RAINFALL-STATE EVOLUTION FOR MULTI-STEP TROPICAL CYCLONE PRECIPITATION FORECASTING

Anonymous authors
Paper under double-blind review

ABSTRACT

Tropical cyclone rainfall fields undergo spatial deformation and local intensity changes beyond the motion of the storm center. Multi-step forecasting must therefore respond to nonuniform evolution while retaining spatial skill at later lead times. Direct regression of rainfall fields leaves the evolution relation implicit in the output mapping, whereas explicit net-residual updates do not separately coordinate the weights assigned to existing rainfall and additions within each step. We propose the Storm-relative Deformation–Survival–Renewal Network (SDRNet), which treats the current rainfall field as an explicit state and constructs a coupled survival–renewal recurrence derived from closed-form integration along characteristics. The model redistributes existing rain areas through backward warping, uses shared decay to coordinate exponential retention and nonnegative effective renewal, and feeds predicted rainfall fields back to refresh structural conditions and the hidden state for subsequent evolution. In four-step forecasting over 24 h on the Corrected and Raw rainfall products of LagTCPMoS, using given retrospective best tracks, SDRNet improves the aggregate fractions skill score (FSS) by approximately 6% over the best external comparison method on both products and achieves the highest critical success index at 20 mm/6 h (CSI20). At +24 h, it retains 93.5% and 93.7% of its first-step FSS, respectively. Comparisons using the same backbone and analyses of rainfall evolution further support the roles of explicit state updates and dynamic feedback in retaining spatial skill across forecast steps.

1 INTRODUCTION

Accurately forecasting the distribution and evolution of tropical cyclone rainfall is essential for assessing the spatial extent of heavy rainfall and the persistence of its associated risks. Muller et al. (2026) analyzed direct fatalities from Atlantic tropical cyclones in the continental United States and nearby coastal waters during 1963–2024. Rainfall-related freshwater flooding accounted for 36% of these fatalities, exceeding the 33% attributed to storm surge. The center track determines the storm's position, while the surrounding rain areas continue to change in shape and intensity. Multi-step rainfall forecasting must therefore identify where heavy rainfall will occur and maintain the quality of spatial predictions as the rainfall field evolves.

Nonuniform rainfall evolution requires spatial redistribution and local intensity adjustment to be handled jointly. Even after alignment by the storm center, the Florence example in Figure 1 exhibits a contraction of heavy-rainfall coverage alongside an increase in the local peak (Pérez-Alarcón et al., 2025). These differing changes in area and peak intensity cannot be adequately characterized by uniform translation or intensity scaling; simply persisting the initial rainfall field also misses subsequent adjustments. Forecasting must allow existing spatial structures to continue contributing to future states as their location and intensity change.

Spatiotemporal representation learning captures complex rainfall evolution through mappings from historical to future fields. SimVP organizes sequence information through convolutional encoding and temporal modeling (Gao et al., 2022b), while Earthformer captures dependencies across locations and times through local spatiotemporal attention and global interactions (Gao et al., 2022a). These methods improve representations of motion and intensity changes, but how the current rainfall forms

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

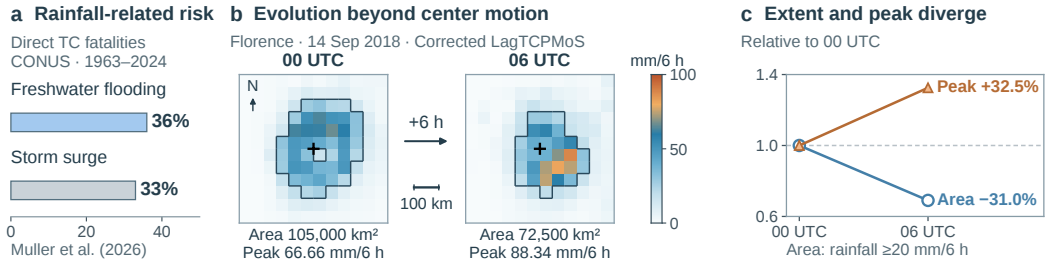

Figure 1: **Tropical cyclone rainfall risk and nonuniform evolution beyond center motion.** (a) Shares of direct fatalities from Atlantic tropical cyclones in the continental United States and nearby coastal waters attributable to rainfall-related freshwater flooding and storm surge (Muller et al., 2026). (b) Corrected LagTCPMoS rainfall fields for Florence at 00 and 06 UTC on 14 September 2018, each aligned by the storm center. Dark boundaries delineate heavy-rainfall areas with rainfall of at least 20 mm/6 h (Pérez-Alarcón et al., 2025). (c) Changes in heavy-rainfall area and local peak relative to 00 UTC, showing simultaneous contraction in coverage and intensification of the peak. The displayed region extends $\pm$300 km from the center; statistics are computed over the valid circular domain with a radius of 500 km.

the next state remains largely implicit in the output mapping. Making the update relation of the rainfall field explicit allows existing structures to contribute directly to successive predictions.

Explicit evolution further applies spatial motion and intensity changes directly to the current rainfall field. NowcastNet constructs an evolution field by combining advection with an intensity residual (Zhang et al., 2023). LUPIN predicts net increments after a Lagrangian transformation and feeds generated rainfall fields back during rollout to re-estimate motion (Pavlík et al., 2025). While local intensity can be represented through net changes, rainfall retained over the full step and additions accumulated within the step have different temporal weights. Coordinating these two contributions within a single state relation, while updating subsequent conditions from the current rain areas, is central to further structuring multi-step evolution.

We therefore propose the Storm-relative Deformation–Survival–Renewal Network (SDRNet), which constructs an explicit rainfall recurrence through integration along characteristics of a storm-relative transport–reaction relation. The model first adjusts the position and shape of existing rain areas through backward warping, then uses shared decay to coordinate the retention of existing rainfall and local additions. The new prediction refreshes the structural conditions and hidden state and becomes the rainfall field updated at the next step. This design connects nonuniform changes within a step with the propagation of contributions across steps, allowing the model to retain spatial skill at later lead times while continuously adjusting the rain areas.

Our main contributions are as follows:

1. We introduce SDRNet, which explicitly organizes the evolution from the current rainfall field to future states through storm-relative deformation and survival–renewal coupling derived from integration along characteristics.
2. We develop feedback that refreshes structural conditions and the hidden state from predicted rainfall fields, allowing each generated spatial distribution to inform both subsequent state updates and the estimation of evolution quantities.
3. We improve multiscale spatial matching and heavy-rainfall identification in four-step, 24 h forecasting on both LagTCPMoS products and retain higher spatial skill at longer lead times. Comparisons using the same backbone further support the predictive value of the update parameterization and dynamic conditioning.

2 Related Work

Tropical cyclone rainfall forecasting and data resources. LagTCPMoS provides tropical cyclone rainfall estimates and moisture-source data, together with a corrected rainfall product constructed

using MSWEP as a reference (Pérez-Alarcón et al., 2025). TCP-Diffusion integrates historical rainfall, environmental information, and numerical weather predictions, reconstructing future sequences through residuals between adjacent rainfall fields and accumulation of the resulting increments (Huang et al., 2025). These studies advance tropical cyclone rainfall forecasting through data resources, multiple conditioning sources, and change modeling, providing complementary foundations for studying rainfall evolution.

Spatiotemporal representations and latent dynamics. ConvLSTM (Shi et al., 2015), SimVP (Gao et al., 2022b), and Earthformer (Gao et al., 2022a) learn historical-to-future dependencies through convolutional recurrence, convolutional sequence mappings, and spatiotemporal attention, respectively. PhyDNet further combines a dynamics branch constrained by partial differential equations with a ConvLSTM residual branch to organize evolution in latent space (Le Guen & Thome, 2020). This line of work focuses on enhancing spatiotemporal representations or constraining hidden-state dynamics. SDRNet instead concentrates its structural design on the computational relation between the current rainfall field of a product and the next field.

Explicit rainfall evolution. Pysteps provides a Lagrangian nowcasting workflow that combines cascade decomposition, autoregression at individual scales, and stochastic perturbations (Pulkkinen et al., 2019). NowcastNet constructs an evolution field by combining motion with intensity residuals, then uses it to guide conditional generation (Zhang et al., 2023). LUPIN uses a differentiable Lagrangian transformation to predict net source–sink increments after transport and feeds generated rainfall fields back into subsequent predictions (Pavlík et al., 2025). SDRNet links exponential retention of the warped field with the accumulation of nonnegative renewal through shared decay, then uses dynamic conditioning to organize subsequent evolution. Contribution propagation and local adjustment are thus incorporated into the same state recurrence.

3 Problem Formulation

We study multi-step rainfall forecasting conditioned on a given storm track. Let $k = 0$ denote the last historical time, and let Δt, L, and K denote the time interval, history length, and number of forecast steps, respectively. At each time, a local north-up grid is centered on the storm. The field $P_k \in \mathbb{R}_{\geq 0}^{H_g \times W_g}$ represents rainfall accumulated over the corresponding time interval, where H_g, W_g are the spatial dimensions. The inputs are the historical rainfall fields $P_{-L+1:0}$ and the track $\mathbf{T}_{-L+1:K}$ covering both the historical and forecast periods. Here, $\mathbf{T}_k \in \mathbb{R}^2$ represents the east–west and north–south offsets of the storm center relative to the forecast origin.

The track and grid jointly determine the geometry $G_k = (a_k, M_k)$, where $a_{k,i}$ is the effective area of cell i within the forecast domain and $M_{k,i} = \mathbf{1}[a_{k,i} > 0]$ is the valid-domain mask. The forecasting task is

$$\begin{aligned} \widehat{P}_{1:K} &= f_\theta(P_{-L+1:0}, \mathbf{T}_{-L+1:K}, G_{-L+1:K}), \\ \widehat{P}_k &= M_k \odot \widehat{P}_k \geq 0, \qquad k = 1, \ldots, K. \end{aligned} \tag{1}$$

Here, $\odot$ denotes elementwise multiplication. Predictions and targets lie on the same storm-centered grid at each corresponding time and cover all rainfall intensities within the valid domain. Future centers and grid geometries are determined by the given track; future rainfall fields are used only for supervision and evaluation. The rainfall products, spatial scale, and forecast horizon are specified in the experimental setup.

4 Methodology

SDRNet treats the current rainfall field as an explicit state and learns backward displacement $\mathbf{d}_k$, accumulated decay $\Lambda_k \geq 0$, and effective renewal $Q_k \geq 0$. Starting from $\widehat{P}_0 = M_0 \odot P_0$, it applies the same update relation at each step:

$$\widehat{P}_{k+1} = M_{k+1} \odot \left[\underbrace{e^{-\Lambda_k} \odot \mathcal{W}(M_k \odot \widehat{P}_k; \mathbf{d}_k)}_{\substack{\text{retention of existing} \\ \text{rainfall after warping}}} + \underbrace{\phi_1(-\Lambda_k) \odot Q_k}_{\substack{\text{accumulation of within-step} \\ \text{effective renewal}}} \right]. \tag{2}$$

Here, $k = 0, \ldots, K-1$, $\mathcal{W}$ is a bilinear backward-warping operator, and $\phi_1(-z) = \int_0^1 e^{-(1-u)z}\,\mathrm{d}u$. Shared decay coordinates the two contributions, while the new prediction enters both the rainfall state and the conditions for evolution at the next step (Figure 2).

