

AI-BASED NOWCASTING IN EXTREME WEATHER PREDICTION

WMO AI for Nowcasting Pilot Project (AINPP) Workshop

Session 3 – AI based Nowcasting

25 September 2025

Wai Kin Wong

E-mail: wkwong@hko.gov.hk

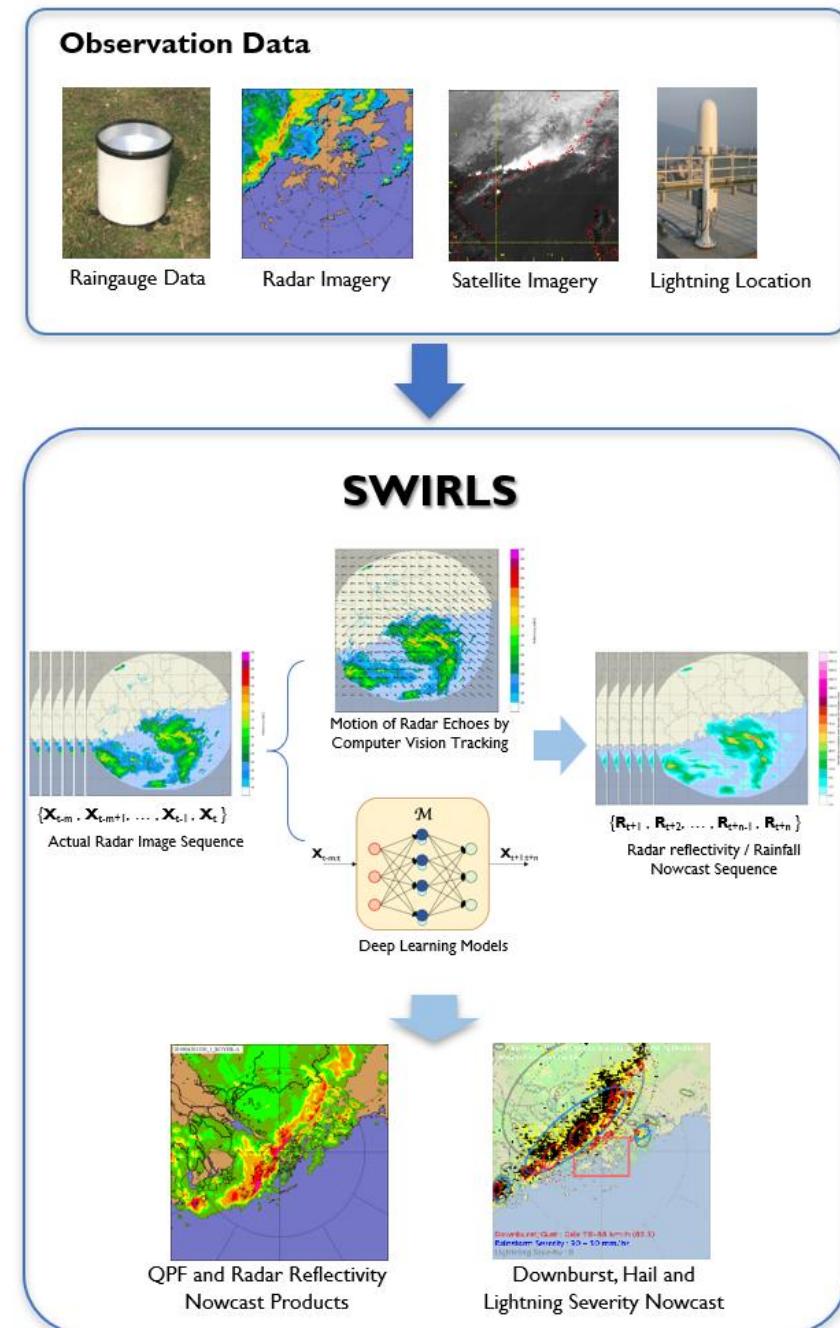


Outline

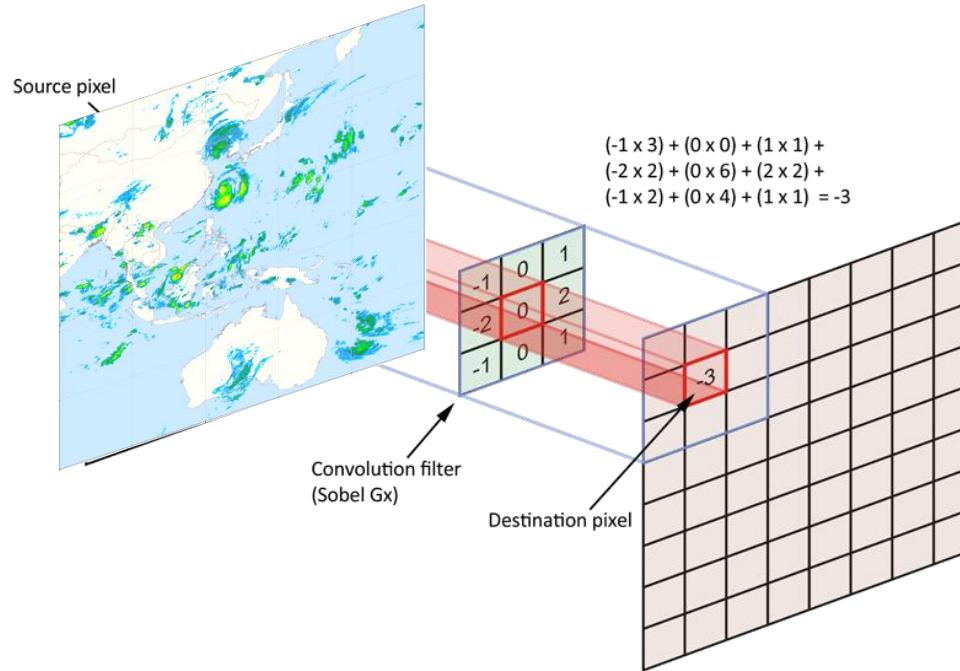
1. Evolution of Deep Learning Precipitation Nowcast Models in HKO Nowcasting System SWIRLS
2. Regional Convection / Precipitation Nowcast Supports using Deep Learning Models
3. Ensemble nowcast and AI-based guidance for rainstorm / extreme precipitation
4. Way forward on AI nowcast support

SWIRLS

Short-range Warning of Intense Rainstorms in Localized Systems



Encoding-forecasting ConvLSTM network



Convolution operator

- Good at feature extraction
- More intuitive model architecture

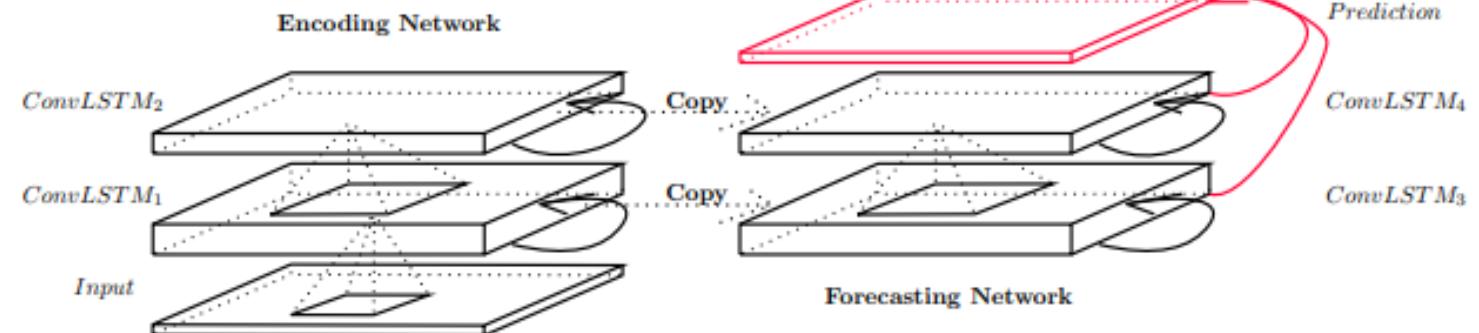
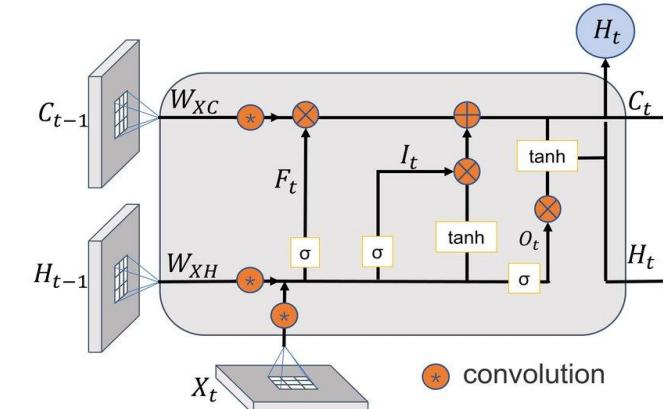


Figure 3: Encoding-forecasting ConvLSTM network for precipitation nowcasting

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. NIPS 2015.

<https://arxiv.org/abs/1506.04214>

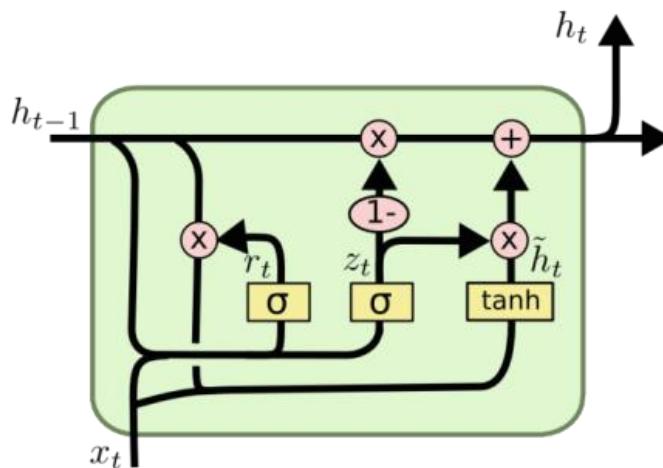
Trajectory Gated Recurrence Unit (TrajGRU)

Xingjian Shi, Zhihan Gao, Leonard Lausen, Hao Wang, Dit-Yan Yeung, Wai-kin Wong, and Wang-chun Woo, 2017: Deep learning for precipitation nowcasting: A benchmark and a new model. <https://arxiv.org/pdf/1706.03458.pdf>

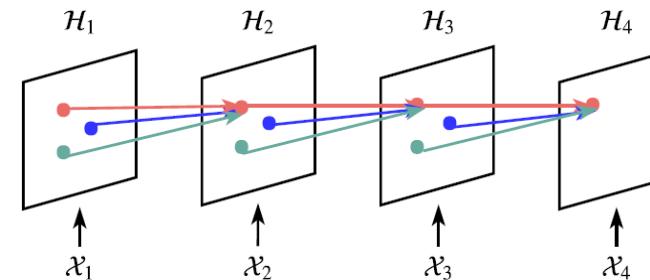
TrajGRU replaces LSTM, introduces “Trajectory” and adopts weighted error function

“HKO-7” dataset available to community to promote development of AI in nowcasting

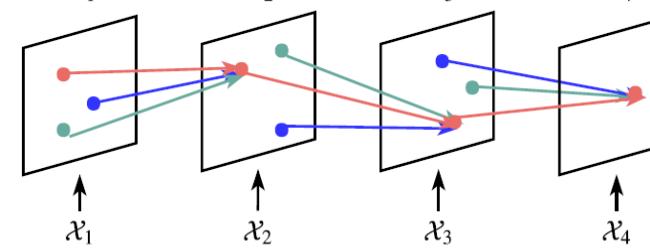
GRU (Gated Recurrent Unit) includes *reset gate* and *update gate*, similar to LSTM but more efficient.



Trajectory:
Recurrent connections are dynamically determined



(a) ConvRNN: Links are fixed over time/location.



(b) TrajRNN: Links are dynamically determined.

