GeoRad-3D: Factorized Geometry Transport and Residual Radiometry for 3D Radar Nowcasting

Yiting Li, Zihan Zhou, Shengkai Chen, Jing Zhang, Jianyu Liang, Ka Wing Chui, Chen Chen, Fayao Liu, Yifang Yin, Shili Xiang

IAIC, A*STAR · Nanyang Technological University · Macau University of Science and Technology · Centre for Climate Research Singapore, NEA

NeurIPS 2026

3D radar nowcasting on coherent Gaussian parcels

Task

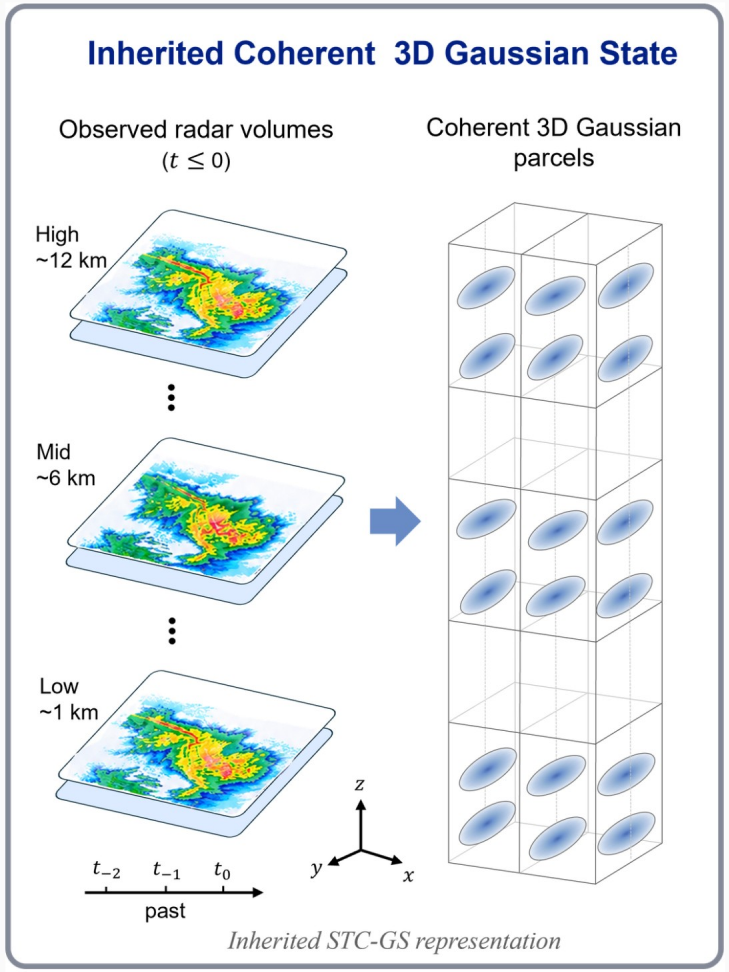
Observe 5 past 3D radar volumes and forecast the next 20: a 100-minute horizon at 5-minute steps.

Why 3D

Vertical growth, height-dependent drift and overhanging strong echoes are lost in 2D products. Dense 3D grids are costly in memory and compute.

STC-GS makes it practical

Each volume becomes N tracked Gaussian parcels (Wang et al., ICLR 2025). A parcel carries geometry (center, scale, rotation) and 6-D radiometry. Forecasts are rendered back to reflectivity in dBZ.



One unified regressor, three mismatches

STC-GS evolves all future Gaussian attributes with a single deterministic regressor.

GEOMETRY

Transport mismatch

Parcels of one storm share a global motion trend with local deformation on top. Independent errors accumulate into trajectory drift.

RADIOMETRY

Uncertainty mismatch

Strong echoes grow and decay. A deterministic forecast averages these alternatives and underestimates intensity.