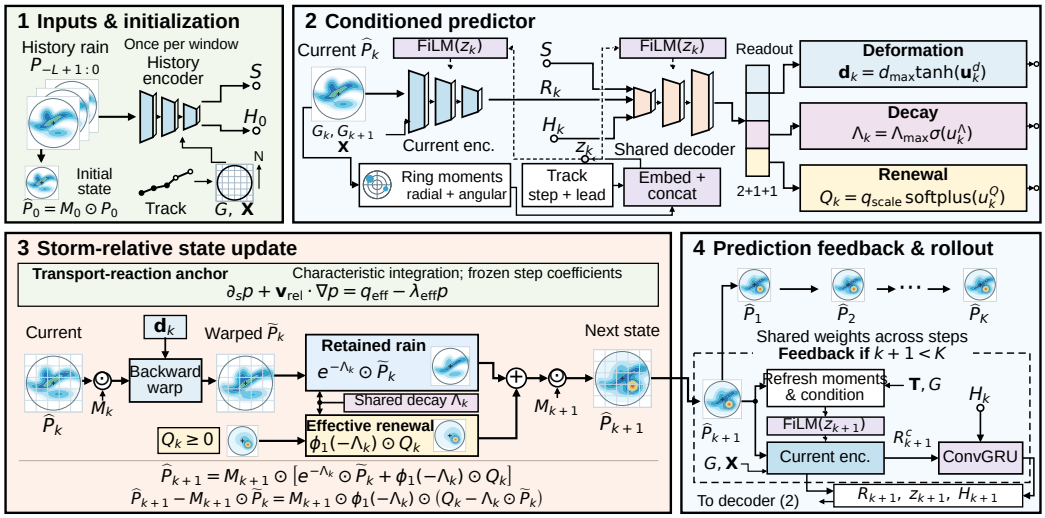

Figure 2: **Explicit state updates and prediction feedback in SDRNet.** Historical encoding produces fixed multiscale reference features S and an initial hidden state H_0. The current rainfall structure, track, and lead time form the conditioning vector z_k, which modulates the current encoder and shared decoder to generate displacement, decay, and effective renewal. The retention term of the warped rainfall field and the integral-weighted renewal contribution are added to produce the next prediction. This prediction then refreshes the structural conditions and hidden state for subsequent steps. Network parameters are shared across steps; the rainfall fields and track are computational illustrations.

4.1 Storm-Relative Deformation

After center alignment, rain areas can still drift and deform relative to the center. SDRNet establishes spatial correspondences between the current rainfall field and the next grid through cellwise backward sampling. For a cell center $\mathbf{x}_i$,

$$\widetilde{P}_{k,i} = \mathcal{W}(M_k \odot \widehat{P}_k; \mathbf{d}_k)_i = \sum_j b_j(\mathbf{x}_i - \mathbf{d}_{k,i}) M_{k,j} \widehat{P}_{k,j}, \tag{3}$$

where b_j are bilinear weights. The learned displacement describes relative changes beyond center motion, allowing both overall drift and local deformation of rain areas. The warped field $\widetilde{P}_k$ thereby provides the spatial basis for the next state. Displacement parameterization and boundary handling are described in Appendices C.1 and C.4.

4.2 Survival–Renewal Coupling

The intensity of the warped rainfall field changes through the combined effects of retaining existing rainfall and adding rainfall locally. To relate their weights, we adopt an effective transport–reaction relation in rainfall-product space:

$$\partial_s p + \mathbf{v}_{\mathrm{rel}} \cdot \nabla p = q_{\mathrm{eff}} - \lambda_{\mathrm{eff}} p, \qquad q_{\mathrm{eff}}, \lambda_{\mathrm{eff}} \geq 0, \tag{4}$$

where s is dimensionless step time, $\Delta s = 1$ corresponds to one forecast interval, and the coefficients describe effective changes in the product's rainfall field. Along a characteristic ending at $\mathbf{x}_i$, we freeze the coefficients within a step and write $\Lambda = \bar{\lambda}\Delta s$ and $Q = \bar{q}\Delta s$. Rainfall over normalized within-step time $u \in [0, 1]$ then satisfies $y' = Q - \Lambda y$. The integrating factor gives

$$y(1) = \underbrace{e^{-\Lambda} y(0)}_{\substack{\text{existing rainfall decays over} \\ \text{the full step}}} + \underbrace{\int_0^1 e^{-(1-u)\Lambda} Q \, \mathrm{d}u}_{\substack{\text{within-step additions decay} \\ \text{over the remaining interval}}} = e^{-\Lambda} y(0) + \phi_1(-\Lambda) Q. \tag{5}$$

Existing rainfall and within-step additions enter the evolution at different times. They therefore receive exponential and integral weights, respectively, while sharing the same decay quantity. Approximating rainfall at the characteristic's starting point using Eq. (3) and applying the target mask yields Eq. (2). Here, Λ is dimensionless and Q has the same units as the rainfall field. The net change at each location is governed by $Q_k - \Lambda_k \widetilde{P}_k$, allowing local intensification and weakening to occur simultaneously. The general integral form, nonnegativity, and parameterization properties are provided in Appendix A.

4.3 Feedback of Structural Conditions and Multi-Step Propagation

SDRNet retains the initial spatial reference in fixed historical features S, represents the current distribution in the explicit rainfall field $\widehat{P}_k$, and accumulates recurrent context in the hidden state H_k. The history encoder produces S, H_0, while the current encoder extracts R_k at each step. To summarize the current rain areas, we use effective-area-weighted annular means and first- and second-order azimuthal components, combined with the track direction, to form the structural summary $\boldsymbol{\xi}_k = \Gamma(\widehat{P}_k, a_k, \Delta\mathbf{T}_k)$. The full track, structural summary, description of the current track increment $\boldsymbol{\nu}_k$, and lead time are embedded separately to construct the conditioning vector:

$$\begin{aligned} z_k &= \mathrm{Concat}[e_{\mathrm{trk}}(\mathbf{T}_{-L+1:K}), e_{\mathrm{str}}(\boldsymbol{\xi}_k), e_{\mathrm{step}}(\boldsymbol{\nu}_k), e_{\mathrm{lead}}(k)], \\ R_k &= \mathcal{E}_{\mathrm{cur}}\Big(g(\widehat{P}_k), G_k, G_{k+1}, \mathbf{X}; z_k\Big), \\ (\mathbf{u}_k^d, u_k^\Lambda, u_k^Q) &= \mathcal{D}_\theta(S, R_k, H_k; z_k), \\ H_{k+1} &= \mathrm{ConvGRU}\big(R_{k+1}^{\mathrm{coarse}}, H_k\big), \qquad k+1 < K. \end{aligned} \tag{6}$$

Here, $\Delta\mathbf{T}_k = \mathbf{T}_{k+1} - \mathbf{T}_k$, $\mathbf{X}$ denotes fixed grid coordinates, and g is a logarithmic transform applied to the encoder input. The conditioning vector modulates the current encoder and shared decoder through FiLM. Raw outputs are parameterized as bounded displacement, nonnegative decay, and nonnegative renewal before being substituted into Eq. (2). After generating the new rainfall field, the model first refreshes its structural summary and conditioning vector, then computes R_{k+1} and updates the hidden state. The new prediction therefore provides the spatial state for the next step and changes the conditions under which subsequent evolution quantities are generated. The model feeds back its own predictions during both training and inference. Full configurations and algorithms are provided in Appendices B and C, and the expansion of contribution propagation across steps is given in Appendix A.3.

4.4 Storm-Balanced Multiscale Learning

We use a full-storm soft fractions skill score (soft-FSS) objective weighted by effective area to jointly learn rain-area matching across lead times, thresholds, and spatial scales. For a window w and component $c = (j, \tau, r)$, let $\widehat{F}_w^c, F_w^c$ denote the local event fractions of the prediction and target. During training, predictions use a smooth threshold and targets use a hard threshold; neighborhood fractions are normalized by effective area. Let $\mathcal{W}_s$ denote all valid windows of storm s, and let $\mathcal{S}_c$ be the set of training storms with valid target events at the corresponding lead time and threshold. The objective is

$$\mathcal{L}(\theta) = \frac{1}{|\mathcal{C}|} \sum_{c \in \mathcal{C}} \frac{1}{|\mathcal{S}_c|} \sum_{s \in \mathcal{S}_c} \frac{\sum_{w \in \mathcal{W}_s, i} a_{w,j,i} (\widehat{F}^c_{w,i} - F^c_{w,i})^2}{\sum_{w \in \mathcal{W}_s, i} a_{w,j,i} \left[(\widehat{F}^c_{w,i})^2 + (F^c_{w,i})^2 \right]}. \tag{7}$$

The numerator and denominator are first accumulated within each storm. The resulting ratios are then averaged equally over storms and components, so that each storm's overall weight is independent of its number of windows. Event-free windows from eligible storms also enter the aggregation, constraining spurious rainfall in these windows. This objective directly supervises the full recurrence. Event fractions, sampling, and evaluation definitions are provided in Appendices B.2 and C.5.

Table 1: Four-step forecasting results on the two rainfall products.

Method	Corrected			Raw		
	FSS ↑	CSI20 ↑	AW-MAE ↓	FSS ↑	CSI20 ↑	AW-MAE ↓
All-zero	0.000	0.000	3.890	0.000	0.000	4.300
Train-only climatology	0.437	0.000	3.910	0.469	0.000	3.580
Last-frame persistence	0.656	0.206	2.950	0.651	0.198	2.900
Calibrated persistence	0.573	0.147	3.040	0.600	0.142	2.770
EOF-ridge	0.625	0.184	2.660	0.623	0.172	2.490
ConvLSTM	0.654	0.224	2.540	0.646	0.201	**2.400**
SimVP	0.666	0.244	**2.380**	0.656	0.215	2.500
Earthformer	0.681	0.250	2.610	0.675	0.235	2.650
PhyDNet	0.651	0.218	2.630	0.642	0.198	2.570
LUPIN	0.674	0.248	2.550	0.667	0.232	2.490
NowcastNet-Evolution	0.692	0.258	2.560	0.686	0.248	2.520
TCP-Diffusion (restricted)	0.687	0.254	2.530	0.680	0.243	2.500
SDRNet	**0.734**	**0.292**	2.420	**0.727**	**0.281**	2.460

Trainable networks report means over three seeds; deterministic methods report a single result. Models are selected by validation aggregate FSS. Aggregate FSS combines 36 components, CSI20 uses a threshold of 20 mm/6 h, and AW-MAE is measured in mm/6 h. Bold indicates the best value in each column; relative improvements are calculated from unrounded values. Track adaptations and definitions of restricted versions are given in Table C3.

5 Experiments

5.1 Experimental Setup

Datasets. We used the Corrected and Raw rainfall products of LagTCPMoS (Pérez-Alarcón et al., 2025), covering 654 valid storms and 15,056 forecast windows in the North Atlantic during 1980–2023. Storms from 1980–2011, 2012–2015, and 2016–2023 were used for training, validation, and testing, respectively, with all windows from a given storm assigned to the same split. Each sample uses four historical fields of 6 h accumulated rainfall to predict four fields over the next 24 h. We used a north-up, storm-centered 20×20 grid with 50 km spacing and a valid radius of 500 km. Models were trained and evaluated separately on the two products, and track-conditioned methods used the same retrospective best tracks.

Comparison methods. The twelve external comparison methods include five basic approaches—all-zero prediction, training-set climatology, last-frame persistence, calibrated persistence, and track-conditioned EOF-ridge—as well as ConvLSTM (Shi et al., 2015), SimVP (Gao et al., 2022b), Earthformer (Gao et al., 2022a), PhyDNet (Le Guen & Thome, 2020), LUPIN (Pavlík et al., 2025), NowcastNet-Evolution (Zhang et al., 2023), and a restricted version of TCP-Diffusion (Huang et al., 2025). NowcastNet-Evolution is an implementation of the evolution branch based on the paper's

description, while TCP-Diffusion is restricted to the common rainfall and track inputs. We also compare update formulations, components, and dynamic conditions using the same backbone and a common soft-FSS objective, distinguishing comparisons of complete methods from the effects of the state-update design. Implementation sources, adaptations, and interventions are described in Appendices C.6 and D.3.