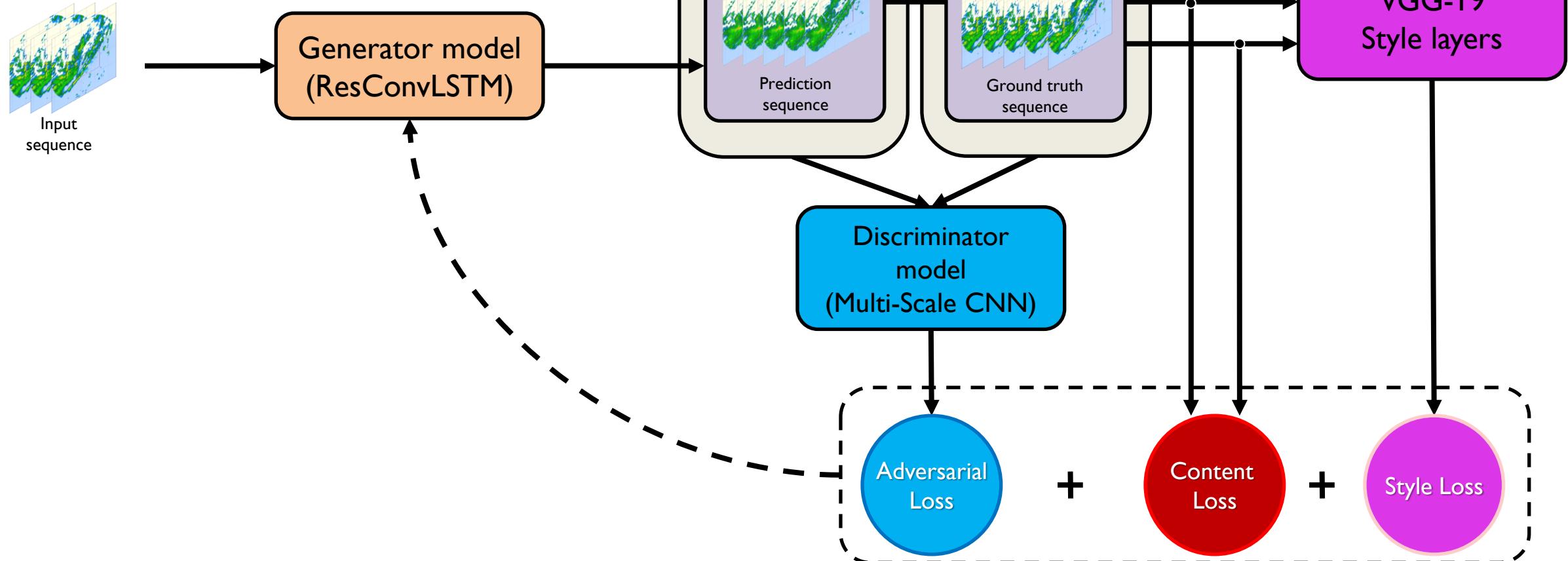
Weighted Error:
optimize performance in heavy rain

$w(x)$ to each pixel according to its rainfall intensity x : $w(x) = \begin{cases} 1, & x < 2 \\ 2, & 2 \leq x < 5 \\ 5, & 5 \leq x < 10 \\ 10, & 10 \leq x < 30 \\ 30, & x \geq 30 \end{cases}$. Also, the

masked pixels have weight 0. The resulting B-MSE and B-MAE scores are computed as $B\text{-MSE} = \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^{480} \sum_{j=1}^{480} w_{n,i,j} (x_{n,i,j} - \hat{x}_{n,i,j})^2$ and $B\text{-MAE} = \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^{480} \sum_{j=1}^{480} w_{n,i,j} |x_{n,i,j} - \hat{x}_{n,i,j}|$, where N is the total number of frames and $w_{n,i,j}$ is the weight corresponding to the (i,j) th pixel in the n th frame. For the conventional MSE and MAE measures, we simply set all the weights to 1 except the masked points.

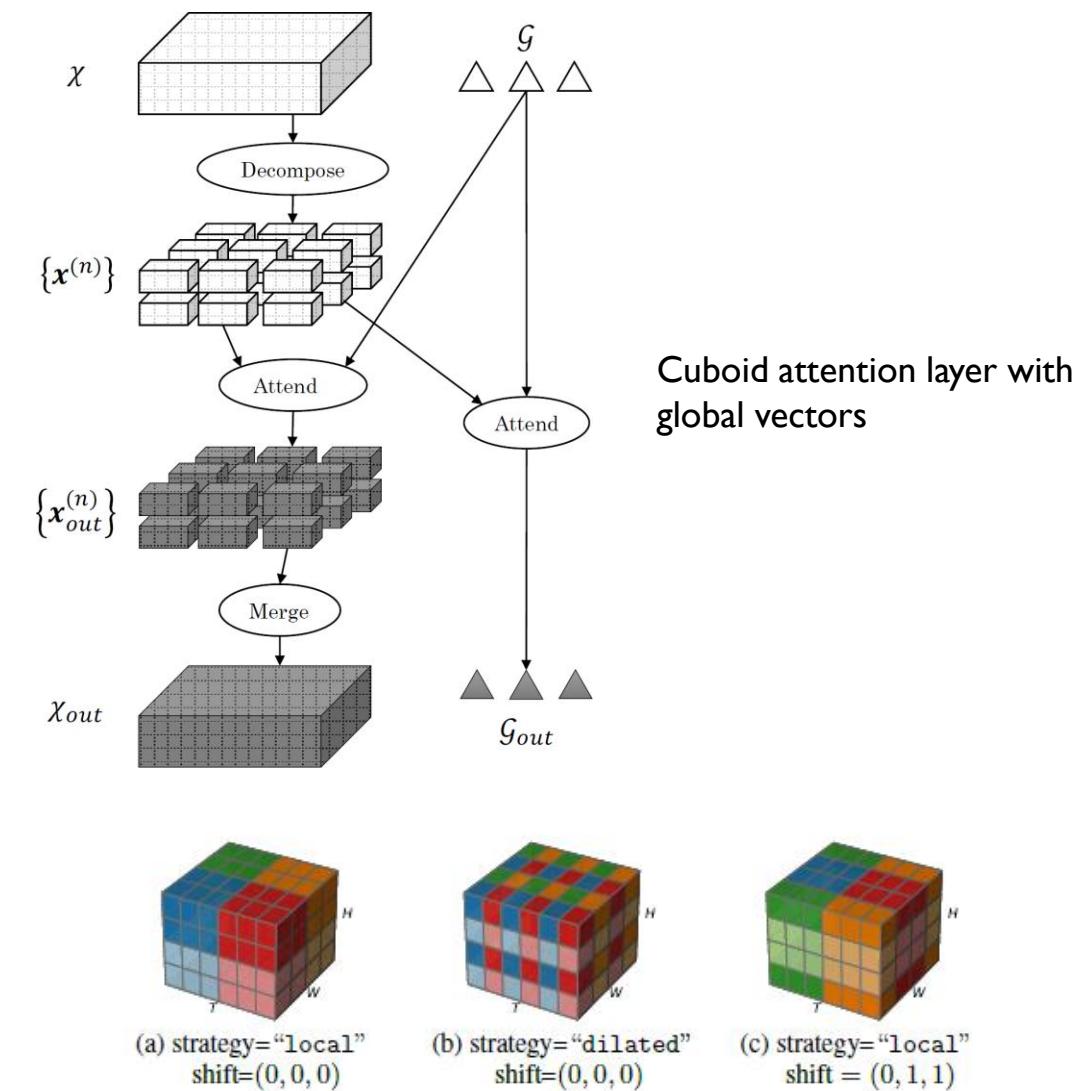
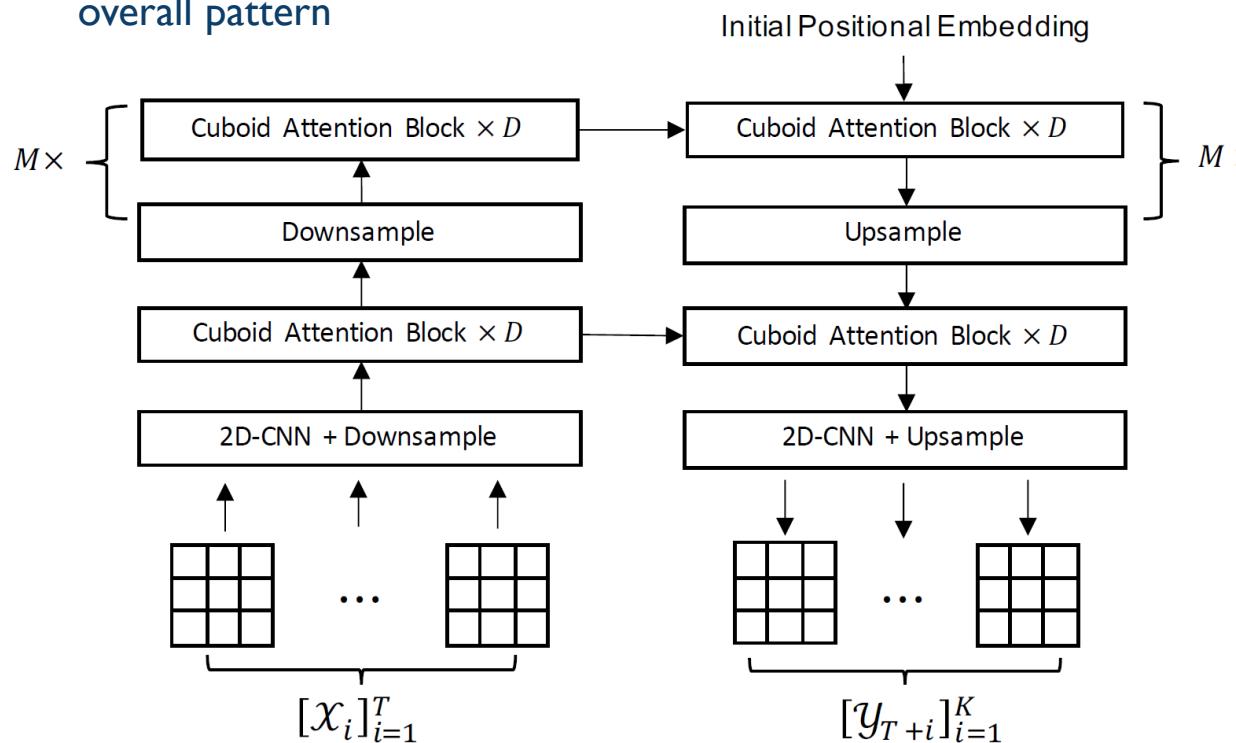
ResConvLSTM-GAN (RCLG)

- **ConvLSTM** with **Residual** connections in encoder-forecaster network
- **Generative Adversarial Network (GAN)** to improve representation of small-scale features



Earthformer

- **Space-time Transformer** model based on Cuboid Attention
- Spatiotemporal data are divided into non-overlapping cuboids, and **self-attention** is applied locally within each cuboid
- **Global vectors** are used to connect the cuboids for capturing the overall pattern



Cuboid decomposition strategies in attention layer

Enhanced feature preservation using Fourier Amplitude and Correlation Loss (FACL)

Ref: <https://arxiv.org/abs/2410.23159>

DFT of image X
of height M and
width N

$$F_{pq} = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X_{mn} e^{-i2\pi(\frac{mp}{M} + \frac{nq}{N})}$$

Fourier
frequencies

$$\text{FAL}(X, \hat{X}) = \frac{1}{MN} \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} (|F|_{pq} - |\hat{F}|_{pq})^2$$

Image
structure

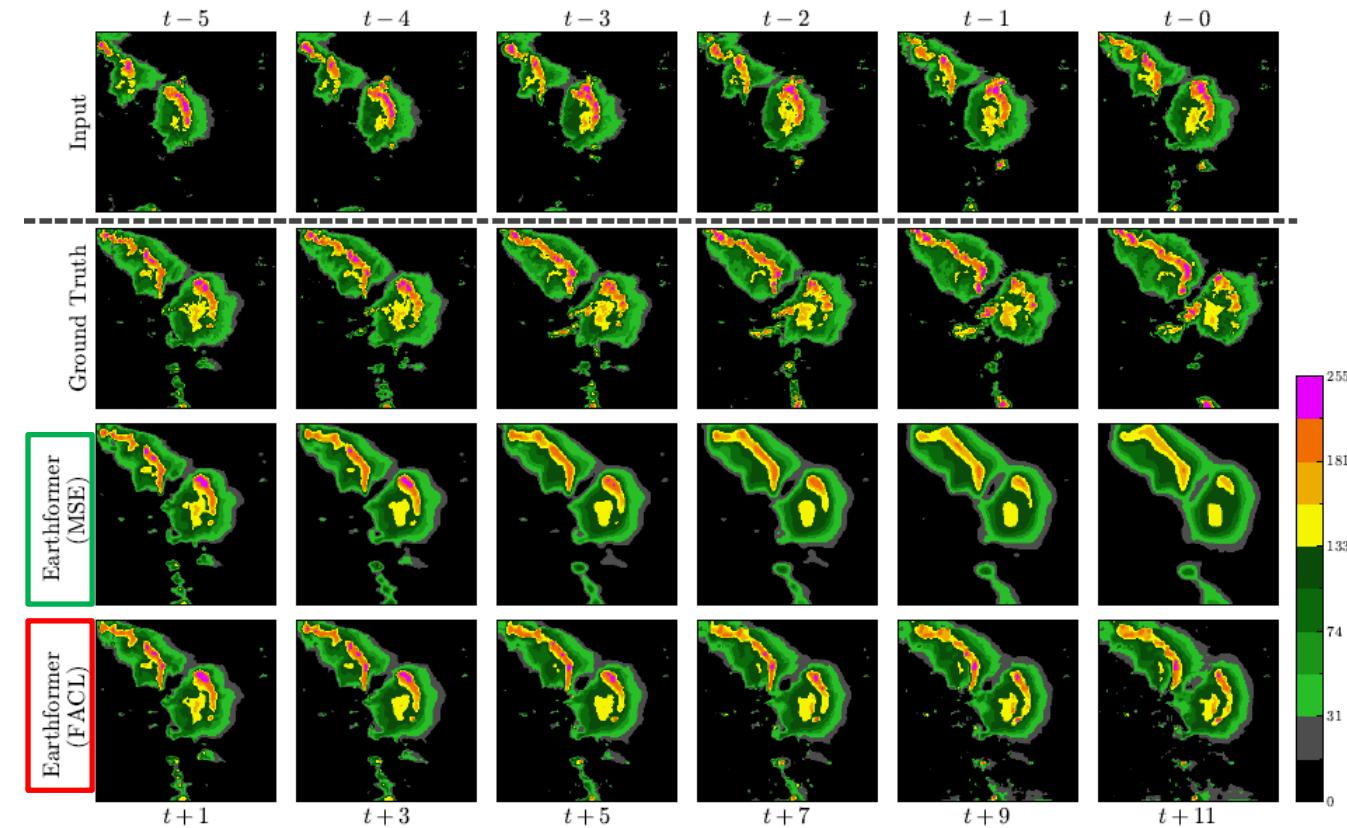
$$\text{FCL}(X, \hat{X}) = 1 - \frac{\frac{1}{2} \sum [F\hat{F}^* + \hat{F}F^*]}{\sqrt{\sum |F|^2 \sum |\hat{F}|^2}}$$