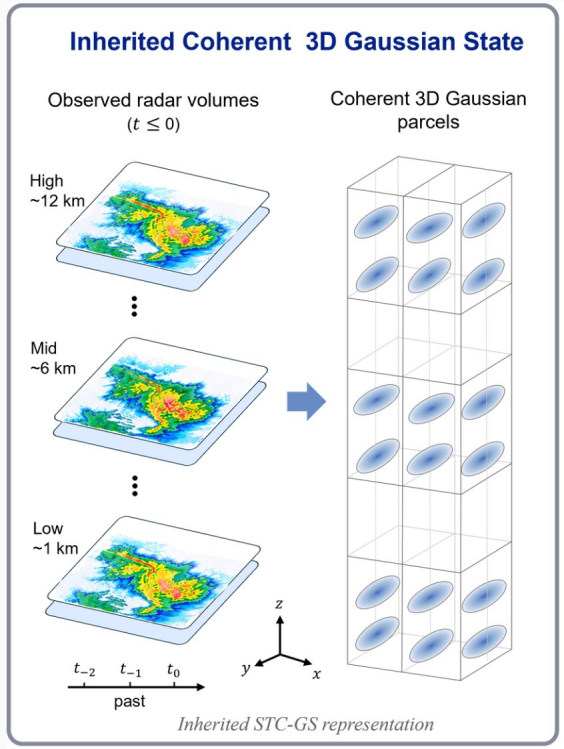
OBJECTIVE

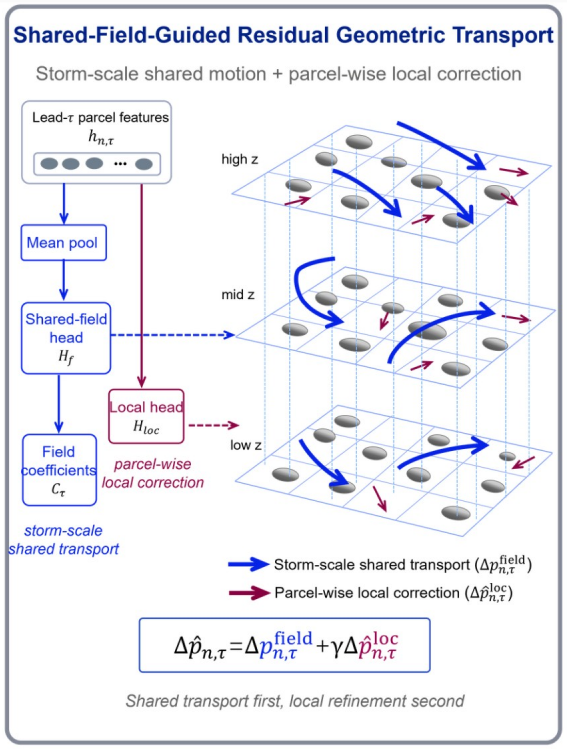
Training vs. evaluation

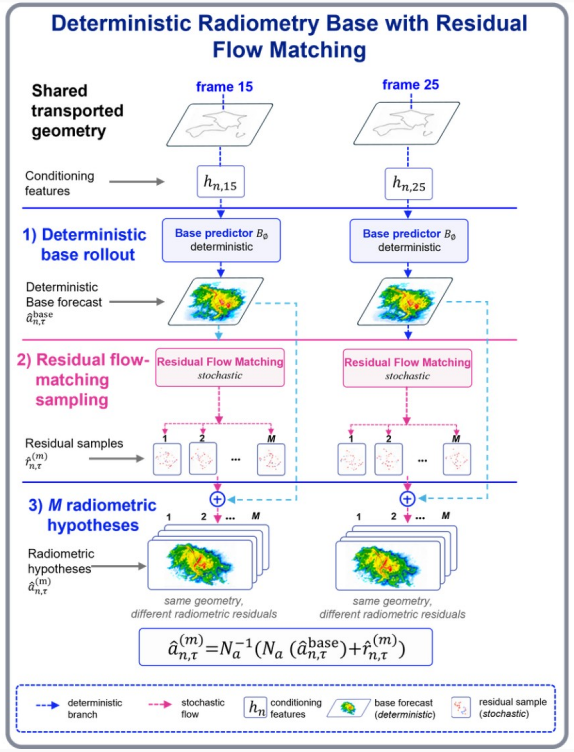
Training targets Gaussian states. Skill is scored on the rendered volume, where small state errors shift echoes across dBZ thresholds.