Experimental details. Aggregate FSS, the critical success index at 20 mm/6 h (CSI20), and area-weighted mean absolute error (AW-MAE) evaluate multiscale spatial matching, heavy-rainfall events, and rainfall amount errors, respectively. Aggregate FSS is computed at the storm level and averages equally over 36 components: four lead times, thresholds of $\{1, 5, 20\}$ mm/6 h, and neighborhood radii of $\{0, 100, 200\}$ km. All neural network methods were trained from random initialization separately on each product using three predefined seeds, 9101, 9102, and 9103. Checkpoints were selected solely by validation aggregate FSS, and a single test evaluation was performed after the configurations and checkpoints were frozen. Tables report means over the three seeds, while sample standard deviations in the figures show variation across seeds; deterministic methods report a single result. SDRNet used AdamW with an initial learning rate of 3×10^{-4} and a training budget of at most 50 epochs. Detailed settings and paired statistics at the storm level are provided in Appendix C.

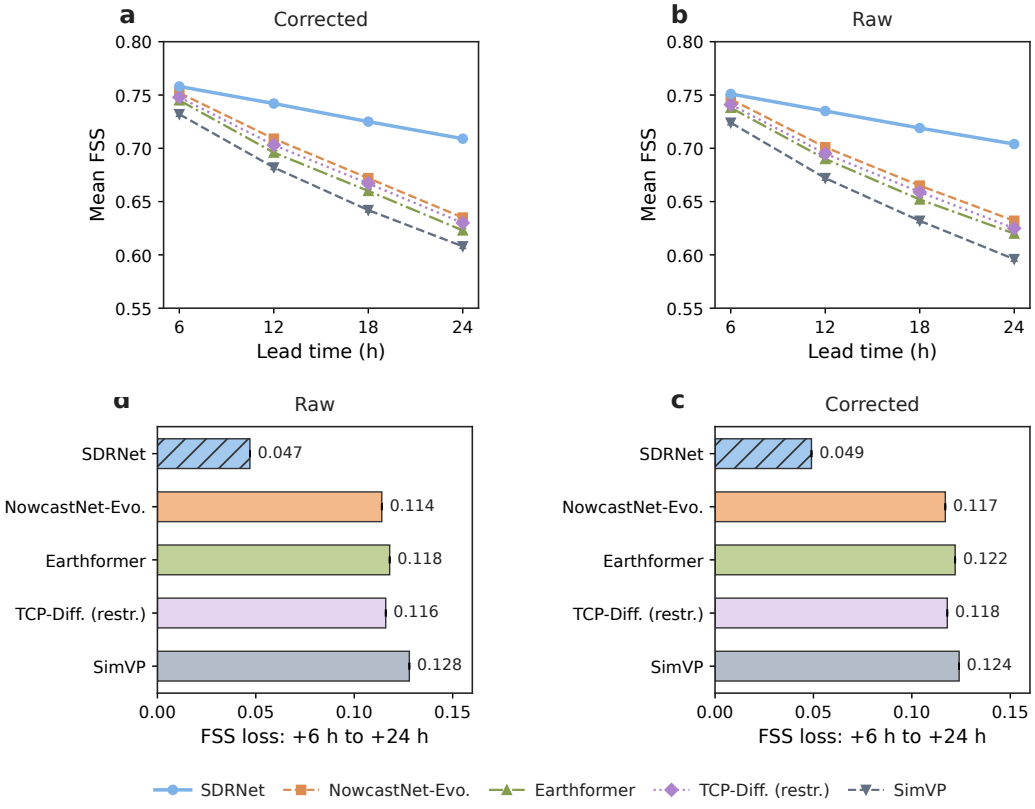

Figure 3: **Spatial skill across forecast steps and first-to-last-step drops.** (a,b) FSS at four lead times for Corrected and Raw, with each point averaging the nine threshold–neighborhood components at that lead time. (c,d) FSS at +6 h minus FSS at +24 h within the same seed. Points and bars show means over three seeds, and error bars indicate sample standard deviations. Numbers of eligible storms and the definition of the retention ratio are provided in Appendix C.5.

5.2 Overall Spatial Matching and Heavy-Rainfall Forecasting

SDRNet achieves the highest aggregate FSS and CSI20 on both rainfall products, improving aggregate FSS by approximately 6% over the best external comparison method, NowcastNet-Evolution (Table 1). Spatial matching across thresholds and neighborhoods improves alongside heavy-rainfall event identification at 20 mm/6 h, indicating better predictions of both rain-area coverage and heavy-rainfall

locations. Its AW-MAE is lower than that of NowcastNet-Evolution, whereas SimVP and ConvLSTM achieve the lowest amplitude errors on the Corrected and Raw products, respectively. The overall advantages are concentrated in spatial organization and heavy-rainfall identification, consistent with the explicit coordination of rainfall redistribution, retention, and additions.

5.3 Retention of Spatial Skill across Forecast Steps

We assess the stability of multi-step forecasting through the retention of spatial skill as lead time increases. In Figure 3, SDRNet's advantage over the main comparison methods widens at longer lead times. At +24 h, FSS remains at 0.709 and 0.704 on the Corrected and Raw products, retaining 93.5% and 93.7% of the corresponding first-step scores, respectively, with smaller first-to-last-step drops. On both products, SDRNet outperforms NowcastNet-Evolution in every threshold–neighborhood combination from +12 to +24 h (Figure D1), indicating that sustained rain-area matching at later lead times contributes broadly to the overall gains. This behavior is consistent with the model's continued use of the current rainfall field and propagation of renewal contributions into subsequent states.

5.4 Contributions of the Update Formulation and Dynamic Conditioning

Comparisons using the same backbone show that output construction directly affects forecast quality (Figure 4). Direct readout generates rainfall fields directly from the common backbone and obtains lower FSS than the full model. Warp+Residual already includes spatial deformation but remains approximately 0.010 and 0.011 lower in FSS on the two products, respectively, indicating that the organization of intensity updates also matters. The Uncoupled results further support the survival–renewal parameterization under shared decay. The full model also achieves the highest CSI20 and lowest AW-MAE among these variants (Table D1), providing consistent evidence across spatial, event, and amplitude metrics.

Among the component comparisons, removing effective renewal causes the largest loss. Removing deformation or fixing the structural summary also reduces skill, indicating that subsequent rainfall fields require spatial redistribution, local additions, and evolution conditions that adapt to the current distribution. No survival in the figure sets decay to zero and incurs a relatively small loss, indicating that decay weighting provides an additional gain. Together, these results support the joint contributions of explicit state construction and dynamic conditioning to forecasting.

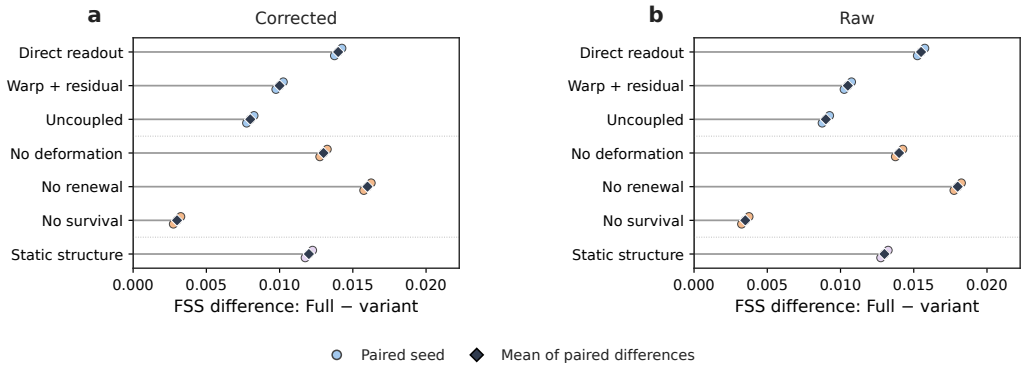

Figure 4: **Comparisons of update formulations, components, and dynamic conditioning using the same backbone.** (a,b) Paired differences in aggregate FSS for Corrected and Raw, computed as Full minus each independently retrained variant within the same seed. Circles show paired differences for the three seeds, and diamonds show their means. No survival sets $\Lambda = 0$, removing decay weighting while retaining the existing rainfall term; Static structure fixes the initial structural summary. Full metrics and interventions are provided in Appendix D.3.

5.5 Divergent Changes in Rainfall Area and Peak Intensity

Joint changes in area and peak rainfall test the model's response to nonuniform evolution. In each of the nine change categories at +24 h, SDRNet achieves higher joint direction accuracy than Earthformer, Direct readout, and Warp+Residual. It also maintains higher accuracy when the heavy-rainfall area contracts while the peak intensifies (Figure D2). CSI20 improves both inside and outside the initial heavy-rainfall area. Together with the retention of FSS at later lead times, these findings show that continued reuse of the rainfall state can be combined with joint adjustments to coverage and local intensity, improving spatial forecast quality as the field evolves.

5.6 Case Studies of Four-Step Rainfall Evolution

Two test windows further illustrate how structural continuity and local adjustment jointly appear in the predicted rainfall fields. For AL022021 in the Corrected product, SDRNet preserves a clearer southwest–northeast rainband than SimVP and adjusts local intensity with lead time. For AL032018 in the Raw product, SDRNet progressively weakens the initial intense core and follows the direction of the target's evolution more closely than last-frame persistence (Figures D5 and D6). These qualitative results are consistent with the aggregate spatial-skill and nonuniform-change analyses, while also showing remaining errors in local peaks and the timing of their changes. The cases use the same validation-based model-selection rule as the other experiments with learned models. Full four-step fields and analyses are provided in Appendix D.5.

6 Conclusion

We introduce SDRNet, which organizes successive evolution of tropical cyclone rainfall fields through storm-relative deformation, survival–renewal coupling under shared decay, and prediction feedback. In four-step, 24 h forecasting on both LagTCPMoS products with given tracks, the method improves multiscale spatial matching and heavy-rainfall identification while retaining higher skill at longer lead times. Comparisons using the same backbone, area–peak analyses, and four-step rainfall cases further support the value of the update relation, dynamic conditioning, and nonuniform-change modeling. These results show that treating the current rainfall field as an explicit state that can be continually adjusted is an effective approach to improving multi-step rainfall forecasting.

AI use statement

Generative AI tools were used to improve the clarity and readability of the manuscript and to assist with limited code editing and debugging for the experiments. All AI-assisted text and code were reviewed and verified by the authors.

References

Z. Gao, X. Shi, H. Wang, Y. Zhu, Y. Wang, M. Li, and D.-Y. Yeung. Earthformer: Exploring Space-Time Transformers for Earth System Forecasting. In *Advances in Neural Information Processing Systems*, volume 35, pp. 25390–25403, 2022a. doi: 10.52202/068431-1841.

Z. Gao, C. Tan, L. Wu, and S. Z. Li. SimVP: Simpler Yet Better Video Prediction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3170–3180, 2022b.

C. Huang, P. Mu, C. Bai, and P. A. G. Watson. TCP-Diffusion: A Multi-modal Diffusion Model for Global Tropical Cyclone Precipitation Forecasting with Change Awareness. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 25634–25653. PMLR, 2025.

V. Le Guen and N. Thome. Disentangling Physical Dynamics From Unknown Factors for Unsupervised Video Prediction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11474–11484, 2020.

J. Muller, A. M. Idzik, D. Benzi, P. J. Klotzbach, H. S. Acosta, D. R. Chavas, S. A. Urdaneta, C. T. Gray, J. E. Morris, S. B. Girimurugan, and D. Bhatt. Continental United States direct

Atlantic tropical cyclone fatalities: 1963–2024. *npj Natural Hazards*, 3:38, 2026. doi: 10.1038/s44304-026-00178-8.

P. Pavlík, M. Výboh, A. Bou Ezzeddine, and V. Rozinajová. Fully differentiable Lagrangian convolutional neural network for physics-informed precipitation nowcasting. *Applied Computing and Geosciences*, 28:100296, 2025. doi: 10.1016/j.acags.2025.100296.

A. Pérez-Alarcón, R. M. Trigo, R. Nieto, and L. Gimeno. Lagrangian Tropical Cyclone Precipitation Estimates and Moisture Sources (LagTCPMoS) Dataset. *Scientific Data*, 12:1161, 2025. doi: 10.1038/s41597-025-05490-y.