A new metric **RHD (Regional Histogram Divergence)** to
improve over traditional L1 loss, L2 loss and FSS (Fraction Skill Score)
to measure patch-wise similarly between two spatiotemporal patterns

$$\text{FSS} = 1 - \frac{\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} (F_{i,j} - O_{i,j})^2}{\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} F_{i,j}^2 + \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} O_{i,j}^2}$$

$$\text{RHD} = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} D_{\text{KL}}(O'_{i,j} \| F'_{i,j}) = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \sum_{x \in \mathcal{X}} O'_{i,j}(x) \log \frac{O'_{i,j}(x)}{F'_{i,j}(x)}$$

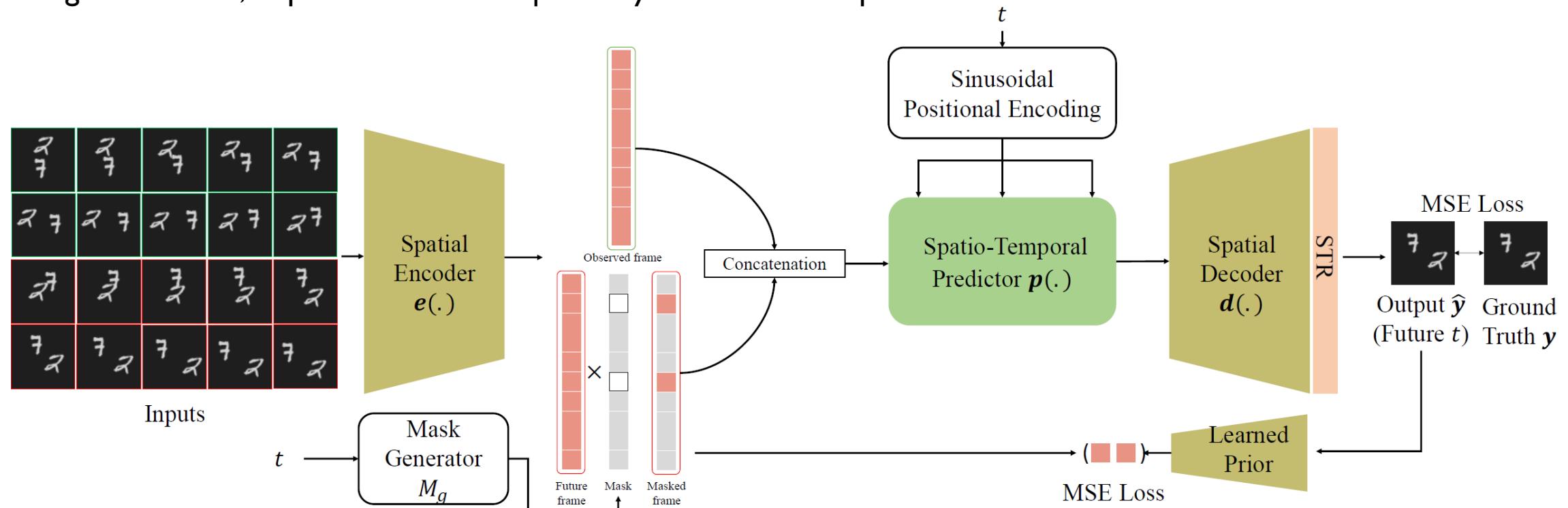
Predicted frames of the Earthformer models trained with **MSE** and **FACL**

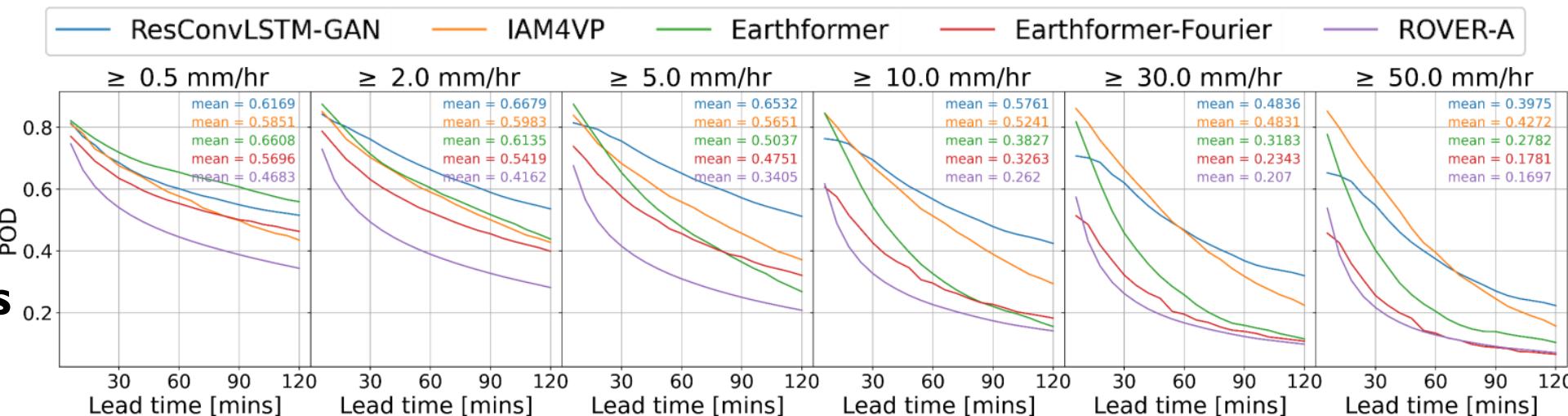


F' and O' is predicted and observed discrete probability distribution respectively
→ RHD penalizes more on “blur” pattern

Implicit Stacked Autoregressive Model for Video Prediction (IAM4VP)

A stacked autoregressive method applied in an implicit video prediction model resulting in improved performance at longer lead time, as predictions are sequentially stacked in the queue

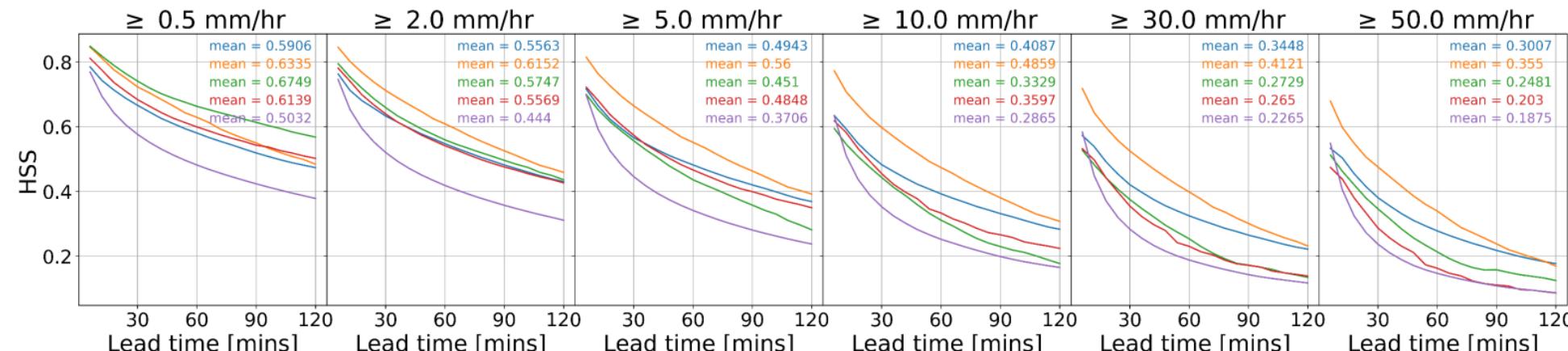
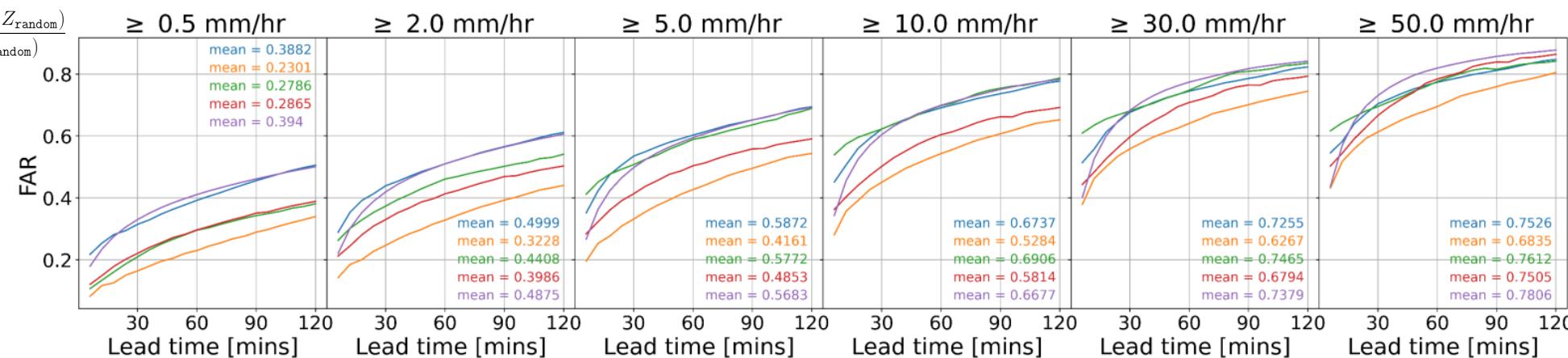




Verification Metrics

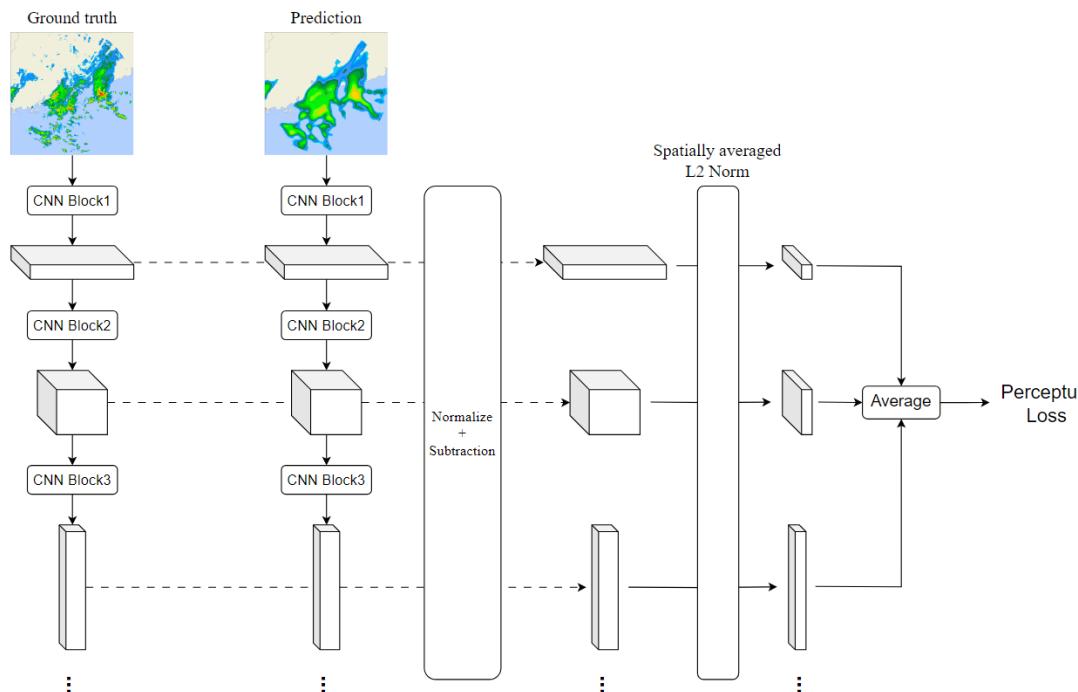
POD, FAR, HSS

$$HSS = \frac{POC - POC_{\text{random}}}{1 - POC_{\text{random}}} = \frac{(H + Z) - (H_{\text{random}} + Z_{\text{random}})}{N - (H_{\text{random}} + Z_{\text{random}})}$$