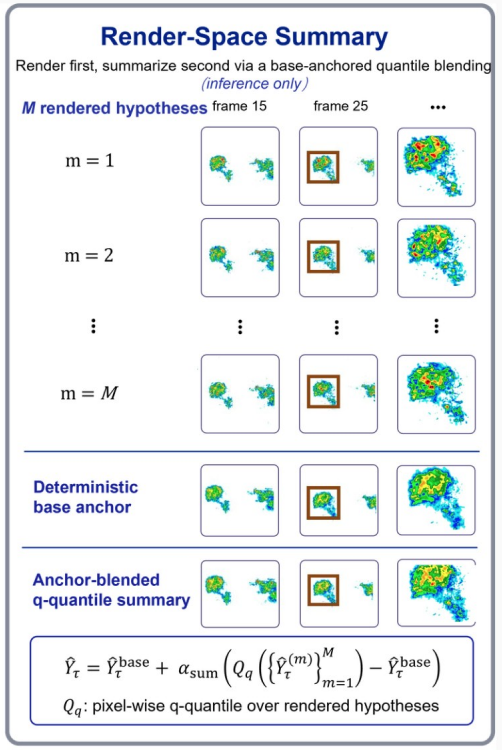
Our position: **factorize geometry and radiometry**, and form the final forecast in radar-product space.

GeoRad-3D: a factorized forecasting head







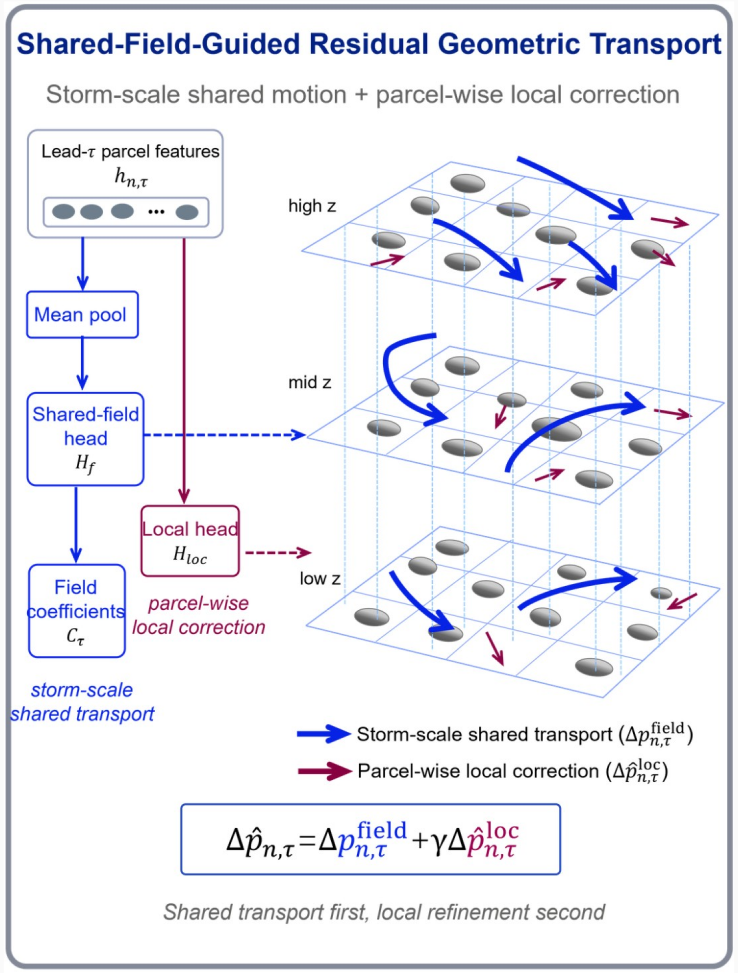


1 · Transport. Move parcel geometry with a shared field plus parcel-wise corrections.

2 · Radiometry. Predict a deterministic base, then sample residuals around it.

3 · Summary. Render every hypothesis and blend a quantile with the base anchor.

Stage 1: shared-field-guided residual transport



Transport first

Predict lead-to-lead increments from the last observed state. Each parcel is carried along one trajectory through all 20 leads.