S. Pulkkinen, D. Nerini, A. A. Pérez Hortal, C. Velasco-Forero, A. Seed, U. Germann, and L. Foresti. Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1.0). *Geoscientific Model Development*, 12:4185–4219, 2019. doi: 10.5194/gmd-12-4185-2019.

X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. In *Advances in Neural Information Processing Systems*, volume 28, pp. 802–810, 2015.

Y. Zhang, M. Long, K. Chen, L. Xing, R. Jin, M. I. Jordan, and J. Wang. Skilful nowcasting of extreme precipitation with NowcastNet. *Nature*, 619:526–532, 2023. doi: 10.1038/s41586-023-06184-4.

A Derivation and Propagation Properties of the State Update

A.1 Integration Along Characteristics and Survival–Renewal Weights

Let $p(\mathbf{x}, s)$ be a continuous representation of the product rainfall field. A characteristic $\boldsymbol{\chi}_i(u)$ ending at $\mathbf{x}_i$ satisfies

$$\frac{\mathrm{d}\boldsymbol{\chi}_i(u)}{\mathrm{d}u} = \mathbf{v}_{\mathrm{rel}}(\boldsymbol{\chi}_i(u), u), \qquad \boldsymbol{\chi}_i(s + \Delta s) = \mathbf{x}_i. \tag{A.1}$$

Let $y_i(u) = p(\boldsymbol{\chi}_i(u), u)$. Spatial transport is then incorporated into the trajectory, and rainfall along the trajectory satisfies

$$\frac{\mathrm{d}y_i(u)}{\mathrm{d}u} = q_i(u) - \lambda_i(u) y_i(u), \quad q_i(u) = q_{\mathrm{eff}}(\boldsymbol{\chi}_i(u), u), \quad \lambda_i(u) = \lambda_{\mathrm{eff}}(\boldsymbol{\chi}_i(u), u). \tag{A.2}$$

Define the decay kernel $\mathscr{S}_i(u, v) = \exp[-\int_u^v \lambda_i(\xi)\,\mathrm{d}\xi]$. The integrating factor gives the Duhamel representation:

$$p(\mathbf{x}_i, s + \Delta s) = \underbrace{\mathscr{S}_i(s, s + \Delta s) p(\boldsymbol{\chi}_i(s), s)}_{\substack{\text{contribution present at the} \\ \text{start of the step}}} + \underbrace{\int_s^{s+\Delta s} \mathscr{S}_i(u, s + \Delta s) q_i(u)\,\mathrm{d}u}_{\substack{\text{accumulated contribution from} \\ \text{renewal within the step}}}. \tag{A.3}$$

The two terms share the same decay kernel: existing rainfall undergoes decay over the entire step, whereas rainfall added during the step undergoes decay over the remaining interval after its addition. Freezing the effective coefficients along the trajectory within a single step at $\bar{q}_i, \bar{\lambda}_i$, letting $\Lambda_i = \bar{\lambda}_i \Delta s$ and $Q_i = \bar{q}_i \Delta s$, and substituting $v = (u - s)/\Delta s$ yields

$$\begin{aligned} p(\mathbf{x}_i, s + \Delta s) &= e^{-\Lambda_i} p(\boldsymbol{\chi}_i(s), s) + Q_i \int_0^1 e^{-(1-v)\Lambda_i}\,\mathrm{d}v \\ &= e^{-\Lambda_i} p(\boldsymbol{\chi}_i(s), s) + \phi_1(-\Lambda_i) Q_i, \end{aligned} \tag{A.4}$$

where

$$\phi_1(-z) = \begin{cases} (1 - e^{-z})/z, & z > 0, \\ 1, & z = 0. \end{cases} \tag{A.5}$$

Λ is dimensionless, and Q has the same units as accumulated rainfall. Approximating rainfall at the starting point of the characteristic using learned displacement and bilinear sampling, then applying the target mask, gives Eq. (2) in the main text. The closed-form integration corresponds to the effective intensity relation with coefficients frozen within the step, while spatial evolution is discretized through backward sampling.

A.2 Local Changes, Nonnegativity, and the Zero-Decay Limit

Using $1 - e^{-\Lambda} = \Lambda \phi_1(-\Lambda)$, the net change in the prediction relative to the warped rainfall field at valid target grid points is

$$\widehat{P}_{k+1,i} - \widetilde{P}_{k,i} = \phi_1(-\Lambda_{k,i}) \left(Q_{k,i} - \Lambda_{k,i} \widetilde{P}_{k,i} \right). \tag{A.6}$$

The sign of the net change is determined by the relative magnitudes of $Q_{k,i}$ and $\Lambda_{k,i} \widetilde{P}_{k,i}$. Rainfall can intensify or weaken at different locations, allowing coverage contraction and local intensification to occur together in the same output field.

For any $z \geq 0$, $e^{-(1-u)z} \in [e^{-z}, 1]$, and therefore

$$0 < e^{-z} \le \phi_1(-z) \le 1. \tag{A.7}$$

Starting from a nonnegative final input frame, all predicted fields remain nonnegative because the masks, interpolation weights, weights on the two intensity contributions, and Q_k are all nonnegative. This update allows effective additions and decay to jointly change total rainfall, and bilinear backward warping uses non-conservative sampling.

The continuous limit at zero decay is

$$\phi_1(-z) = 1 - \frac{z}{2} + \frac{z^2}{6} + O(z^3), \qquad \lim_{\Lambda_k \to 0} \widehat{P}_{k+1} = M_{k+1} \odot (\widetilde{P}_k + Q_k). \tag{A.8}$$

The No survival comparison in Figure 4 uses $\Lambda = 0$, setting both the retention and renewal weights to 1 and yielding a state update without decay. The numerical implementation near zero is described in Appendix C.4.

A.3 Multi-Step Propagation and Non-Expansiveness with Fixed Evolution Quantities

Vectorize the rainfall field and all grid quantities in the same order. Along a given forecast trajectory, let W_k denote the bilinear sampling matrix corresponding to the generated displacement, and define

$$\begin{aligned} \mathsf{A}_k &= \mathrm{Diag}(M_{k+1} \odot e^{-\Lambda_k}) \mathsf{W}_k \, \mathrm{Diag}(M_k), \\ \mathbf{b}_k &= M_{k+1} \odot \phi_1(-\Lambda_k) \odot Q_k. \end{aligned} \tag{A.9}$$

Then $\widehat{\mathbf{p}}_{k+1} = \mathsf{A}_k \widehat{\mathbf{p}}_k + \mathbf{b}_k$. Let $\mathsf{A}_{b:a} = \mathsf{A}_b \cdots \mathsf{A}_a$, with an empty product equal to the identity operator. Successive expansion gives

$$\widehat{\mathbf{p}}_n = \underbrace{\mathsf{A}_{n-1:0} \widehat{\mathbf{p}}_0}_{\substack{\text{continued propagation of the} \\ \text{initial rainfall field}}} + \underbrace{\sum_{j=0}^{n-1} \mathsf{A}_{n-1:j+1} \mathbf{b}_j}_{\substack{\text{subsequent propagation of} \\ \text{renewal contributions from} \\ \text{each step}}}, \qquad n = 1, \dots, K. \tag{A.10}$$

This expansion makes the propagation paths of the two types of contribution explicit: the initial rainfall field continues to contribute to the forecast, and additions at each step, once incorporated into the explicit state, likewise undergo subsequent deformation and decay.

Non-expansiveness of the explicit propagation operator with fixed evolution quantities. Fix the generated displacement, decay, renewal, and grid geometry. The bilinear sampling matrix has nonnegative entries, and with zero padding, the weights in each row sum to at most 1; masking and exponential retention preserve this bound. Let $\alpha_k = \|\mathsf{A}_k\|_\infty$. Then $0 \le \alpha_k \le 1$, and for two rainfall fields,

$$\begin{aligned} \|\mathsf{A}_k(\mathbf{p} - \mathbf{p}')\|_\infty &\le \alpha_k \|\mathbf{p} - \mathbf{p}'\|_\infty, \\ \|\mathsf{A}_{n-1:0}(\mathbf{p} - \mathbf{p}')\|_\infty &\le \left(\prod_{k=0}^{n-1} \alpha_k \right) \|\mathbf{p} - \mathbf{p}'\|_\infty. \end{aligned} \tag{A.11}$$

Equation (A.10) also gives an amplitude bound over a finite number of steps:

$$\|\widehat{\mathbf{p}}_n\|_\infty \le \left(\prod_{k=0}^{n-1} \alpha_k \right) \|\widehat{\mathbf{p}}_0\|_\infty + \sum_{j=0}^{n-1} \left(\prod_{k=j+1}^{n-1} \alpha_k \right) \|\mathbf{b}_j\|_\infty. \tag{A.12}$$

These properties characterize the explicit propagation branch with fixed evolution quantities. The actual recurrence updates $\mathsf{A}_k, \mathbf{b}_k$ at each step based on the current rainfall field and hidden state; its multi-step forecast quality is evaluated using FSS across lead times in the main text.

A.4 Parameterization Implications of Shared Weighting

Full uses $\phi_1(-\Lambda_k)Q_k$ as the renewal contribution, whereas Uncoupled uses Q_k directly. Because $\phi_1(-\Lambda_k) > 0$ and nonnegative renewal has no additional fixed upper bound, setting $Q'_k = \phi_1(-\Lambda_k)Q_k$ allows Full to be written as $e^{-\Lambda_k}\widetilde{P}_k + Q'_k$. The two formulations therefore differ in their parameterization and in how the contributions are learned under shared weighting, rather than in an expanded set of nonnegative outputs attainable in a single step; Figure 4 and Table D1 evaluate the forecasting performance of this parameterization within the given network. $\mathbf{d}_k, \Lambda_k, Q_k$ are learned jointly under rainfall-field supervision; their interpretation concerns effective changes in the product rainfall field and the propagation of contributions.

B Complete Forecasting and Training Algorithms

B.1 Shared State Recurrence

The explicit rainfall field is always updated in the original rainfall units, and the input transformation g is used only for feature encoding. The historical reference remains fixed, while the current encoder, conditioning generator, decoder, and recurrent unit share their respective parameters across steps.

Algorithm B1 Multi-step rainfall forecasting with SDRNet.

Input: Historical rainfall fields $P_{-L+1:0}$; track $\mathbf{T}$ over the historical and future periods; geometry G; parameters θ; number of forecast steps K.

Output: Future rainfall fields $\widehat{P}_{1:K}$.

Process the transformed rainfall fields, valid domains, areas, and fixed coordinates with the historical encoder to obtain S and H_0.
Initialize $\widehat{P}_0 \leftarrow M_0 \odot P_0$.
Compute a radial–azimuthal structural summary from $\widehat{P}_0$, and combine it with the track, first-step track increment, and lead time to construct z_0.
Modulate the current encoder with z_0 to obtain current features R_0.
for $k = 0, \ldots, K-1$ **do**
  $(u_k^d, u_k^\Lambda, u_k^Q) \leftarrow \text{shared decoder}(S, R_k, H_k; z_k)$.
  $\mathbf{d}_k \leftarrow d_{\max}\tanh(u_k^d)$.
  $\Lambda_k \leftarrow \Lambda_{\max}\,\text{sigmoid}(u_k^\Lambda)$.
  $Q_k \leftarrow q_{\text{scale}}\,\text{softplus}(u_k^Q)$.
  $\widetilde{P}_k \leftarrow \mathcal{W}(M_k \odot \widehat{P}_k; \mathbf{d}_k)$.
  $\widehat{P}_{k+1} \leftarrow M_{k+1} \odot [e^{-\Lambda_k} \odot \widetilde{P}_k + \phi_1(-\Lambda_k) \odot Q_k]$.
  Save $\widehat{P}_{k+1}$.
  if $k+1 < K$ **then**
    Re-extract the structural summary from $\widehat{P}_{k+1}$ and the effective cell areas.
    Combine the summary with the next track increment and lead time to construct z_{k+1}.
    $R_{k+1} \leftarrow \text{current encoder}(g(\widehat{P}_{k+1}), G_{k+1}, G_{k+2}, X; z_{k+1})$.
    $H_{k+1} \leftarrow \text{ConvGRU}(R_{k+1}^{\text{coarse}}, H_k)$.
  end if
end for
return all predicted rainfall fields.