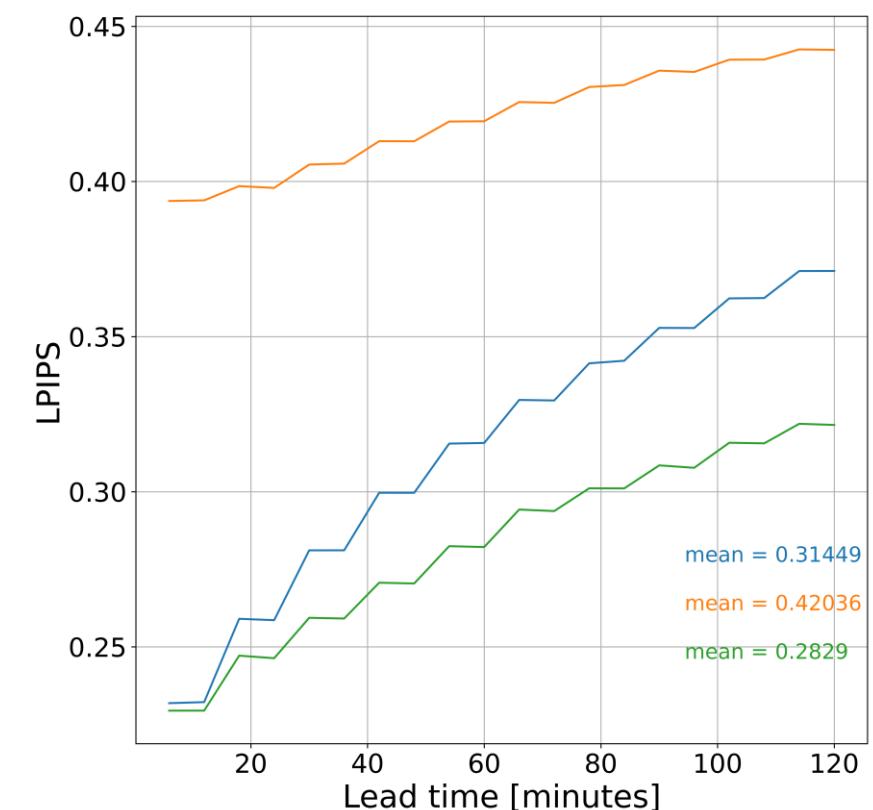
Verification Metric - Perceptual Similarity

- Learned Perceptual Patch Similarity (LPIPS) metric
 - A Neural Network(NN)-based metric to match human perception, lower the better
 - Measure the difference of activations in a pre-defined NN between 2 images
 - Pre-defined NN: AlexNet
- More realistic features generated in ResConvLSTM-GAN nowcast compared to TrajGRU and optical flow extrapolation



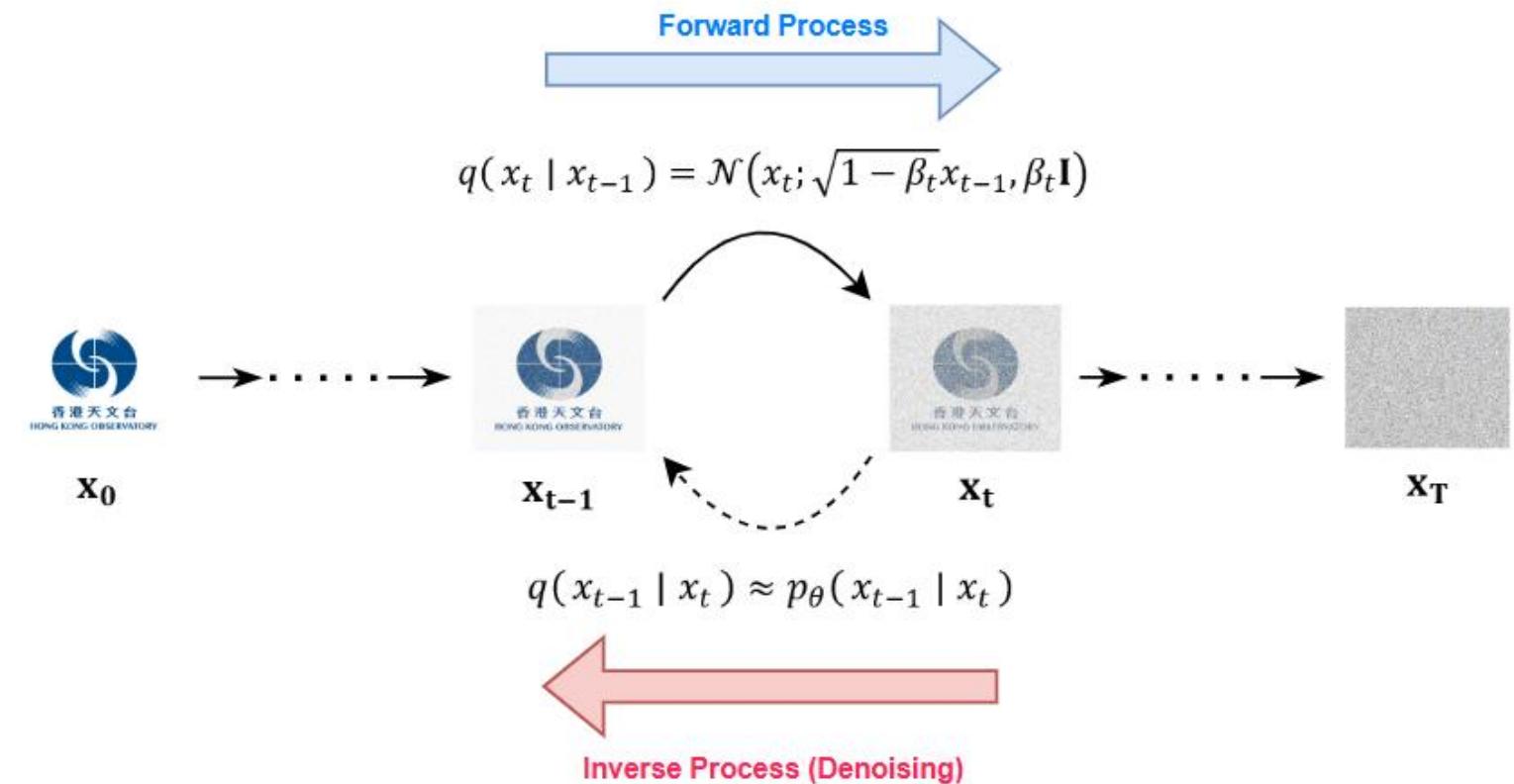
— rover_nonlinear
— trajGRU
— ResConvLSTM-GAN

LPIPS - Lead time plot



Denoising Diffusion Probabilistic Model (DDPM)

- Diffusion Model consists of two processes: **Forward Process** (noising) and **Inverse Process** (denoising)
- To capture the data distribution by training a neural network to undo a Markov noising process that gradually distorts the data



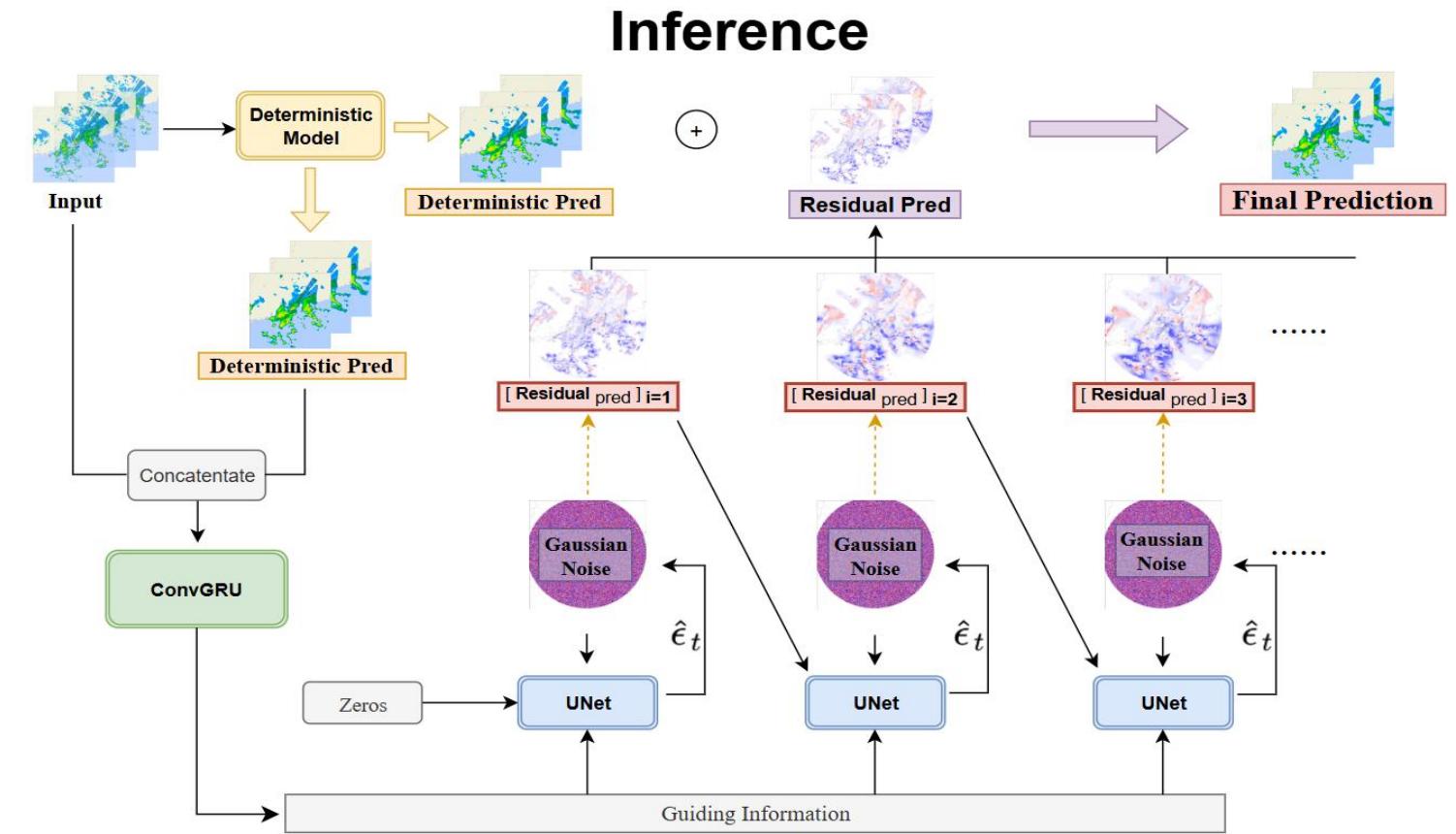
DDPM implementation



$$y = \hat{y}(x) + r$$

target blurry prediction residual

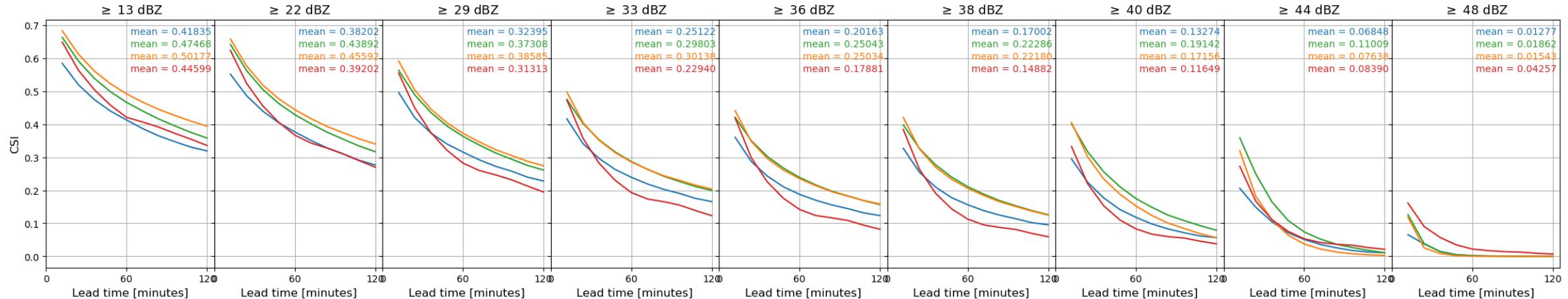
- “Deterministic model”: any model that makes “blurry prediction” such as CNN or Transformer based models
- Diffusion model in learning to generate the residual based on past radar frames and forecast from the “Deterministic model” and able to generate more realistic forecast



In the denoising process, different outputs could be generated and thus producing an **ensemble prediction**