Height-conditioned shared field

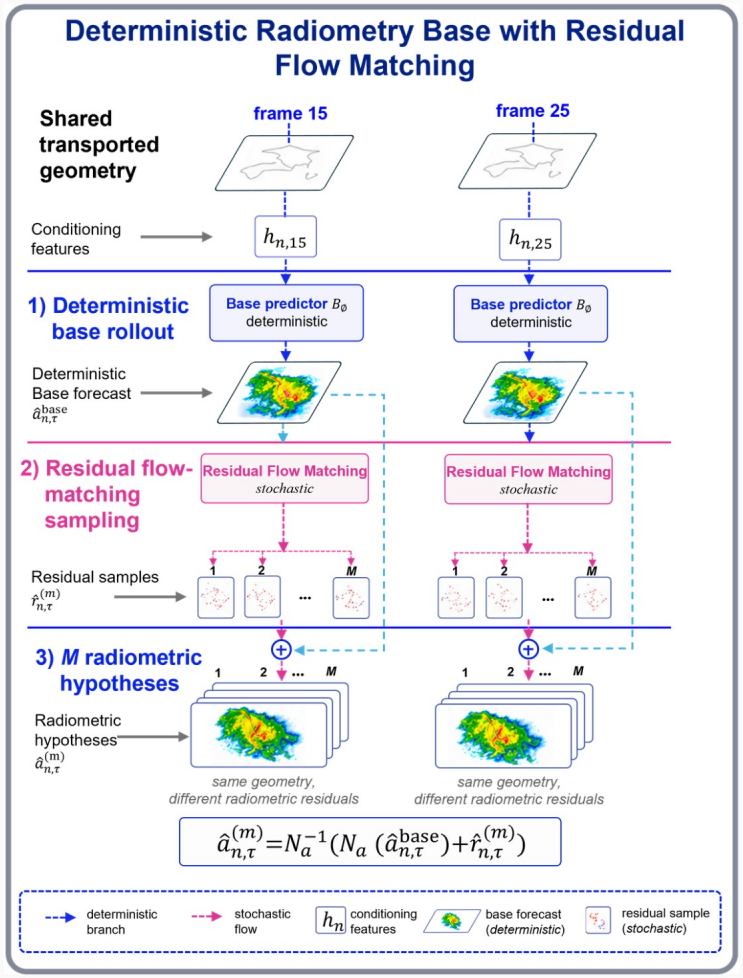
One field per lead for the whole volume: translation plus horizontal scaling, shear and rotation, varying smoothly with height through K = 4 cubic bases. 28 coefficients per lead.

Parcel-wise residual

A linear head corrects each parcel's center, scale and orientation. γ = 0.15 keeps the shared field in charge of the main motion.

Fixed parcel set: the same N indexed Gaussians from the last input to the final lead. Splitting and merging emerge in the rendered echoes.

Stage 2: deterministic base with residual flow matching



Deterministic radiometry base

A base predictor outputs the 6-D radiometry of each transported parcel. Base geometry and base radiometry drive the autoregressive rollout.

Residual flow matching

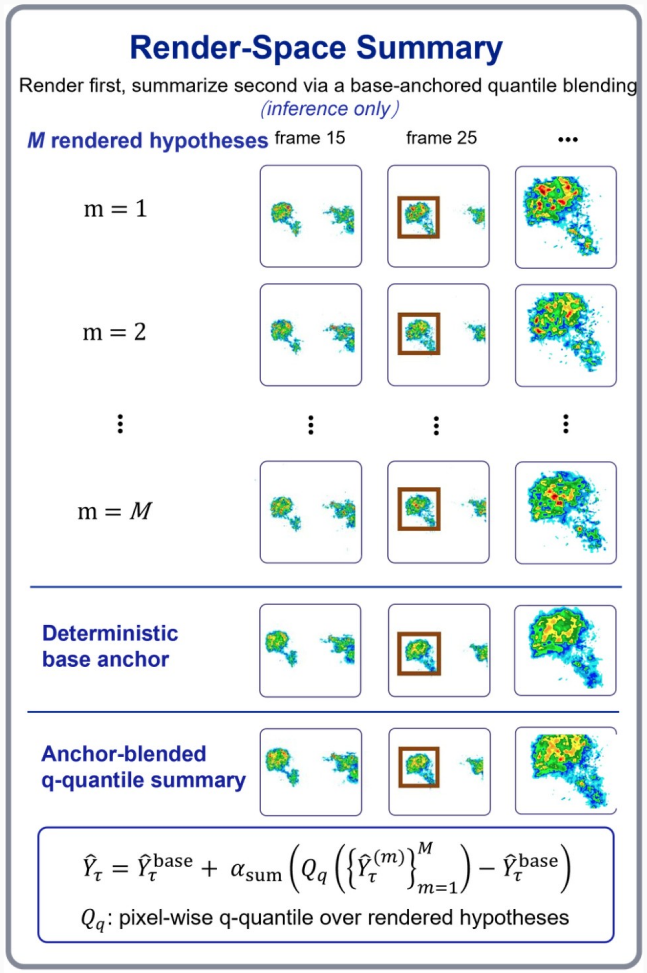
Conditional flow matching samples the normalized residual between truth and base, conditioned on the parcel feature and the base state. 8 Euler steps at evaluation.