Four-step forecasting produces four rainfall-field outputs and three hidden-state updates. Structural conditioning, current features, and the hidden state are always constructed from the predictions already generated.

B.2 Full-Storm Objective and Validation-Based Model Selection

Training uses 454 component–storm sampling iterations per epoch and at most 50 epochs. The full-storm objective first accumulates the numerator and denominator and then takes their ratio; the

same aggregation order is maintained when forward passes are computed in chunks of windows. Uniform sampling of components and eligible storms corresponds to the expected objective in Eq. (7) of the main text.

Algorithm B2 Storm-balanced soft-FSS training.

Input: Training and validation sets; component set $\mathcal{C}$; predefined seeds; optimizer and training budget.
Output: A checkpoint selected on the validation set for each seed.

```
Estimate normalization quantities using only the training set.
For each component c, construct the set S_c of training storms containing the corresponding target event.
Independently initialize the network and optimizer for each predefined seed.
for epochs within the training budget do
    for the prescribed number of repetitions do
        Sample a component c uniformly from C, then sample a storm s uniformly from S_c.
        Use Algorithm B1 to predict all valid windows of s.
        Compute smooth event fractions for the predictions and hard-threshold event fractions for the targets.
        N ← sum of area-weighted squared differences between predicted and target fractions over all windows and grid cells.
        D ← sum of area-weighted squares of the predicted and target fractions over the same windows and grid cells.
        L ← N/D.
        Backpropagate through the full recurrence, clip gradients, and update parameters.
    end for
    Adjust the learning rate and compute hard-threshold aggregate FSS on the validation set.
    Save the checkpoint with the best validation FSS; retain the earliest epoch in an exact tie.
    Stop training when the predefined early-stopping conditions are met.
end for
Use each selected checkpoint to perform the fixed test evaluation, and save the per-seed results and summary statistics.
```

C Implementation, Data, and Evaluation Protocols

C.1 Encoders, Shared Decoding, and Parameterization

The history encoder takes historical rainfall fields, the valid domain, normalized effective cell areas, and fixed east–west, north–south, and radial coordinates as inputs. It produces multiscale reference features S and initializes H_0 from the coarsest-scale features. The current encoder takes $g(\widehat{P}_k)$, the geometry at the current and next time steps, and the fixed coordinates as inputs, producing R_k at each step. The two encoders learn representations for historical aggregation and the current state, respectively.

The shared decoder fuses H_k with the coarsest-scale current features, progressively upsamples the result, and incorporates historical and current features at the corresponding scales. Its four raw output channels represent two-dimensional displacement, accumulated decay, and effective renewal, parameterized as $\mathbf{d}_k = d_{\max}\tanh(\mathbf{u}_k^d)$, $\Lambda_k = \Lambda_{\max}\sigma(u_k^\Lambda)$, and $Q_k = q_{\text{scale}}\,\text{softplus}(u_k^Q)$. The bound $d_{\max}$ constrains each displacement component; q_{scale} is a positive numerical scale, and Q_k remains nonnegative without a fixed upper bound on its magnitude. Table C1 lists the configuration.

The input transform is $g(P) = \log(1 + \max(P, 0)/s_{\text{rain}})$, which compresses the dynamic range for encoding; state updates always use the original rainfall units.

C.2 Radial–Azimuthal Structural Summary

Let r_i, ϑ_i denote the distance of a cell center from the storm center and its azimuth measured from due east. For five radial bands of width 100 km, $\mathcal{B}_b = \{i : 100b \le r_i/\text{km} < 100(b+1)\}$, $b = 0, \ldots, 4$, we compute

Table C1: Shared implementation settings and product-specific numerical scales for SDRNet.

Setting	Corrected	Raw
History/current encoder input channels	15 / 8	15 / 8
Multiscale feature channels	48, 96, 192	48, 96, 192
Hidden-state/conditioning dimensions	192 / 208	192 / 208
Displacement component bound $d_{\max}$ (km)	200	200
Accumulated decay bound $\Lambda_{\max}$	4	4
Input rainfall scale s_{rain} (mm/6 h)	4.51656251	4.40006986
Renewal scale q_{scale} (mm/6 h)	12.52319431	10.75134754
Smoothing temperatures for thresholds 1/5/20 (mm/6 h)	0.5 / 1 / 2	0.5 / 1 / 2
Optimizer	AdamW	AdamW
Initial learning rate	3×10^{-4}	3×10^{-4}
Weight decay	10^{-4}	10^{-4}
Maximum epochs / warmup epochs	50 / 2	50 / 2
Minimum learning rate under cosine scheduling	10^{-5}	10^{-5}
Component–storm draws per epoch	454	454
Gradient norm clipping / training precision	1 / FP32	1 / FP32
Minimum epochs before early stopping / patience epochs	10 / 10	10 / 10

Models are trained independently on the two products. Input and renewal scales retain the precision of the original configuration; $d_{\max}$ constrains each displacement component, and q_{scale} is the numerical scale for renewal.

$$\begin{aligned} \mu_{k,b} &= \frac{\sum_{i\in\mathcal{B}_b} a_{k,i}\widehat{P}_{k,i}}{\sum_{i\in\mathcal{B}_b} a_{k,i}}, \\ C_{k,b,m} &= \frac{\sum_{i\in\mathcal{B}_b} a_{k,i}\widehat{P}_{k,i}\cos(m\vartheta_i)}{\sum_{i\in\mathcal{B}_b} a_{k,i}}, \\ D_{k,b,m} &= \frac{\sum_{i\in\mathcal{B}_b} a_{k,i}\widehat{P}_{k,i}\sin(m\vartheta_i)}{\sum_{i\in\mathcal{B}_b} a_{k,i}}, \qquad m\in\{1,2\}. \end{aligned} \tag{C.1}$$

The means μ summarize radial intensity, while C, D summarize low-order azimuthal asymmetry. The means are transformed using $\log(1+\mu/s_{\mathrm{rain}})$, and the signed azimuthal components are transformed using $\operatorname{asinh}(\cdot/s_{\mathrm{rain}})$.

We construct features that combine the azimuthal components with the track direction using the length ℓ_k and direction angle ψ_k of the current track increment. Letting $\rho_k = \min\{\ell_k/(50\,\mathrm{km}), 1\}$ gives

$$\begin{bmatrix}\widetilde{C}_{k,b,m}\\ \widetilde{D}_{k,b,m}\end{bmatrix} = \rho_k \begin{bmatrix}\cos(m\psi_k) & -\sin(m\psi_k)\\ \sin(m\psi_k) & \cos(m\psi_k)\end{bmatrix}\begin{bmatrix}C_{k,b,m}\\ D_{k,b,m}\end{bmatrix}. \tag{C.2}$$

When the track displacement is near zero, the additional directional components are set to zero. Both the original and additional azimuthal summaries are retained, with the latter further transformed using asinh. Together, they provide the model with combined information about structure in the north-up frame and the track direction; cellwise evolution quantities are generated from the full spatial features.

C.3 Conditional Modulation and Feedback Order

The full track, structural summary, current track increment descriptors, and lead time are embedded separately and concatenated into z_k. The current encoder and shared decoder use channelwise modulation:

$$\mathrm{FiLM}(X; z_k) = [1 + 0.1\tanh\gamma(z_k)] \odot X + 0.1\tanh\beta(z_k). \tag{C.3}$$

After generating a new rainfall field, the model first refreshes the structural summary and conditioning vector, then encodes the field to obtain the current features and updates the hidden state. Structural conditioning explicitly summarizes the current rainfall area, while the recurrent state retains the rollout context; the two are combined during decoding at the next step. Static structure and Static hidden examine these two information paths, respectively.

C.4 Grid Geometry, Area Weights, and Numerical Computation

Each frame uses north-up local azimuthal equidistant coordinates centered on its own storm center, with a valid circular domain of radius 500 km, a 20×20 grid, and 50 km grid spacing. Area weights are computed from the overlap among source cells, target cells, and the circular domain on the WGS84 ellipsoid. Let A_{ij} denote the overlap area between source cell i and target cell j within the valid domain, and let $a_j = \sum_i A_{ij}$. Under a piecewise-constant assumption within source cells, area-weighted remapping is given by

$$P_j^{\mathrm{dst}} = \frac{\sum_i A_{ij} P_i^{\mathrm{src}}}{a_j}, \qquad a_j > 0. \tag{C.4}$$

Zero rainfall is valid data, and cells with $a_j = 0$ are excluded from the loss and evaluation; 50 km is the common representation scale. Within the model, backward warping first applies the source mask, uses bilinear interpolation with zero padding outside the grid, and then applies the target mask to restrict the output. Each displacement component is constrained to $[-d_{\max}, d_{\max}]$, while the spatial mean of the displacement is learned freely, preserving net drift of the rainfall area relative to the center.

To compute $\phi_1(-\Lambda)$, we use $1 - \Lambda/2 + \Lambda^2/6$ when $\Lambda < 10^{-4}$ and $-\operatorname{expm1}(-\Lambda)/\Lambda$ elsewhere. No survival directly uses $\phi_1(0) = 1$.

The Florence rainfall fields in Figure 1 use the same 50 km remapping and display the region within $\pm$300 km of the center. Area and peak statistics use the valid circular domain of radius 500 km; all heavy-rainfall cells in both frames lie within the displayed region.

C.5 Full-Storm Metrics and Multi-Step Skill Statistics

Neighborhood event fractions. For a circular discrete neighborhood $\mathcal{N}_r(i)$, fractions are normalized by effective area:

$$\mathcal{N}_{r,a}[E]_i = \frac{\sum_{h \in \mathcal{N}_r(i)} a_h E_h}{\sum_{h \in \mathcal{N}_r(i)} a_h}. \tag{C.5}$$

The fraction is set to zero when the neighborhood has zero effective area, and $r = 0$ includes only the center cell. For $c = (j, \tau, r)$, the local event fractions during training are

$$\widehat{F}_w^c = \mathcal{N}_{r,a_{w,j}}\left[\sigma\left(\frac{\widehat{P}_{w,j} - \tau}{s_\tau}\right)\right], \qquad F_w^c = \mathcal{N}_{r,a_{w,j}}[\mathbf{1}(P_{w,j} \geq \tau)], \tag{C.6}$$

where $s_\tau > 0$ is the smoothing temperature. During evaluation, both predictions and targets use hard thresholds.

Full-storm FSS. An eligible storm contains at least one valid target event at the corresponding lead time and threshold. All of its valid windows contribute to that component, including windows without target events. The numerator and denominator are first accumulated within each storm; the resulting scores are then averaged equally over storms and the 36 components:

$$\mathrm{FSS}_{s,c} = 1 - \frac{\sum_{w\in\mathcal{W}_s,i} a_{w,j,i}(\widehat{F}^c_{w,i} - F^c_{w,i})^2}{\sum_{w\in\mathcal{W}_s,i} a_{w,j,i}[(\widehat{F}^c_{w,i})^2 + (F^c_{w,i})^2]},$$
$$\mathrm{FSS}_c = \frac{1}{|\mathcal{S}_c|}\sum_{s\in\mathcal{S}_c}\mathrm{FSS}_{s,c}, \qquad \mathrm{FSS} = \frac{1}{36}\sum_{c\in\mathcal{C}}\mathrm{FSS}_c\,. \tag{C.7}$$

Continuous errors and heavy-rainfall events. AW-MAE for each window is $\sum_i a_i|\widehat{P}_i - P_i|/\sum_i a_i$, followed by successive averaging over windows within each storm, storms, and the four lead times. CSI20, POD20, and FAR20 use the same threshold of 20 mm/6 h and per-storm area-based contingency statistics, but are calculated and aggregated separately. For each storm and lead time, hit (TP), miss (FN), and false-alarm (FP) areas are first accumulated over all valid windows of that storm. The ratios $TP/(TP+FN+FP)$, $TP/(TP+FN)$, and $FP/(TP+FP)$ are then calculated separately. Each metric is subsequently averaged equally over only the storms for which its own denominator yields a defined ratio; undefined cases remain missing. Because the three ratios have different denominators and may be defined over different sets of storms, aggregate CSI20 cannot be reconstructed algebraically from aggregate POD20 and FAR20.