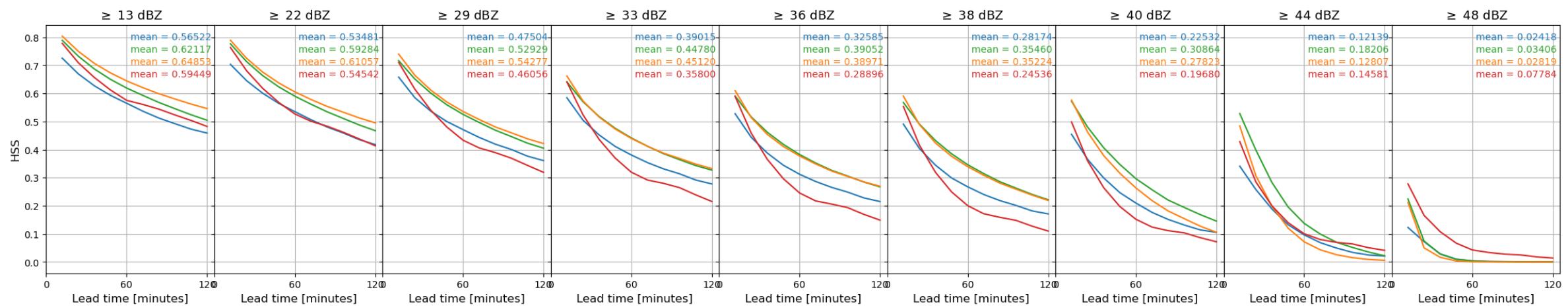
RCLG
Earthformer
VMRNN
Diffusion

CSI

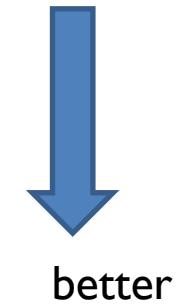
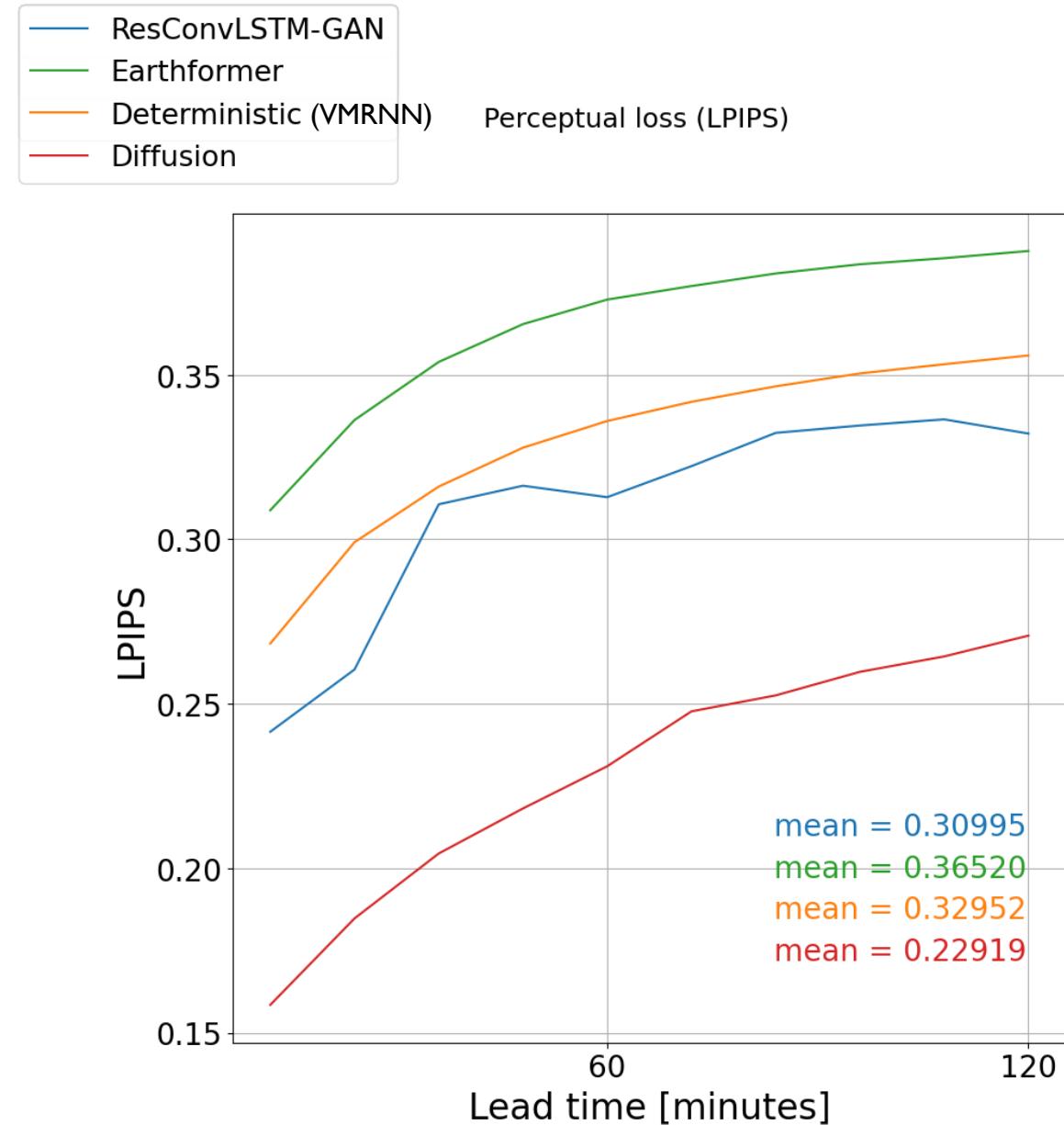


HSS

$$HSS = \frac{POC - POC_{\text{random}}}{1 - POC_{\text{random}}} = \frac{(H + Z) - (H_{\text{random}} + Z_{\text{random}})}{N - (H_{\text{random}} + Z_{\text{random}})}$$

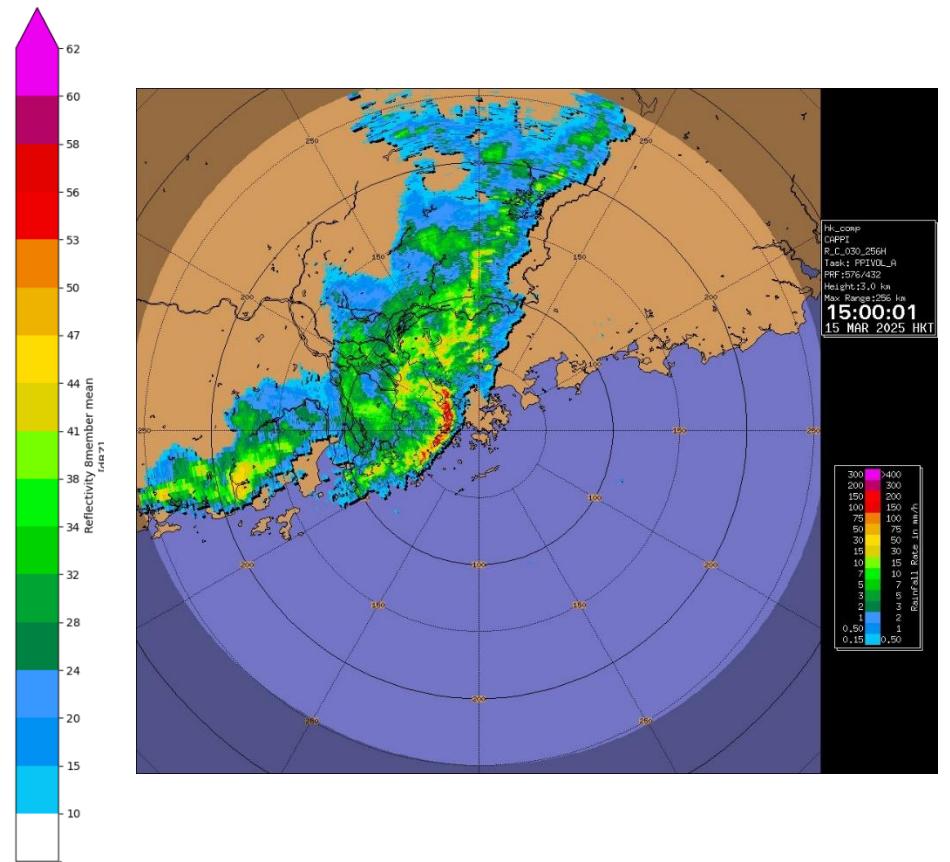
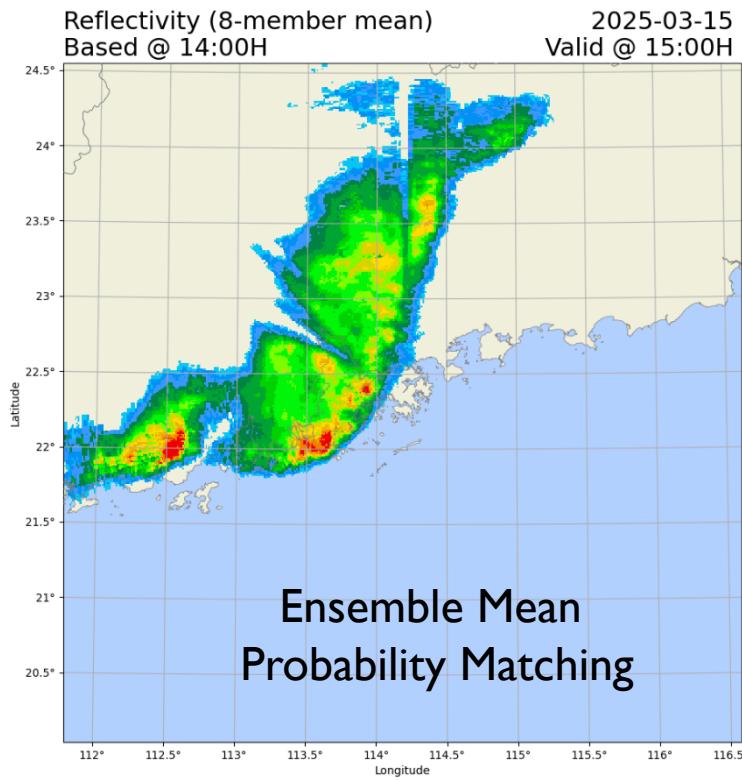
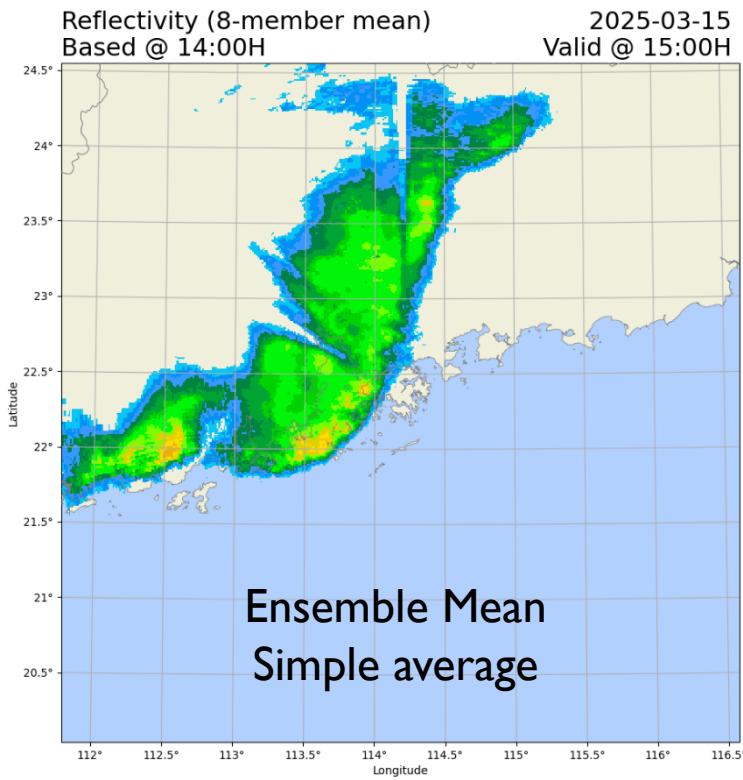
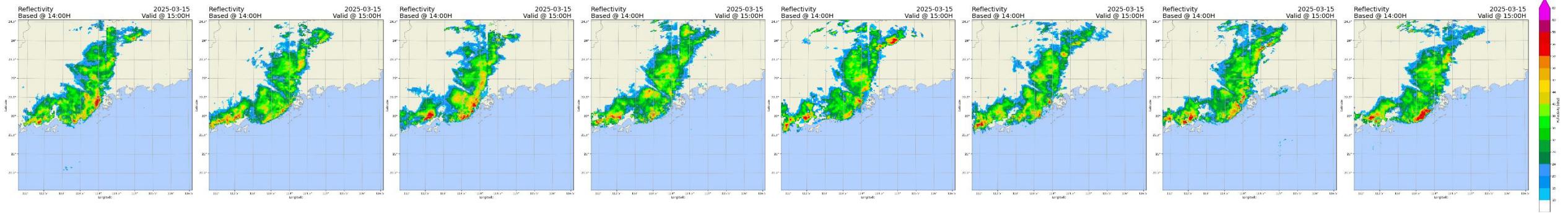