Geometry shared across samples

All M hypotheses reuse the transported geometry and differ only in radiometric residuals, so strong echoes keep their place.

Sampling the full Gaussian state would also perturb parcel positions. The ablation shows what that costs.

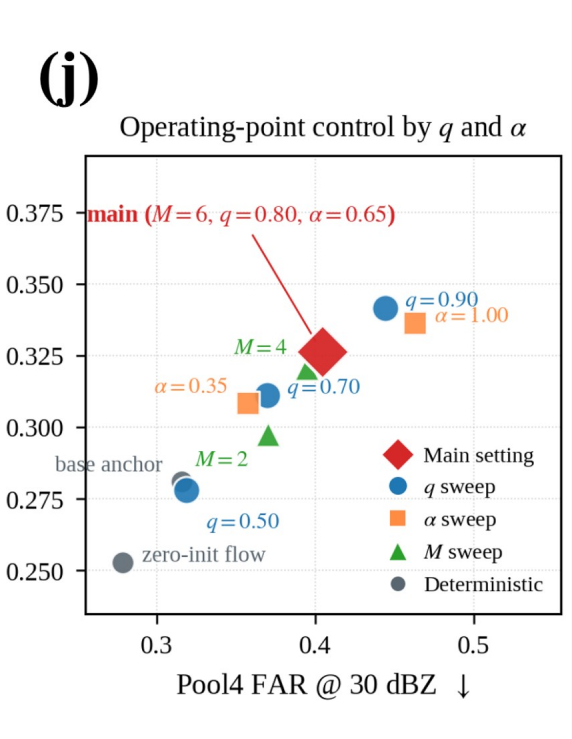
Stage 3: render-space ensemble summary



Each hypothesis is rendered to a radar volume. The final forecast blends a per-voxel q-quantile of the M members with the deterministic base anchor.

- Base anchor: conservative, fewer false alarms
- Quantile: keeps strong echoes present in only some members
- q and α set the hit and false-alarm operating point with no retraining

Main setting, selected on validation: M = 6, q = 0.80, α = 0.65. At 30 dBZ, base anchor only gives CSI 0.281 at FAR 0.316. The main setting gives 0.326 at 0.404.



Results: stronger high-threshold skill

Method (NEXRAD)	LPIPS-Radar	CSI-20	CSI-30	CSI-40
ConvGRU	1.621	0.306	–	–
PhyDNet	2.058	0.311	0.089	0.002
SimVP	2.925	0.085	0.088	0.018
DiffCast	4.057	0.049	0.021	0.021
Mamba	0.699	0.309	0.165	0.074
STC-GS	0.665	0.326	0.179	0.078
GeoRad-3D (ours)	0.459	0.473	0.326	0.175

Pool4 CSI, higher is better. LPIPS-Radar, lower is better. TRRD (tropical convection, Singapore), CSI-20 / 30 / 40: STC-GS 0.127 / 0.077 / 0.045, GeoRad-3D 0.186 / 0.112 / 0.055.