Area-integrated rainfall bias is computed over unique target times for each storm and lead time, avoiding repeated accumulation of the same target rainfall field. Peak bias is the predicted maximum minus the target maximum within the valid domain. These two metrics assess the domain total and local extremes, respectively.

Lead-time FSS and retention ratio. For model m and seed ζ, let $F_{m,\zeta}(j)$ denote the mean of the nine threshold–neighborhood components at lead time j. We define

$$\overline{F}_m(j) = \frac{1}{3}\sum_{\zeta=1}^{3} F_{m,\zeta}(j),$$
$$D_{m,\zeta} = F_{m,\zeta}(6\,\mathrm{h}) - F_{m,\zeta}(24\,\mathrm{h}), \tag{C.8}$$
$$R_m = \frac{\overline{F}_m(24\,\mathrm{h})}{\overline{F}_m(6\,\mathrm{h})}.$$

The first-to-last-step FSS drop in Figure 3 is calculated within each seed, and its mean and sample standard deviation are then reported. The retention ratio is the ratio of the last-step seed mean to the first-step seed mean. Eligible storms are determined from the corresponding target events at each lead time; the three radii share the same eligibility set at a given threshold. Table C2 lists the support counts.

Table C2: Numbers of eligible storms for FSS at each threshold and lead time.

Product	Threshold (mm/6 h)	+6 h	+12 h	+18 h	+24 h
Corrected	1	147	147	147	147
Corrected	5	147	147	147	147
Corrected	20	144	143	140	138
Raw	1	147	147	147	147
Raw	5	147	147	147	147
Raw	20	137	136	134	132

The full test set contains 147 storms and 3,763 windows. At a given lead time and threshold, the three neighborhood radii share the same target-event eligibility set; the eligible set for heavy-rainfall FSS varies with lead time.

Repeated training. Tables report means over three predefined seeds, and seed error bars in the figures show sample standard deviations. For per-storm comparisons, the paired differences for the same storm are first averaged over the three seeds, and the distribution across storms is then displayed.

Paired storm resampling. We perform 10,000 paired cluster bootstrap replicates over the predefined set of 147 test storms. Each replicate draws 147 storm IDs with replacement, keeping all windows

from a storm together as one cluster. The same sampled IDs and their multiplicities are used for both methods, all three seeds, and all evaluation components. For each seed and threshold–neighborhood–lead-time component, the scores for the storms in the sampled sequence that meet the target-event eligibility criterion are first averaged equally, with repeatedly sampled storms counted according to their multiplicities. The available components are then averaged equally to calculate the difference between methods, and the paired differences are finally averaged over the three seeds. The 95% confidence interval is given by the 2.5th and 97.5th percentiles of the 10,000 replicate differences. This interval characterizes test-storm sampling uncertainty conditional on the three training runs; training variability across seeds is described separately by the sample standard deviation.

C.6 Data Splits, Method Adaptations, and Supplementary Settings

Sample construction and model selection. Each sample consists of eight strictly consecutive frames: four historical frames followed by four target frames. The source data contained 707 storms and 19,990 time points. When eight-frame samples were constructed within each storm with a stride of 1, 51 storms yielded no candidate windows because they had fewer than eight frames. The remaining 656 storms yielded 15,105 candidate eight-frame windows. Temporal continuity checks further excluded 49 windows spanning nine zero-duration intervals across nine storms. Seven of these storms retained valid windows, whereas all candidate windows for AL162008 and AL062009—1 and 5 windows, respectively—were excluded. This left 15,056 valid samples from 654 storms; subsequent spatial and field-validity checks excluded no additional windows. The training, validation, and test sets contained 454/53/147 storms with samples and 10,072/1,221/3,763 windows, respectively. Models were trained independently on the two rainfall products, and each neural network was trained from random initialization using seeds 9101, 9102, and 9103. Checkpoints were selected solely by validation aggregate FSS, retaining the earliest epoch in the event of an exact tie. All models were tested under the common protocol after their configurations and checkpoints had been frozen, and all metrics for a given seed were calculated from the same checkpoint.

Table C3: Prediction formulations and task adaptations of the external comparison methods.

Method	Prediction formulation	Adaptation and comparator identity in the common task
All-zero; Train-only climatology	Zero field and training-set mean rainfall field	Use the common valid domain; climatology is estimated only from the training data.
Persistence; Calibrated persistence	Last-frame persistence and intensity calibration	All four steps use the common center coordinates; calibration parameters are determined using only the training and validation sets.
EOF-ridge	Low-dimensional statistical mapping	Incorporates track conditions; the EOF basis and data normalization are fitted only on the training set.
ConvLSTM	Convolutional recurrent spatiotemporal modeling	Retains the recurrent backbone, with adaptations for four-frame input, four-frame output, and the track interface.
SimVP	Convolutional encoding and temporal mapping	Retains the temporal modeling architecture, with adaptations for the common rainfall fields and track inputs.
Earthformer	Local and global spatiotemporal attention	Adapts to the 20×20 grid and track conditions while retaining the spatiotemporal attention backbone.
PhyDNet	Physics-inspired hidden states and complementary residuals	Retains both branches and their corresponding constraints, with adaptations to the common prediction grid.
LUPIN	Lagrangian transformation and intensity differencing	Retains the dual motion and intensity networks and staged training, adapting the rollout to four lead times at 6 h intervals.
NowcastNet-Evolution	Evolution through motion and intensity residuals	A PyTorch evolution branch implemented from the paper's equations and architectural descriptions, without the conditional generative network.
TCP-Diffusion (restricted)	Changes between adjacent steps and diffusion-based prediction	Adapts to the 24 h/6 h task using the common rainfall fields and track conditions; excludes the additional environmental and numerical forecast inputs used by the original method.

All methods use common data splits, available inputs, target grids, validation-based model-selection rules, and evaluation rules; network architectures and training procedures are adapted for each method.

Implementation sources and task adaptations. ConvLSTM, SimVP, and PhyDNet were adapted from a fixed version of OpenSTL. Earthformer used CuboidTransformer from the official repository. LUPIN and TCP-Diffusion were adapted from fixed versions of the authors' public repositories. NowcastNet-Evolution is a PyTorch implementation of the evolution branch based on the paper's equations and public architectural descriptions; it excludes the conditional generative network, and conclusions drawn from comparisons apply only to this implementation. All methods share the data splits, four historical frames, four target steps, effective areas, available track information, and evaluation procedures. Table C3 lists their prediction formulations.

Training objectives and levels of comparison. External methods use their native or task-adapted objectives: ConvLSTM and SimVP use area-weighted MSE, while Earthformer and SDRNet use full-storm soft-FSS. The external comparisons evaluate the complete adapted methods. Variants using the same backbone share soft-FSS and consistent training and validation-based model-selection rules to analyze the contributions of SDRNet's update formulation and feedback conditions. Table C1 lists SDRNet's optimization settings.

Supplementary experimental settings. Variants using the same backbone were retrained independently; the interventions are defined in D.3. The half-data setting used full-storm training on 227 storms, and the combined objective added an AW-MAE term to soft-FSS. The track-conditioning experiments included stepwise Gaussian perturbations to center positions with an RMS of approximately 25 km and forecast-like track perturbations with a median position error of approximately 45 km at +24 h. The area–peak analysis uses the last input frame of each window to determine the common reference and tolerances; targets are used only for stratified evaluation after testing, as defined in D.2. Both rainfall cases use the same test setting, and predictions from the learned models use checkpoints selected on the validation set.

D Supplementary Experiments

This appendix examines the coverage of spatial gains across evaluation conditions and storms, nonuniform evolution, update formulations, responses to training and conditioning settings, and complete four-step rainfall-field cases. All learned models use checkpoints selected on the validation set. The metrics and statistical procedures are described in Appendix C.

D.1 Coverage of Gains Across Evaluation Conditions and Storms

At longer lead times, the gains are consistent across thresholds and spatial scales. In Figure D1a,b, SDRNet outperforms NowcastNet-Evolution on 30/36 components for Corrected and 32/36 for Raw. All combinations from +12 to +24 h show positive differences, whereas negative differences at the first step are concentrated in some light-rainfall conditions. Thus, SDRNet's advantage becomes more consistent as forecasting proceeds, in line with the retention of FSS at later lead times reported in the main text.

Per-storm comparisons further show how broadly the gains extend across samples. SDRNet achieves positive FSS differences relative to Earthformer for all 147 test storms in both products. Relative to Warp+Residual using the same backbone, the numbers of storms with positive differences are 96/147 and 98/147, respectively (Figure D1c,d). The average improvement therefore extends across a broad set of storms, although gains still vary across storm evolutions when compared with a more closely related update formulation.

D.2 Heavy-Rainfall Area, Peak Rainfall, and Responses in Initial Regions

D.2.1 Definitions of Joint Changes and Performance by Category

Using the last input frame as the reference, we define the heavy-rainfall area and the peak within the valid domain as

$$A_{20}(P_k; a_k) = \sum_i a_{k,i}\, \mathbf{1}[P_{k,i} \geq 20\ \mathrm{mm/6h}], \qquad \Pi(P_k; M_k) = \max_{i:M_{k,i}=1} P_{k,i}. \tag{D.1}$$

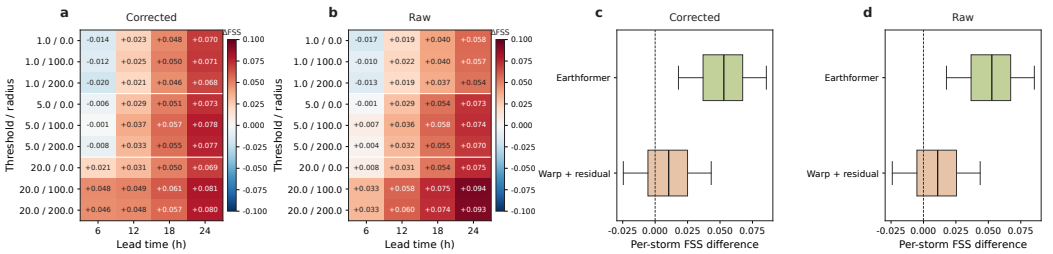

Figure D1: **Coverage of spatial skill gains.** (a,b) Differences in the 36 FSS components, computed as SDRNet minus NowcastNet-Evolution. Rows correspond to thresholds and neighborhood radii, and columns correspond to lead times. Both products share a color scale centered at zero. (c,d) Per-storm FSS differences relative to Earthformer and Warp+Residual, computed as SDRNet minus each comparator, with paired differences first averaged across the three seeds for each storm. Boxes span the 25th–75th percentiles, the central line marks the median, and whiskers extend to the most extreme observations within 1.5 times the interquartile range. Each group contains 147 storms.

For window w, let A_0 and Π_0 denote the heavy-rainfall area and peak in the last input frame, and let $\widetilde{a}_0$ denote the median of the corresponding effective cell areas. Both predictions and targets use this input field as their reference and share the tolerances

$$\varepsilon_A = \max(0.1A_0,\ 4\widetilde{a}_0), \qquad \varepsilon_\Pi = \max(0.1\Pi_0,\ 5\ \mathrm{mm/6h}). \tag{D.2}$$

Both area and $\widetilde{a}_0$ are measured in km^2. At target lead time j, changes in area and peak are calculated relative to the last input frame. A change strictly greater than the positive tolerance is classified as expansion or increase, and a change strictly less than the negative tolerance as contraction or decrease. Changes within the closed interval $[-\varepsilon, \varepsilon]$ are classified as stable. Combining the two types of change gives nine categories. A prediction has the correct joint direction only when both its area category and peak category match those of the target. Target changes are used only for stratified evaluation after testing and are not used as inputs or during training.