Verification Metric - Perceptual Similarity



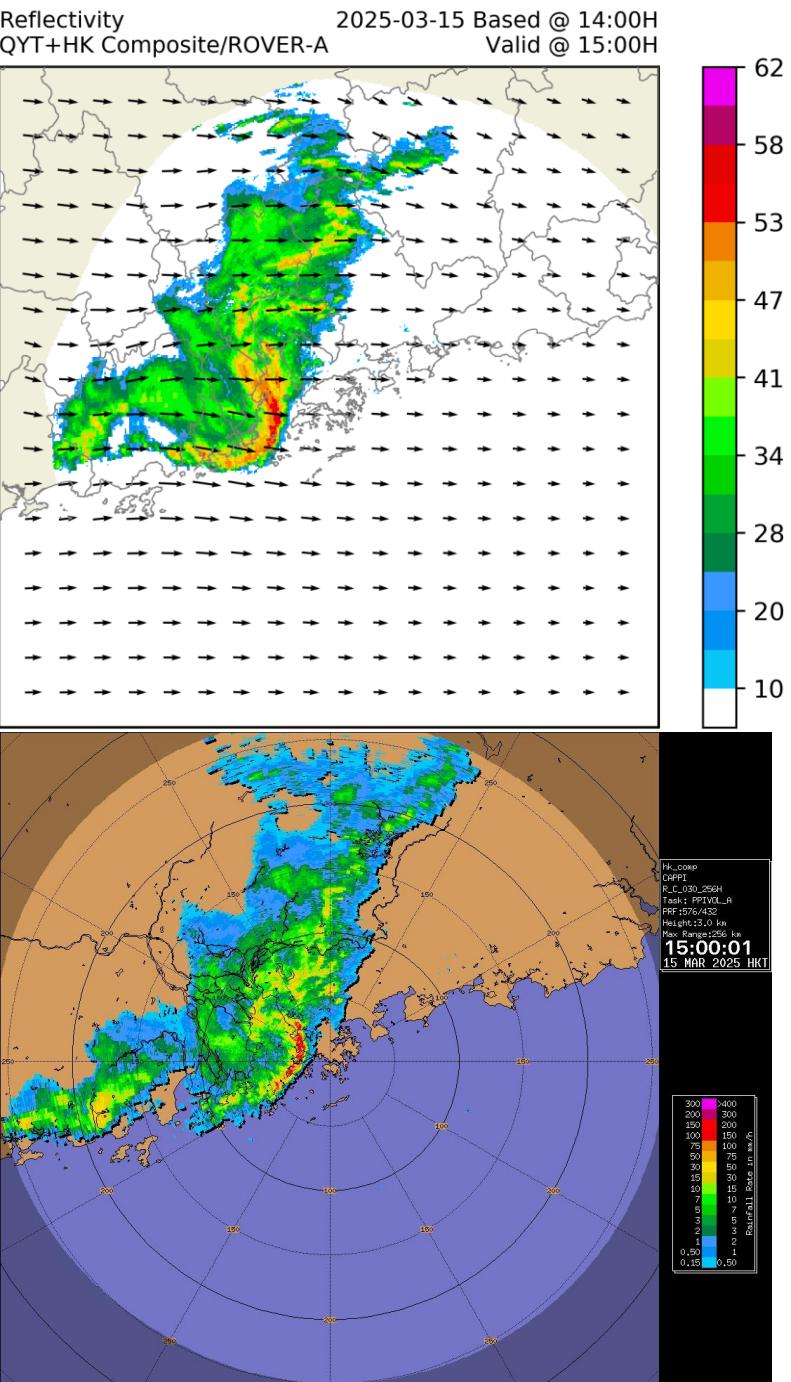
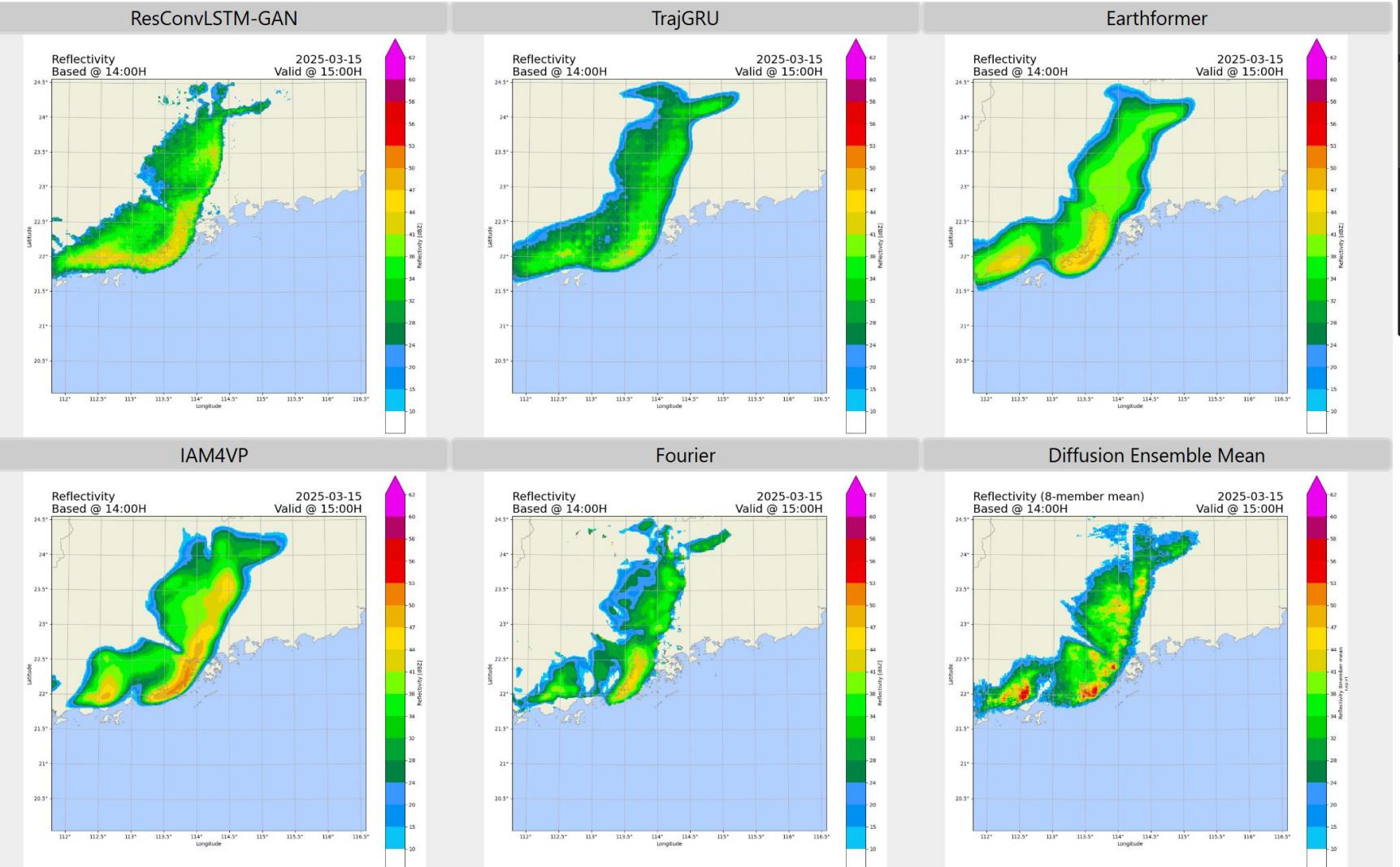
Diffusion Nowcast (8) Ensemble Members at T+60 min

Base time: 2025-03-15 14:00 HKT



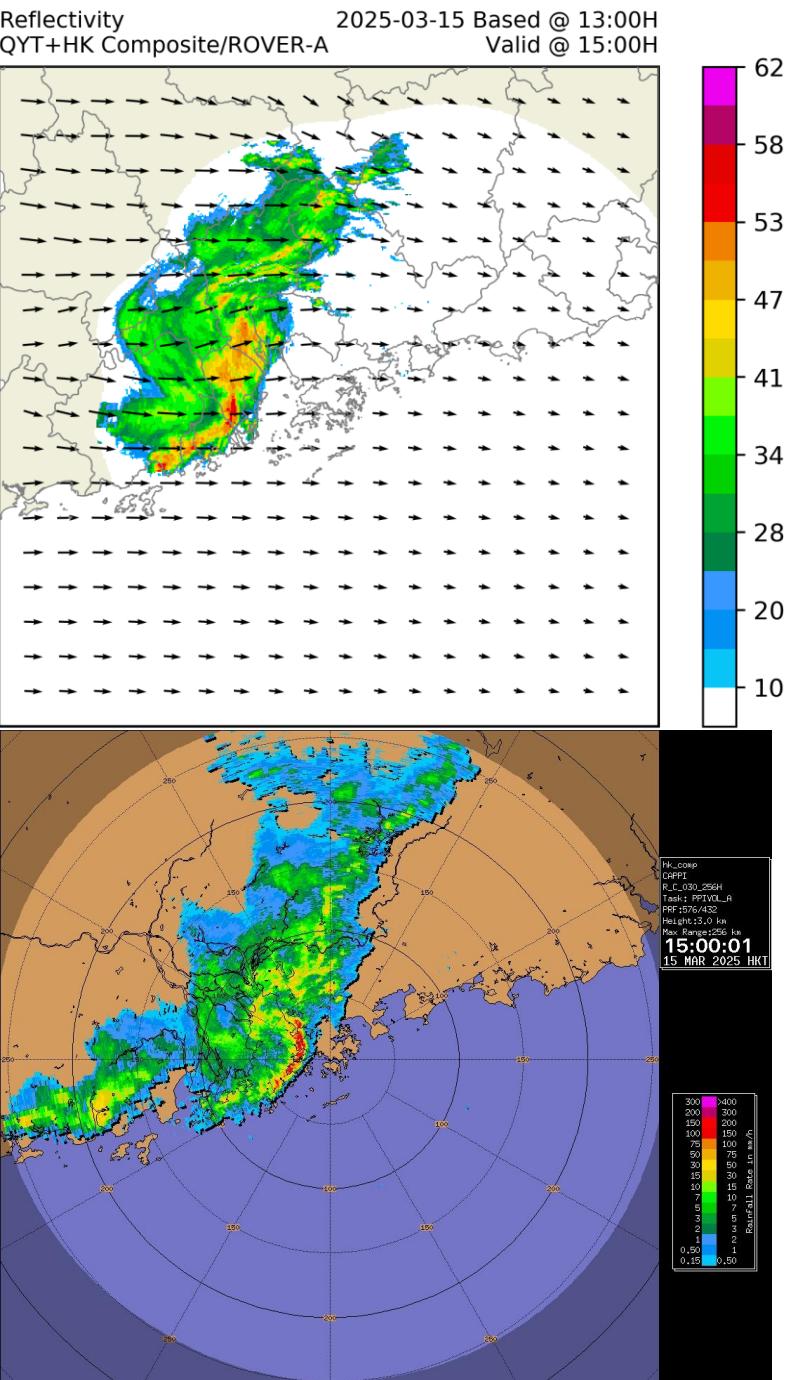
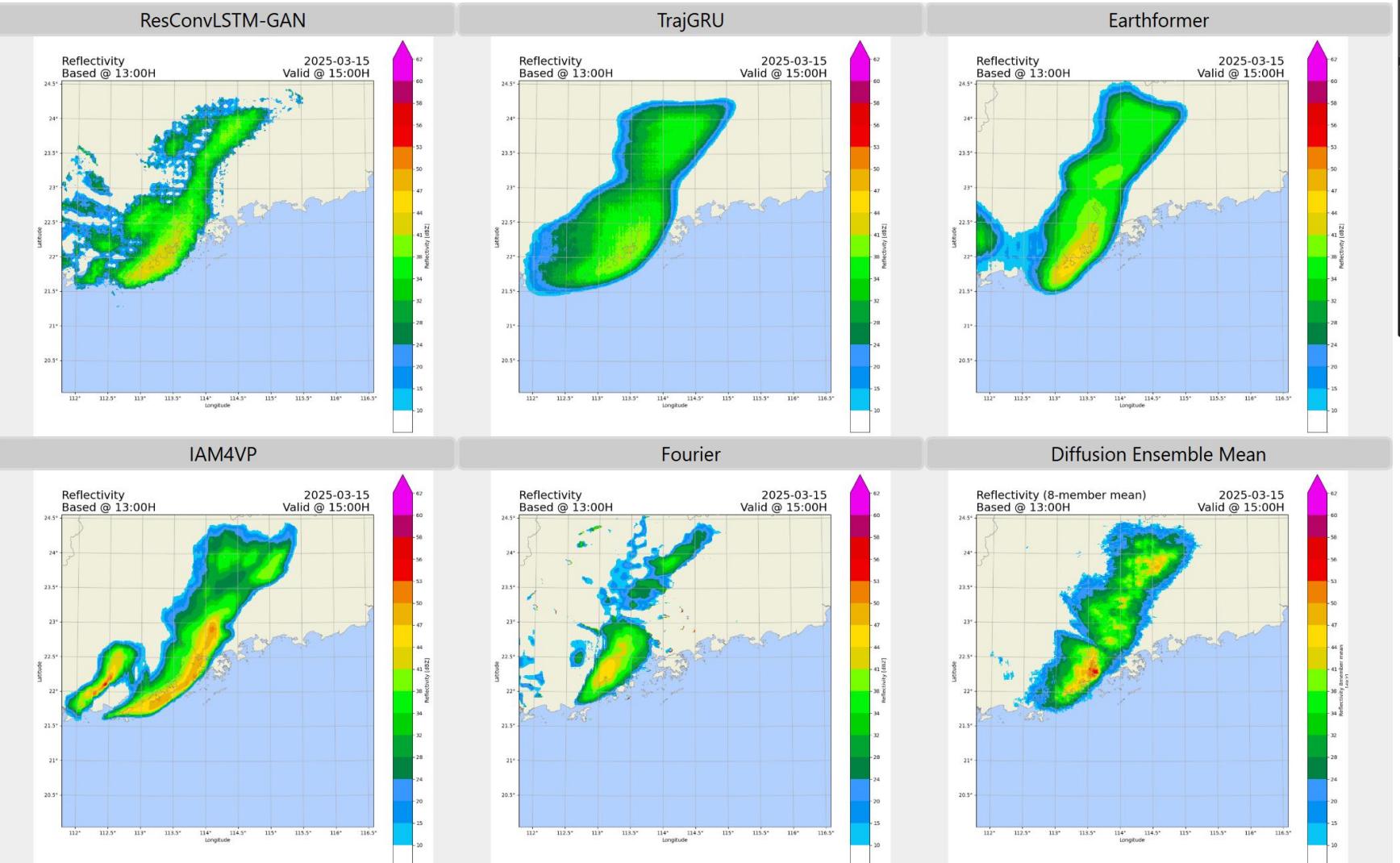
I-hour nowcasts from extrapolation and multi-model deep learning techniques