GeoRad-3D vs. STC-GS

CSI AT 30 DBZ

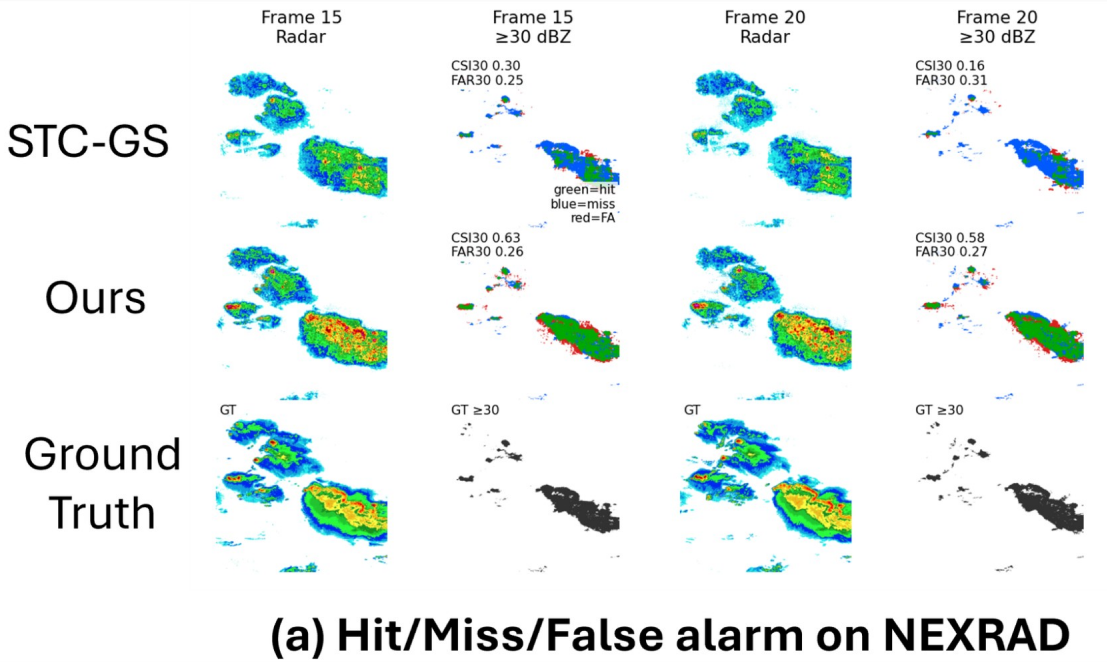
0.179 → 0.326

CSI AT 40 DBZ

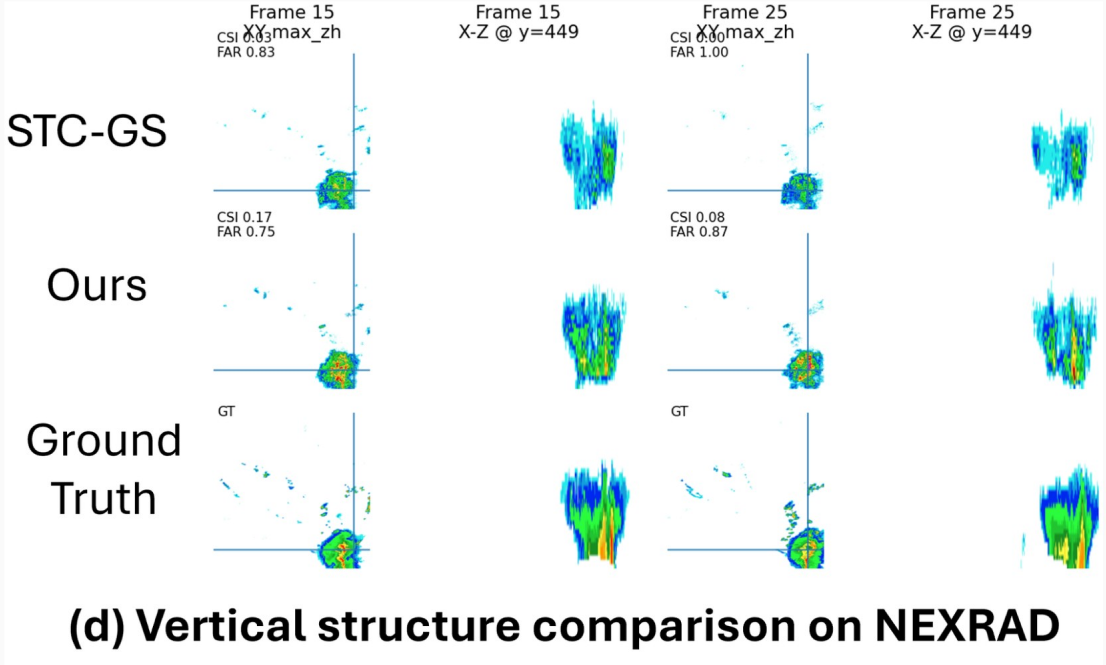
0.078 → 0.175

At matched FAR the gain remains: +0.135 / +0.161 / +0.108 CSI at 20 / 30 / 40 dBZ, with 95% intervals above zero.