Results are aggregated separately at each lead time. Joint direction accuracy is first averaged over windows within each storm–category–lead-time group and then averaged with equal weights across storms. For within-category FSS, the numerator and denominator are first accumulated across windows from the same storm and category, after which the resulting scores are averaged equally over storms that meet the target-event eligibility criteria. Categories are defined at the window–lead-time level, so the same storm can enter different categories at different windows or lead times.

In Figure D2a,b, SDRNet achieves higher joint direction accuracy than Earthformer, Direct readout using the same backbone, and Warp+Residual in all nine change categories. Its advantage covers changes in both the same and opposite directions. At +24 h, the "area contraction–peak increase" category contains 49 storms and 86 windows for Corrected, and 26 storms and 40 windows for Raw. This category requires coverage and local intensity to change in opposite directions, directly connecting the challenge of nonuniform evolution described in the Introduction to model performance.

At +24 h, the nine categories contain a total of 3,763 windows for each product. Storms can overlap across groups, so storm counts across groups should not be added as mutually exclusive totals.

D.2.2 Predictions Inside and Outside the Initial Heavy-Rainfall Region

The initial heavy-rainfall and non-heavy-rainfall regions are defined once from the last input frame. SDRNet achieves higher +24 h CSI20 than Earthformer, Direct readout, and Warp+Residual in both regions (Figure D2c,d). The model improves both subsequent matching at the initial heavy-rainfall locations and the prediction of heavy rainfall outside them. Spatial backward warping and local effective renewal jointly provide a computational means of changing where heavy rainfall occurs. These regional results complement the nine-category joint-change analysis from the perspective of spatial distribution.

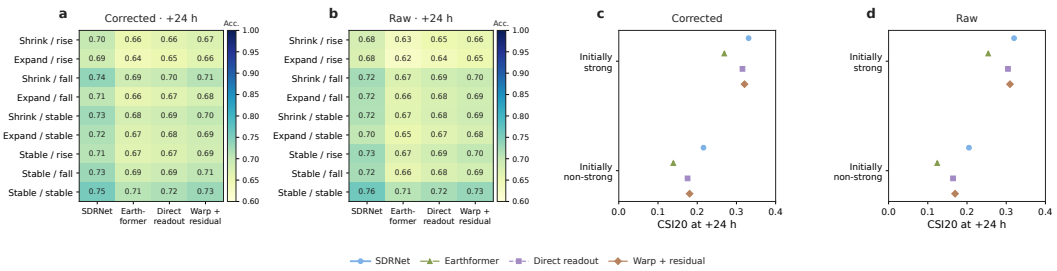

Figure D2: **Joint changes in area and peak rainfall and regional responses at +24 h.** (a,b) Joint direction accuracy for the nine target-change categories in Corrected and Raw. Rows correspond to area/peak combinations, and heatmap values are shown to two decimal places. (c,d) CSI20 in the initial heavy-rainfall and non-heavy-rainfall regions for both products. Values and points are means over three seeds. Regions are determined from the last input frame; the numbers of eligible storms in the two regions are 141 and 146 for Corrected, and 128 and 145 for Raw, respectively.

D.3 Complete Metrics for Update Formulations and Errors at the Final Lead Time

D.3.1 Update Formulations Using the Same Backbone

In Table D1, the full model outperforms all other variants using the same backbone on FSS, CSI20, and AW-MAE. Warp+Residual includes spatial deformation, while Uncoupled retains decay and nonnegative additions. Their differences from Full further isolate how intensity contributions are organized. Let $\widetilde{P}_k = \mathcal{W}(M_k \odot \widehat{P}_k; \mathbf{d}_k)$. Omitting the common target mask gives

$$\begin{aligned} \text{Full}: &\quad e^{-\Lambda_k} \odot \widetilde{P}_k + \phi_1(-\Lambda_k) \odot Q_k, \\ \text{Warp+Residual}: &\quad [\widetilde{P}_k + R_k^{\text{res}}]_+, \\ \text{Uncoupled}: &\quad e^{-\Lambda_k} \odot \widetilde{P}_k + Q_k. \end{aligned} \tag{D.3}$$

Here, R_k^{res} is a signed residual, and $[\cdot]_+$ denotes nonnegative clipping. Direct readout uses a direct nonnegative readout; No deformation sets $\mathbf{d}_k = 0$; No renewal sets $Q_k = 0$; and No survival sets $\Lambda_k = 0$, making both weights equal to 1. Static structure holds the initial structural summary fixed while continuing to update the current rainfall field and hidden state. All variants are independently retrained.

The paired 95% confidence intervals for the FSS difference between Full and Warp+Residual are [0.007, 0.013] and [0.008, 0.014] for the two products, respectively; the corresponding intervals relative to Uncoupled are [0.004, 0.012] and [0.005, 0.013]. The intervals are above zero for both products. Together with the rankings on the three metrics, these results support the predictive gains of the coupled parameterization. The algebraic relationship between Full and Uncoupled is discussed in Appendix A.4.

D.3.2 Heavy-Rainfall Events and Rainfall Amount Calibration at +24 h

At the final lead time, SDRNet improves heavy-rainfall detection while reducing the false alarm ratio. Compared with NowcastNet-Evolution, Earthformer, and restricted TCP-Diffusion, it achieves higher POD20, lower FAR20, and the highest CSI20 (Table D2). These joint changes show that improved heavy-rainfall skill at later lead times is accompanied by a better balance between hits and false alarms.

SDRNet also has an area-integrated rainfall bias closer to zero and a lower absolute peak error. For Corrected, overestimation of area-integrated rainfall coexists with underestimation of the peak, whereas the peak is overestimated for Raw. These results reflect different error characteristics for domain totals and local extremes. By attenuating existing contributions and adding rainfall locally, the survival–renewal relation provides a common mechanism for adjusting both quantities.

Table D1: Comparisons of update formulations, components, and dynamic structure using the same backbone.

Variant	Corrected			Raw		
	FSS ↑	CSI20 ↑	AW-MAE ↓	FSS ↑	CSI20 ↑	AW-MAE ↓
Full	**0.734**	**0.292**	**2.420**	**0.727**	**0.281**	**2.460**
Direct readout	0.720	0.280	2.500	0.712	0.269	2.540
Warp+Residual	0.724	0.284	2.470	0.717	0.273	2.510
Uncoupled	0.726	0.285	2.460	0.718	0.274	2.500
No deformation	0.721	0.282	2.490	0.713	0.271	2.530
No renewal	0.718	0.277	2.520	0.709	0.266	2.560
No survival	0.731	0.289	2.440	0.724	0.278	2.480
Static structure	0.722	0.282	2.480	0.714	0.271	2.520

All variants are independently retrained using the same three seeds and validation-based model selection rule. Full denotes SDRNet. Bold marks the best value in each column. Values are means across seeds, and AW-MAE is measured in mm/6 h. No survival sets $\Lambda = 0$; Static structure holds the initial structural summary fixed while updating the other recurrent states normally.

Table D2: Heavy-rainfall events and rainfall amount errors at +24 h.

Method	CSI20 ↑	POD20 ↑	FAR20 ↓	Rainfall bias (%)	Peak bias (mm/6 h)	Absolute peak error ↓ (mm/6 h)
Corrected						
SDRNet	**0.251**	**0.564**	**0.681**	**+4.200**	**-5.100**	**13.950**
NowcastNet-Evolution	0.201	0.522	0.748	+4.636	-5.162	14.697
Earthformer	0.189	0.513	0.764	+4.751	-5.179	14.895
TCP-Diffusion (restricted)	0.193	0.515	0.757	+4.688	-5.170	14.787
Raw						
SDRNet	**0.240**	**0.557**	**0.692**	**+7.200**	**+3.510**	**14.750**
NowcastNet-Evolution	0.191	0.519	0.758	+7.633	+3.593	15.493
Earthformer	0.174	0.501	0.780	+7.749	+3.615	15.691
TCP-Diffusion (restricted)	0.182	0.509	0.771	+7.696	+3.605	15.601

Values are means over three seeds. Bold marks the best value within each product. Rainfall bias is the signed relative bias in area-integrated rainfall, and peak bias is the prediction minus the target; values closer to zero are better for both. For each event metric, event areas are accumulated within each storm, and the metric is then averaged over the storms for which it is defined.

D.4 Computational Cost and Responses to Training and Conditioning Settings

D.4.1 Model Size and Inference Cost

With 10.800 M parameters, SDRNet completes four-step model inference in 24.000 ms, taking less time than Earthformer, NowcastNet-Evolution, and restricted TCP-Diffusion while achieving higher aggregate FSS (Figure D3 and Table D4). SimVP and ConvLSTM have faster inference, while PhyDNet and LUPIN have fewer parameters. Overall, relative to several high-skill comparison models, SDRNet achieves better spatial matching with a smaller model and lower inference cost.

D.4.2 Trade-Off Between Spatial Skill and Amplitude Error

The training objective can adjust the trade-off between spatial matching and amplitude error. Adding AW-MAE to soft-FSS reduces SDRNet's AW-MAE by 0.080 mm/6 h for both products, with a small decrease in FSS. Earthformer and Warp+Residual show changes in the same direction (Figure D4 and Table D3). This reflects the different emphases of neighborhood threshold coverage and cellwise

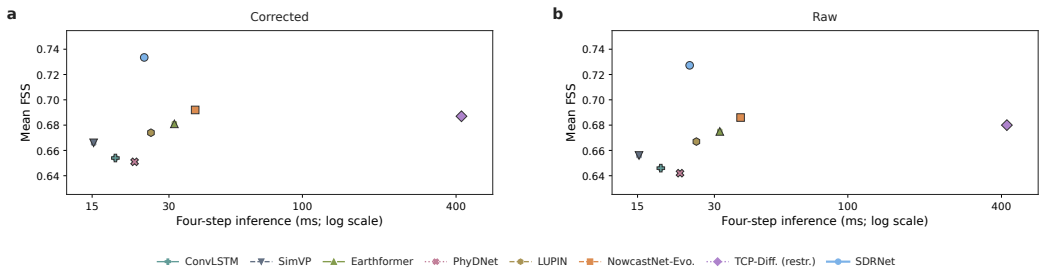

Figure D3: **Spatial skill and model inference cost.** (a,b) Corrected and Raw. The horizontal axis shows four-step model inference time on a logarithmic scale, and the vertical axis shows aggregate FSS. Each point is a mean over three seeds. Model size, peak GPU memory, and training cost are reported in Table D4.

rainfall fitting. SDRNet retains its FSS advantage over both comparison models under the default and combined objectives.

Table D3: Comparisons of training objectives, data volume, hidden states, and track conditions.