Base time: 2025-03-15 14:00 HKT



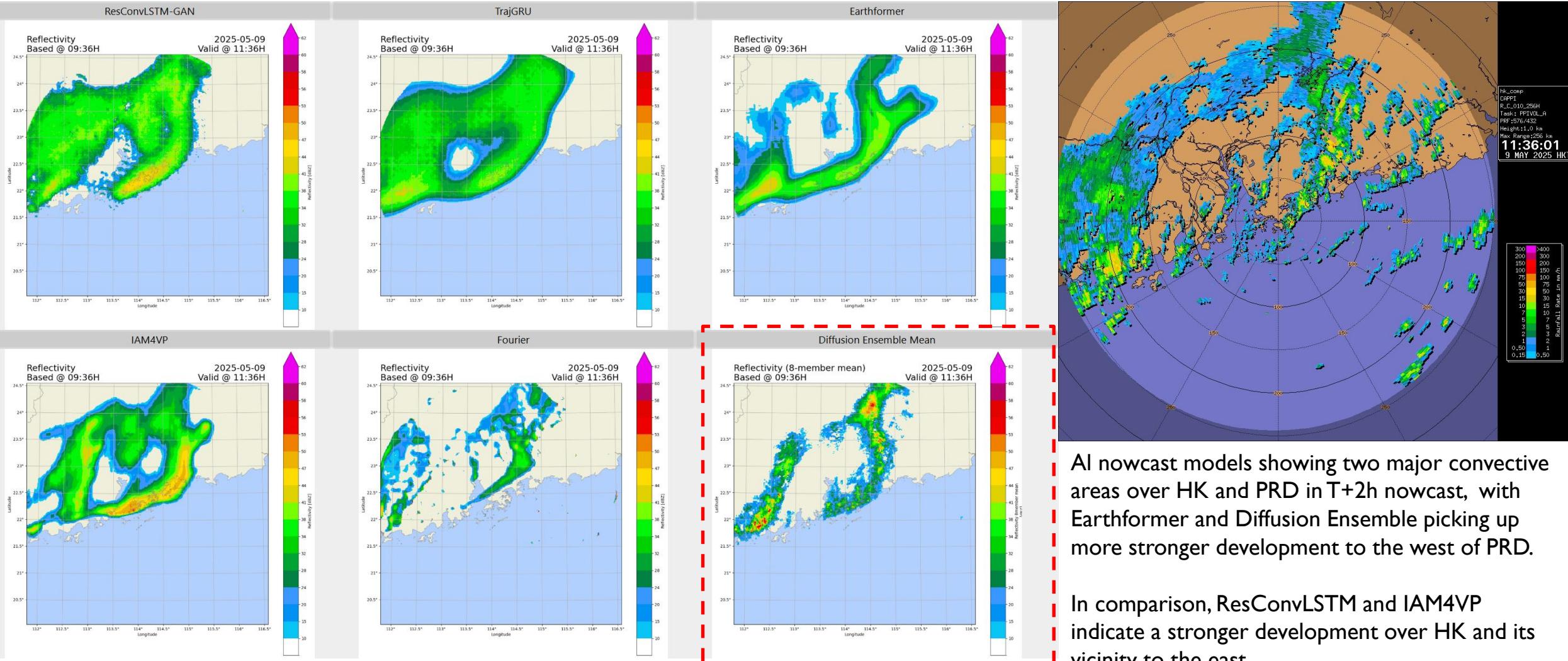
2-hour nowcasts from extrapolation and multi-model deep learning models

Base time: 2025-03-15 13:00 HKT



T+120 min nowcasts from deep learning models

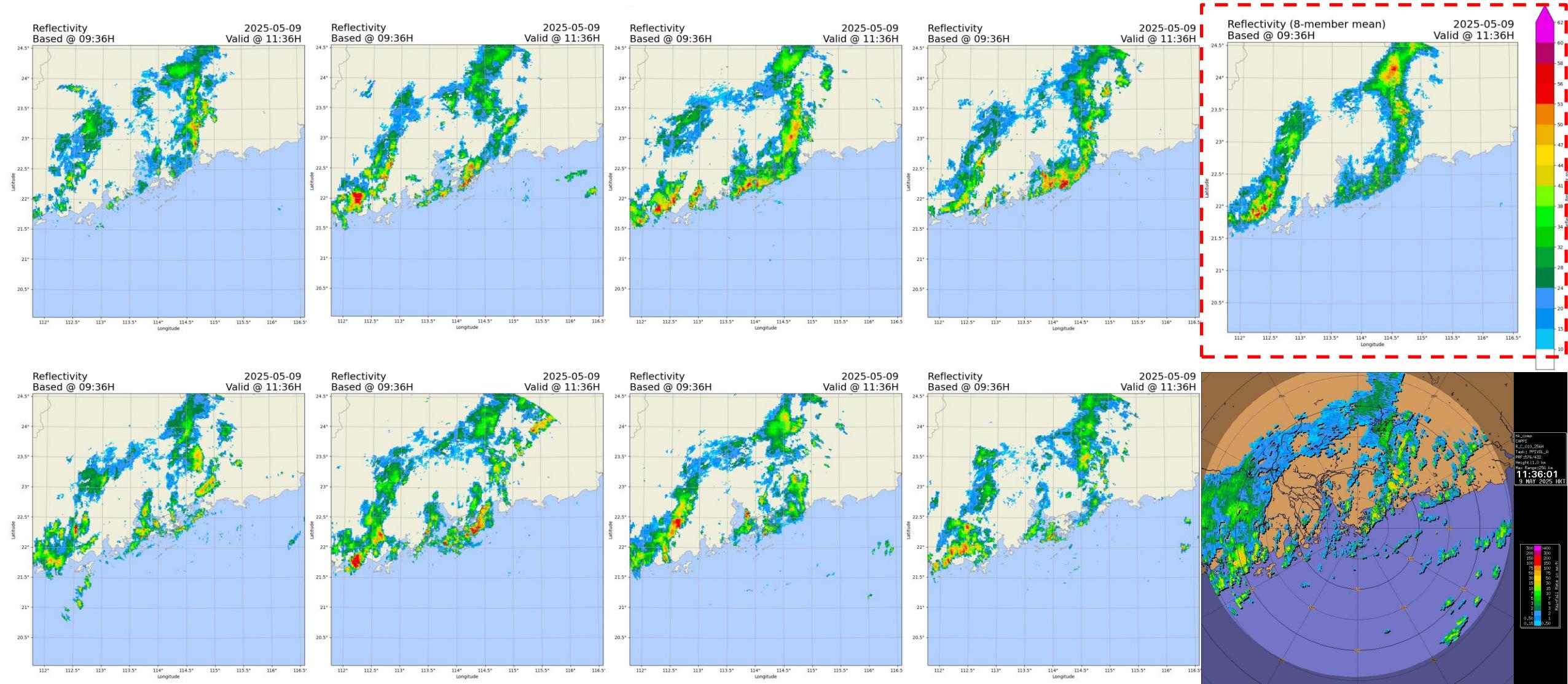
2025-05-09 11:36 HKT



Diffusion ensemble members having different intensity / distribution of convection over HK; while “more agreeable” on the stronger convective development over the western part of the coastal areas → ensemble consensus producing more realistic indication on the difference in their intensities

T+120 min nowcasts from DDPM members

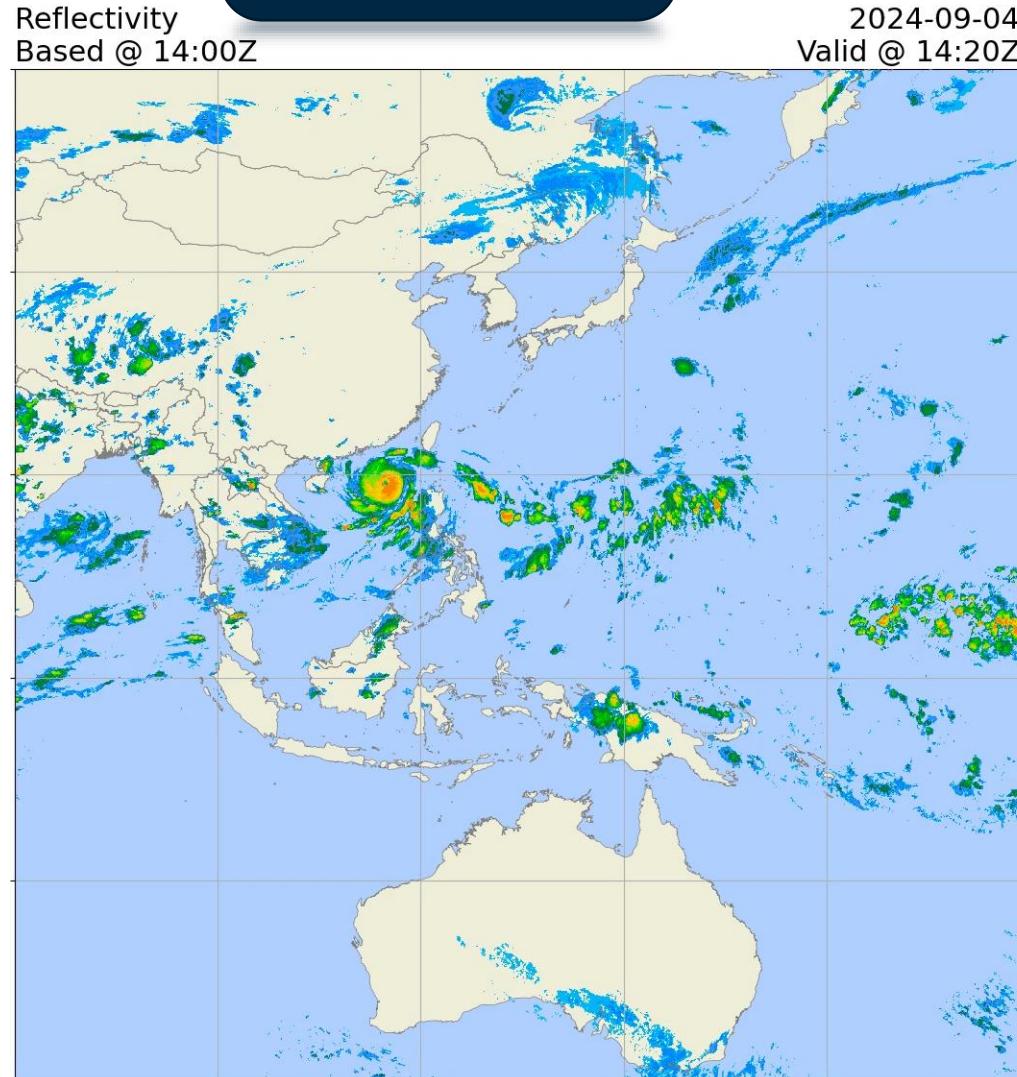
2025-05-09 11:36 HKT



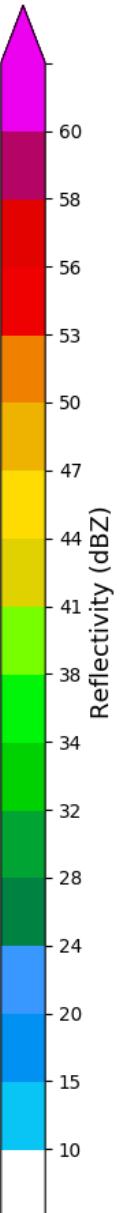
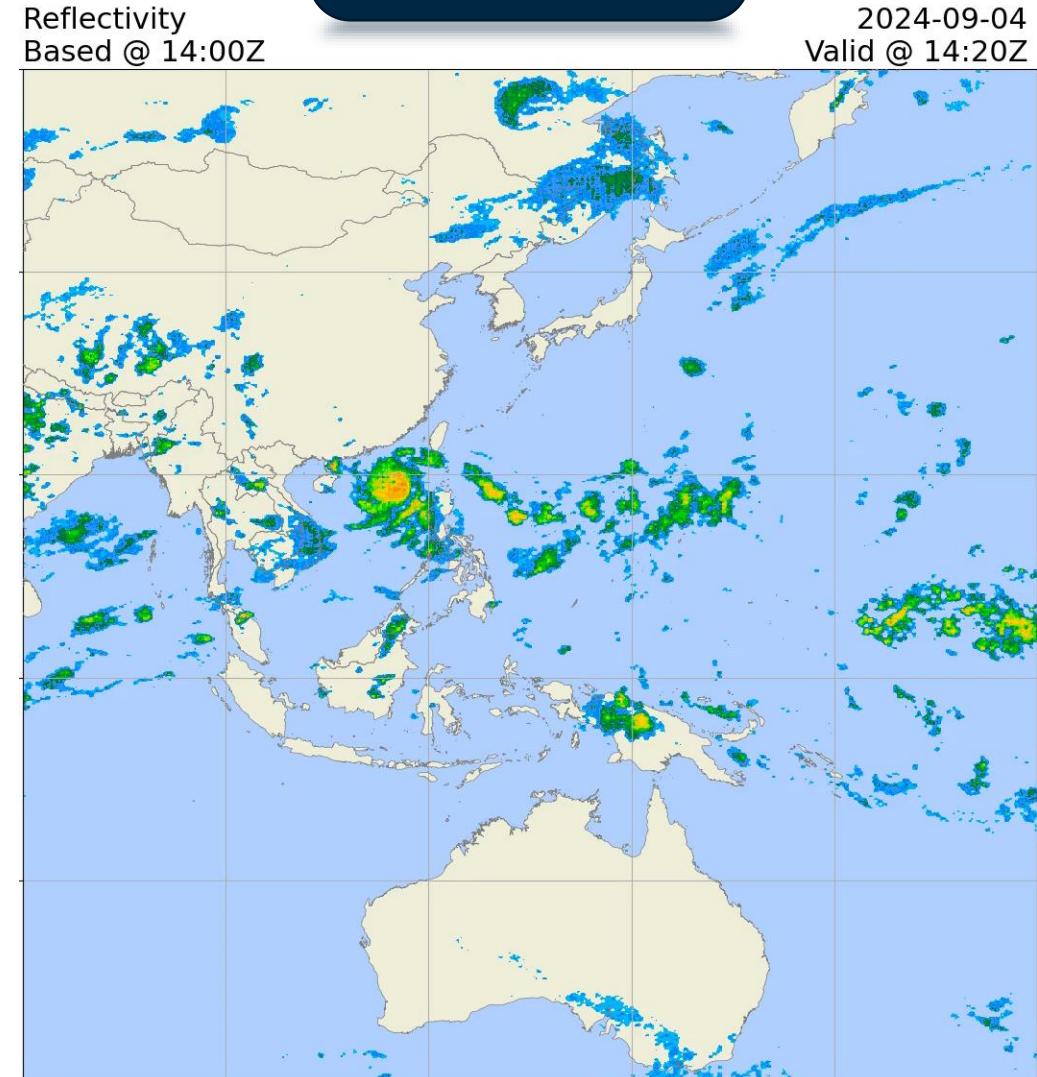
Deep Learning Models in Regional Nowcasting Supports