Qualitative: placement and vertical structure



NEXRAD, frames 15 and 20. Green = hit, blue = miss, red = false alarm at 30 dBZ.

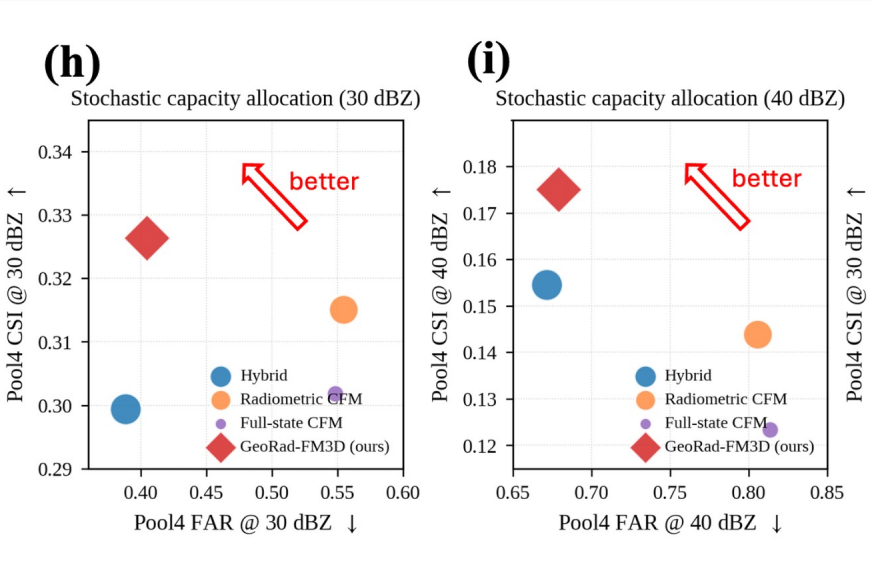


NEXRAD, frames 15 and 25. Column-maximum views and vertical X-Z slices.

GeoRad-3D recovers more strong-echo area with few additional false alarms. Its storm column keeps the vertical extent and tilt, while the STC-GS column collapses toward low altitudes.

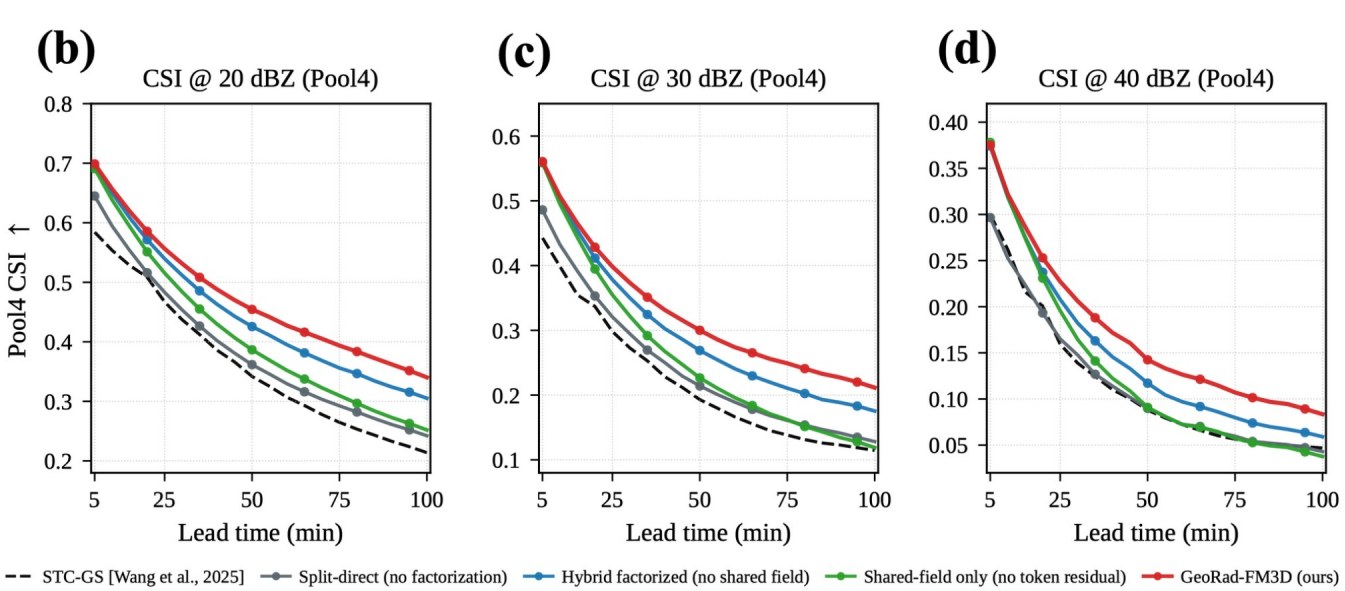
Ablation: where to allocate stochastic capacity