		Corrected		Raw	
Model	Setting	FSS ↑	AW-MAE ↓	FSS ↑	AW-MAE ↓
Training objective					
SDRNet	soft-FSS	**0.734**	**2.420**	**0.727**	**2.460**
SDRNet	soft-FSS + AW-MAE	**0.731**	**2.340**	**0.725**	**2.380**
Earthformer	soft-FSS	0.681	2.610	0.675	2.650
Earthformer	soft-FSS + AW-MAE	0.679	2.530	0.673	2.570
Warp+Residual	soft-FSS	0.724	2.470	0.717	2.510
Warp+Residual	soft-FSS + AW-MAE	0.722	2.390	0.715	2.430
Number of training storms					
SDRNet	50% (227 storms)	**0.715**	**2.560**	**0.707**	**2.600**
SDRNet	100% (454 storms)	**0.734**	**2.420**	**0.727**	**2.460**
Earthformer	50% (227 storms)	0.647	2.830	0.641	2.870
Earthformer	100% (454 storms)	0.681	2.610	0.675	2.650
Direct readout	50% (227 storms)	0.692	2.680	0.684	2.720
Direct readout	100% (454 storms)	0.720	2.500	0.712	2.540
Hidden-state updates					
Static hidden	Initial hidden state held fixed	0.725	2.480	0.717	2.520
Track conditions					
SDRNet	Conditioning vector perturbations	**0.725**	**2.480**	**0.718**	**2.520**
SDRNet	Forecast-like track perturbations	**0.713**	**2.560**	**0.706**	**2.600**
Earthformer	Conditioning vector perturbations	0.665	2.720	0.659	2.760
Earthformer	Forecast-like track perturbations	0.650	2.820	0.644	2.860

Values are means over three seeds, and AW-MAE is measured in mm/6 h. The settings are compared separately. Bold marks the best value among methods under the same setting; the single Static hidden row is not ranked. Conditioning vector perturbations apply Gaussian perturbations to the center position at each step, with an RMS of approximately 25 km. Forecast-like track perturbations have a median position error of approximately 45 km at +24 h.

D.4.3 Data Volume and Hidden-State Updates

When trained with half the data, SDRNet experiences a smaller FSS decrease than Earthformer and Direct readout using the same backbone (Table D3). With full-storm training on 227 storms, explicit state updates retain relatively high spatial skill, supporting the value of reusing existing rainfall fields

Table D4: Model size and computational cost.

Method	Parameters (M)	Peak GPU memory (GB)	Four-step inference (ms)	Training GPU h	
				Corrected	Raw
ConvLSTM	8.200	**6.500**	18.500	**7.800**	**8.100**
SimVP	12.400	8.200	**15.200**	9.600	9.900
Earthformer	15.800	12.800	31.500	14.500	14.800
PhyDNet	**4.600**	7.100	22.000	8.700	9.000
LUPIN	7.300	9.400	25.500	10.800	11.100
NowcastNet-Evolution	19.600	14.500	38.000	18.000	18.300
TCP-Diffusion (restricted)	32.000	17.800	420.000	30.000	30.300
SDRNet	10.800	10.200	24.000	12.500	12.800

Values are means over three seeds. Bold marks the lowest value in each cost column. Parameter count, peak GPU memory, and four-step model inference time are identical for the two products and are reported jointly; training GPU hours are listed separately.

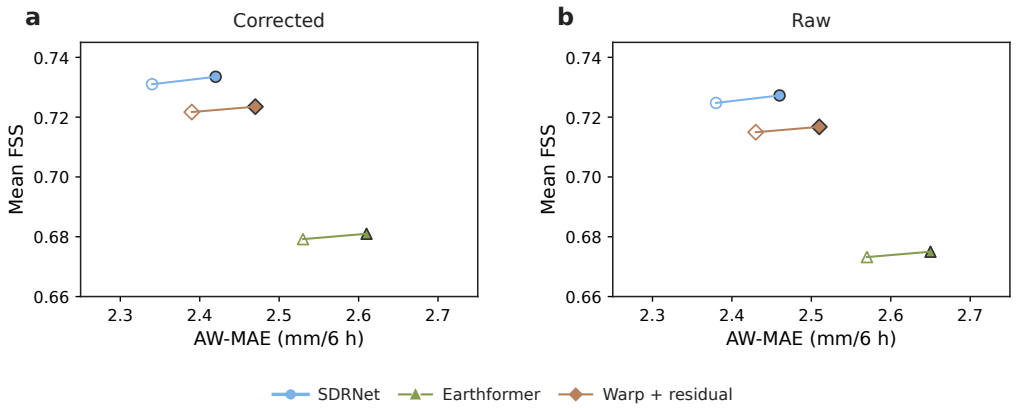

Figure D4: **The spatial–amplitude trade-off induced by the training objective.** (a,b) Corrected and Raw. The horizontal axis shows AW-MAE, and the vertical axis shows aggregate FSS. Filled points indicate soft-FSS, and open points indicate the combined soft-FSS and AW-MAE objective. Lines connect the two settings for the same model. Each point is a mean over three seeds.

and learning subsequent adjustments at this training scale. Holding the initial hidden state fixed in Static hidden reduces skill. Together with the Static structure results, this indicates that the recurrent context and the summary of the current rainfall region each provide useful predictive information.

D.4.4 Perturbations to Track Conditions

Both types of track perturbation decrease FSS and increase AW-MAE. SDRNet experiences a smaller skill decrease than Earthformer and retains higher FSS for both products (Table D3). These results show that track conditions play a role in rainfall-field prediction, while explicit deformation and intensity updates retain their relative advantage under the specified conditioning perturbations. The perturbation types and magnitudes are described in Appendix C.6.

D.5 Cases of Four-Step Rainfall-Field Evolution

We selected two test windows to examine two types of evolution: rainband organization with local intensification, and weakening of an initially intense core. Both cases show the last input frame and all four forecast lead times. The learned models follow the same validation-based model selection rule as in the other experiments. Rainfall fields use a north-up, storm-centered grid, and both figures

share the same rainfall color scale to facilitate comparison of continuous changes in structure and intensity.

D.5.1 Rainband Organization and Local Intensification in the Corrected Product

From +6 to +18 h, the target rainfall field for AL022021 gradually develops a more concentrated heavy-rainfall core oriented southwest–northeast. As lead time increases, SimVP produces smoother, more dispersed rainfall regions and does not fully capture the target's localized, concentrated intensification. SDRNet retains a clearer directional structure and continues to adjust intensity during rollout, showing that explicit reuse of the current rainfall field can be combined with subsequent local changes (Figure D5). Its intense core remains too weak at +18 h, with more pronounced intensification occurring at +24 h. This indicates room to improve the prediction of local peaks and the timing of intensification.

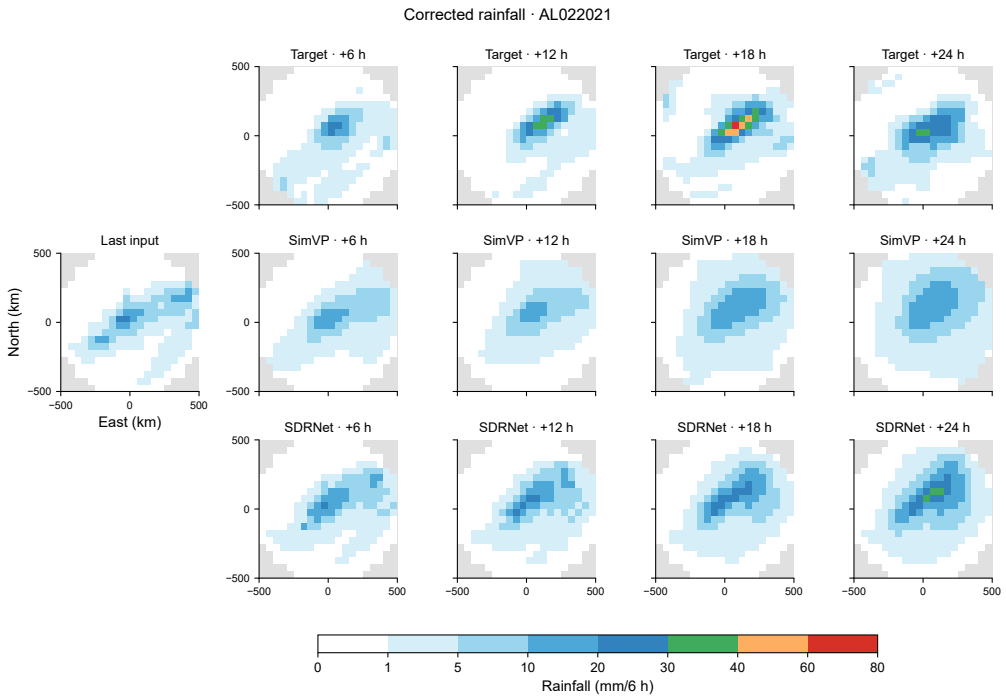

Figure D5: **Four-step rainfall evolution for AL022021 in the Corrected product.** The left panel shows the last input frame at 12 UTC on 14 June 2021. The three rows on the right show the target, track-conditioned SimVP, and SDRNet, respectively. The four columns correspond to +6, +12, +18, and +24 h, spanning 18 UTC on 14 June to 12 UTC on 15 June. The grid is north-up with its origin at the storm center, and coordinates are in km. Gray denotes regions without valid support, and white corresponds to 0–1 mm/6 h. The color scale is shared with Figure D6.

D.5.2 Weakening of an Initially Intense Core in the Raw Product

The target rainfall field for AL032018 shows gradual weakening of the initially intense core, with only a weaker rainfall region at a somewhat shifted location remaining at later lead times. Last-frame persistence retains the initial intense core at all four lead times and therefore increasingly diverges from the target as lead time grows. SDRNet continually reduces the core intensity and more closely follows the target's weakening at later lead times, reflecting the continuation and adjustment of existing rainfall (Figure D6). The residual core in the prediction remains too intense, and its location also differs from the target. This case shows both the improvement over persistence and local changes that are not yet fully reproduced.

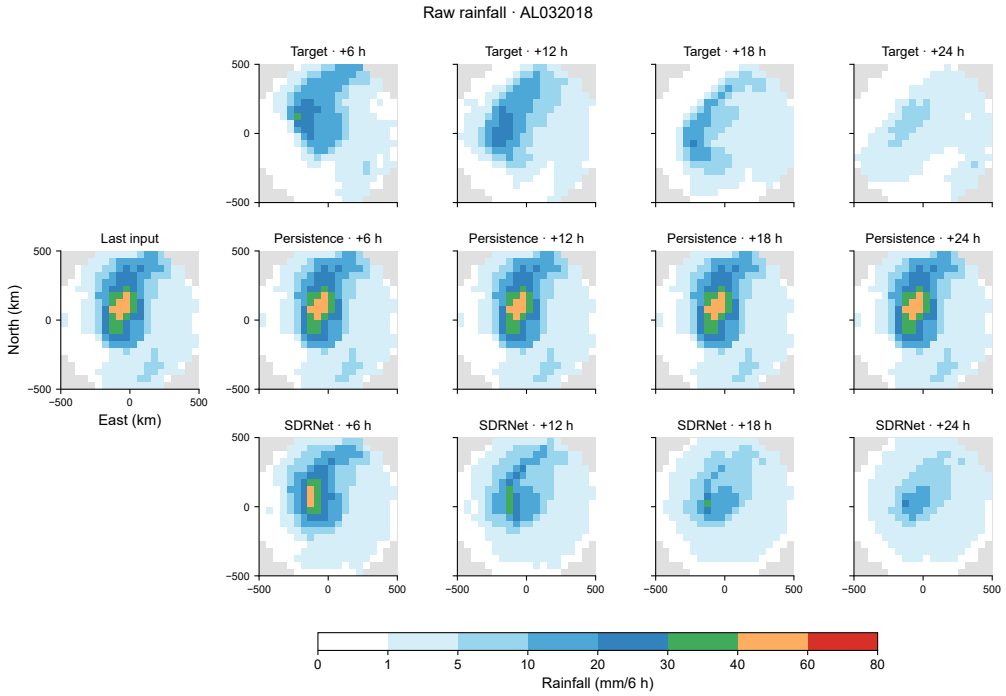

Figure D6: **Four-step rainfall evolution for AL032018 in the Raw product.** The left panel shows the last input frame at 12 UTC on 12 July 2018. The three rows on the right show the target, last-frame persistence, and SDRNet, respectively. The four columns correspond to +6, +12, +18, and +24 h, spanning 18 UTC on 12 July to 12 UTC on 13 July. Coordinates, valid-domain conventions, and color-scale meanings are the same as in Figure D5.

Through both intensification and weakening, these two windows show that SDRNet can continue to change local rainfall amounts while retaining spatial structure. The cases provide a visual complement to the aggregate FSS, heavy-rainfall event, and area–peak results. The contributions of specific design choices are assessed through comparisons using the same backbone.