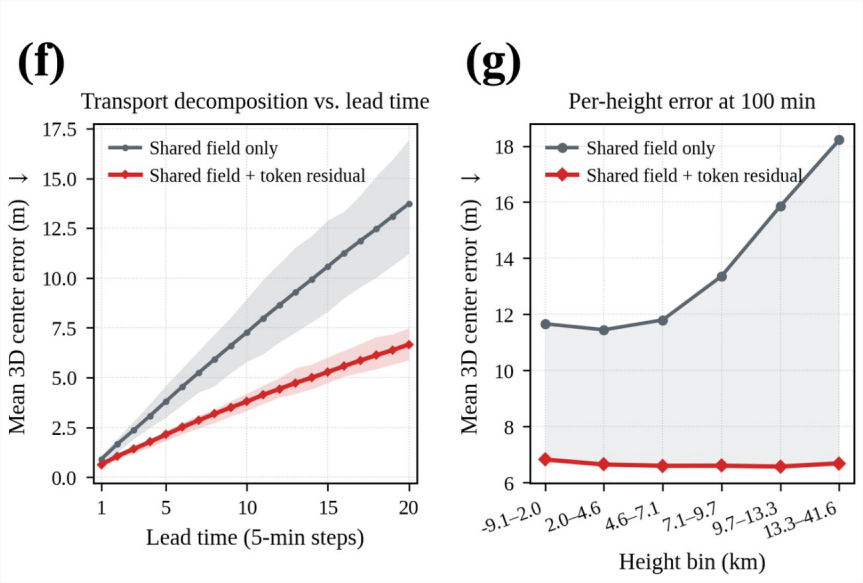
Variant	CSI-30	FAR-30	CSI-40	FAR-40
Split-direct	0.246	0.422	0.124	0.711
Split-transport	0.255	0.312	0.129	0.603
Radiometric CFM only	0.315	0.555	0.144	0.806
Full-state CFM	0.302	0.549	0.123	0.813
Hybrid factorized	0.299	0.388	0.155	0.672
GeoRad-3D (ours)	0.326	0.404	0.175	0.679



- Transport first: FAR drops 28.5 / 26.1 / 15.2% at 20 / 30 / 40 dBZ with no sampling
- Full-state sampling shifts echo positions across members: CSI-40 falls from 0.144 to 0.123
- Residual CFM around a base: FAR-30 from 0.555 to 0.388. Shared field: CSI-30 / 40 up to 0.326 / 0.175

Geometry pays off at long leads





In the legends, GeoRad-FM3D is GeoRad-3D and token residual is the parcel-wise residual.

- CSI-40 at 80–100 min: 0.093 for GeoRad-3D, 0.067 for Hybrid, 0.046 for shared field only
- Gain over Hybrid at CSI-40 grows from 3.5% in steps 1–5 to 38.8% in steps 16–20
- Lead-to-lead transport lowers Gaussian-center RMSE by 16.9% at 100 min vs. direct multi-lead geometry

Takeaways

1. Factorize the forecasting head: geometry follows coherent transport, radiometry carries the local variation of strong echoes.

2. Sample only the radiometric residual on shared geometry, then summarize the hypotheses in radar-product space.

3. NEXRAD Pool4 CSI at 20 / 30 / 40 dBZ: 0.326 / 0.179 / 0.078 → 0.473 / 0.326 / 0.175, with the largest gains for strong echoes and long leads.

Limits: forecast quality follows the quality of the STC-GS parcels, physical conservation is left unenforced, and other radar networks and climates await evaluation.

Thank you

Supported by the National Research Foundation, Singapore and the National Environment Agency, Singapore under the Weather Science Research Programme (WSRP-2025-1R-03-03).