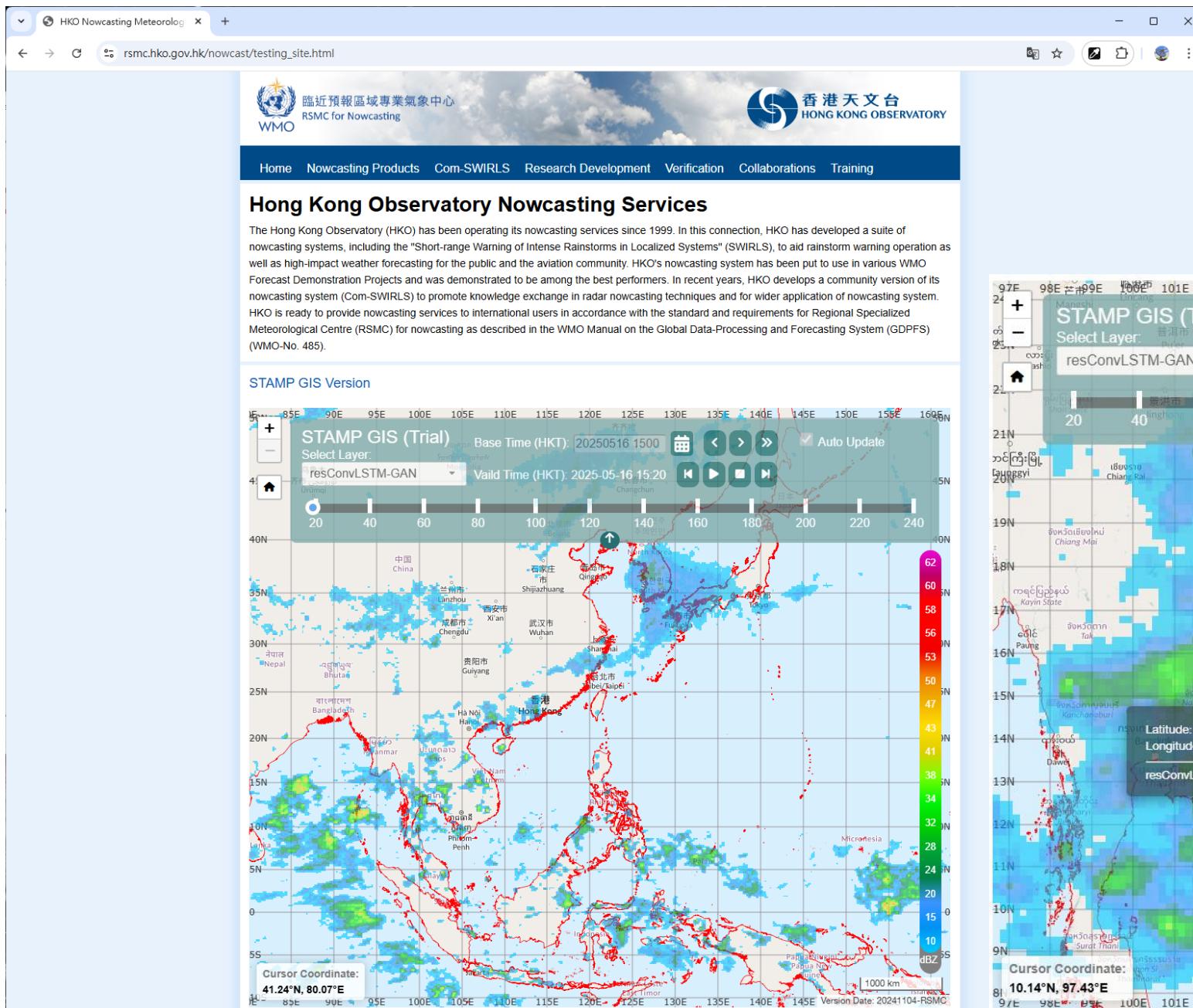
Use of AI/ML in nowcasting – Precipitation and Significant Convection

Actual Observations

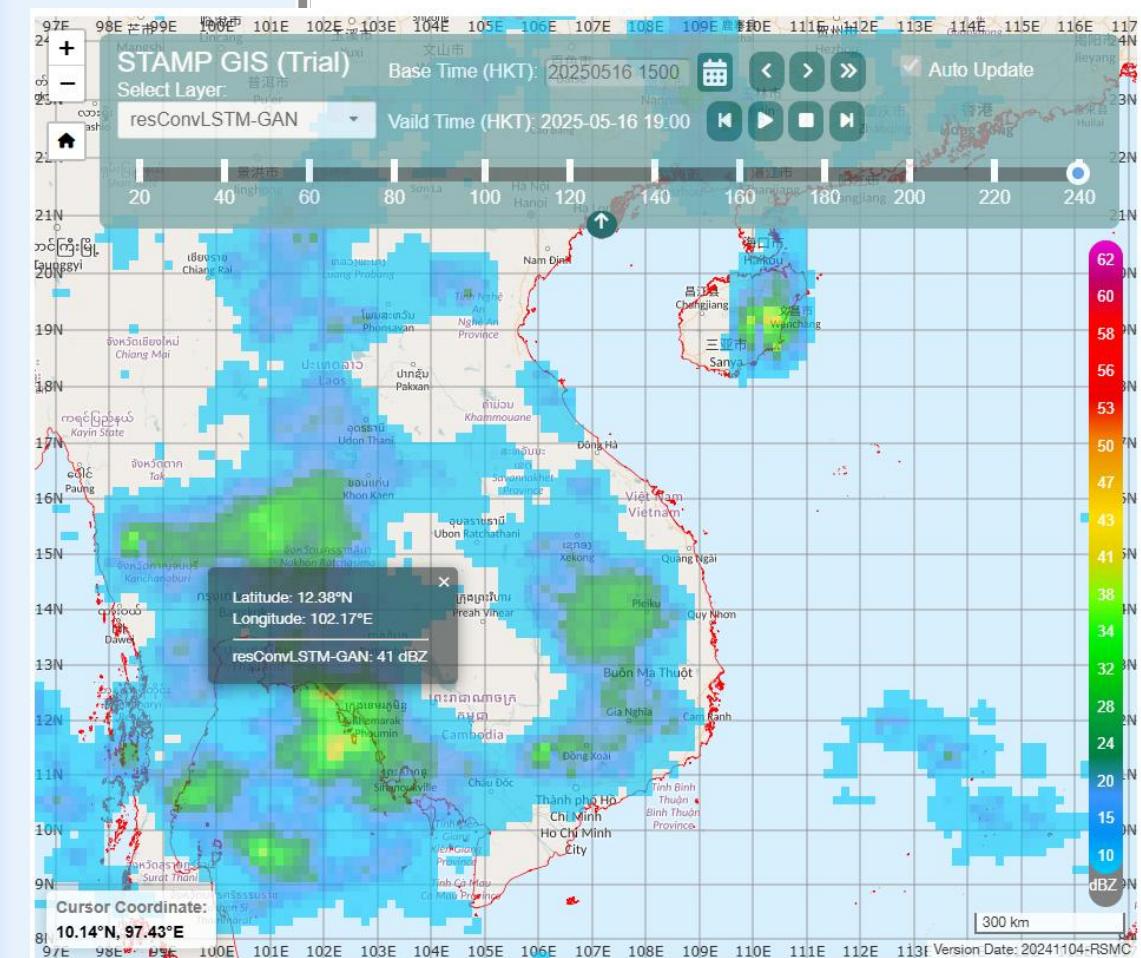


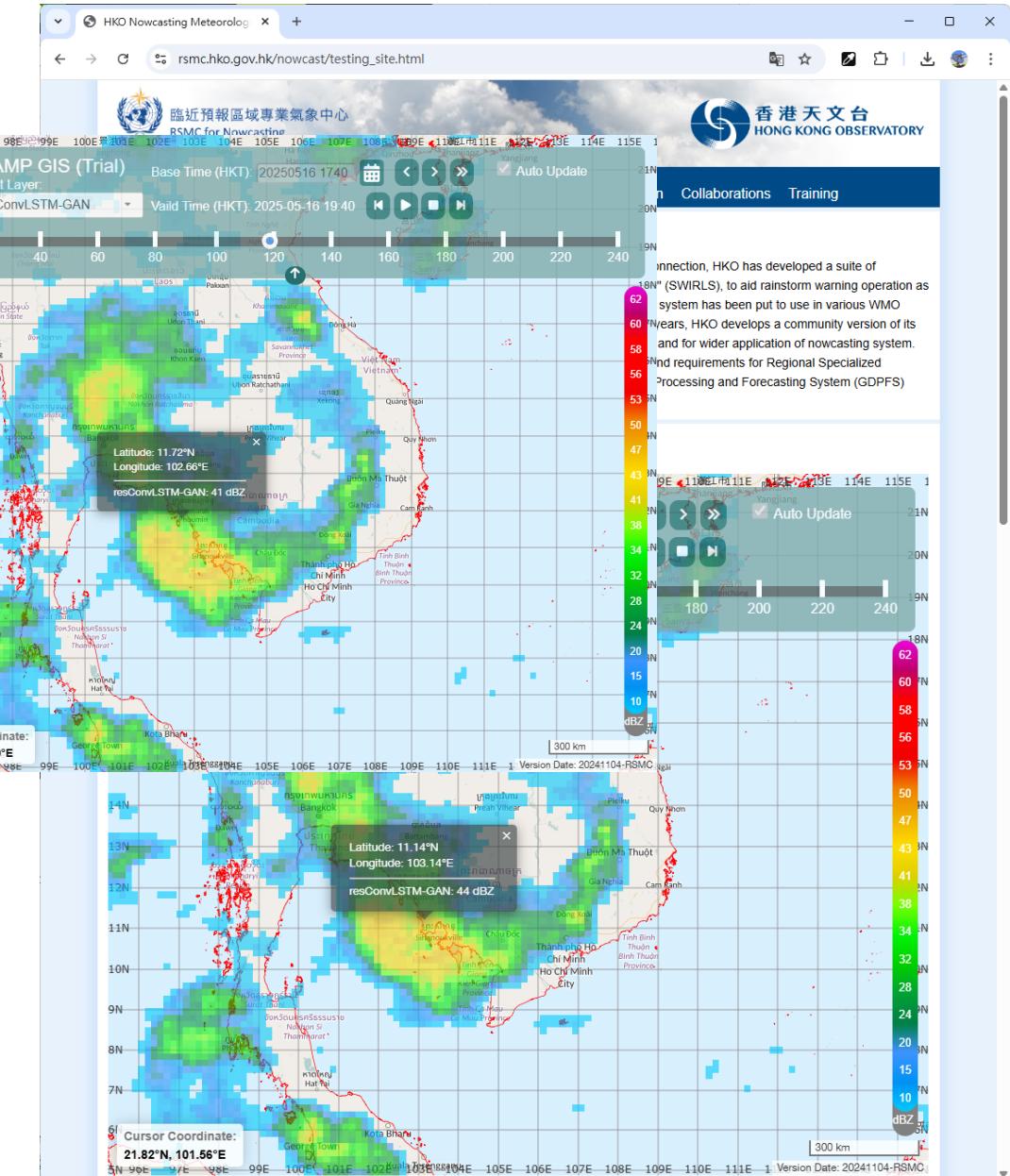
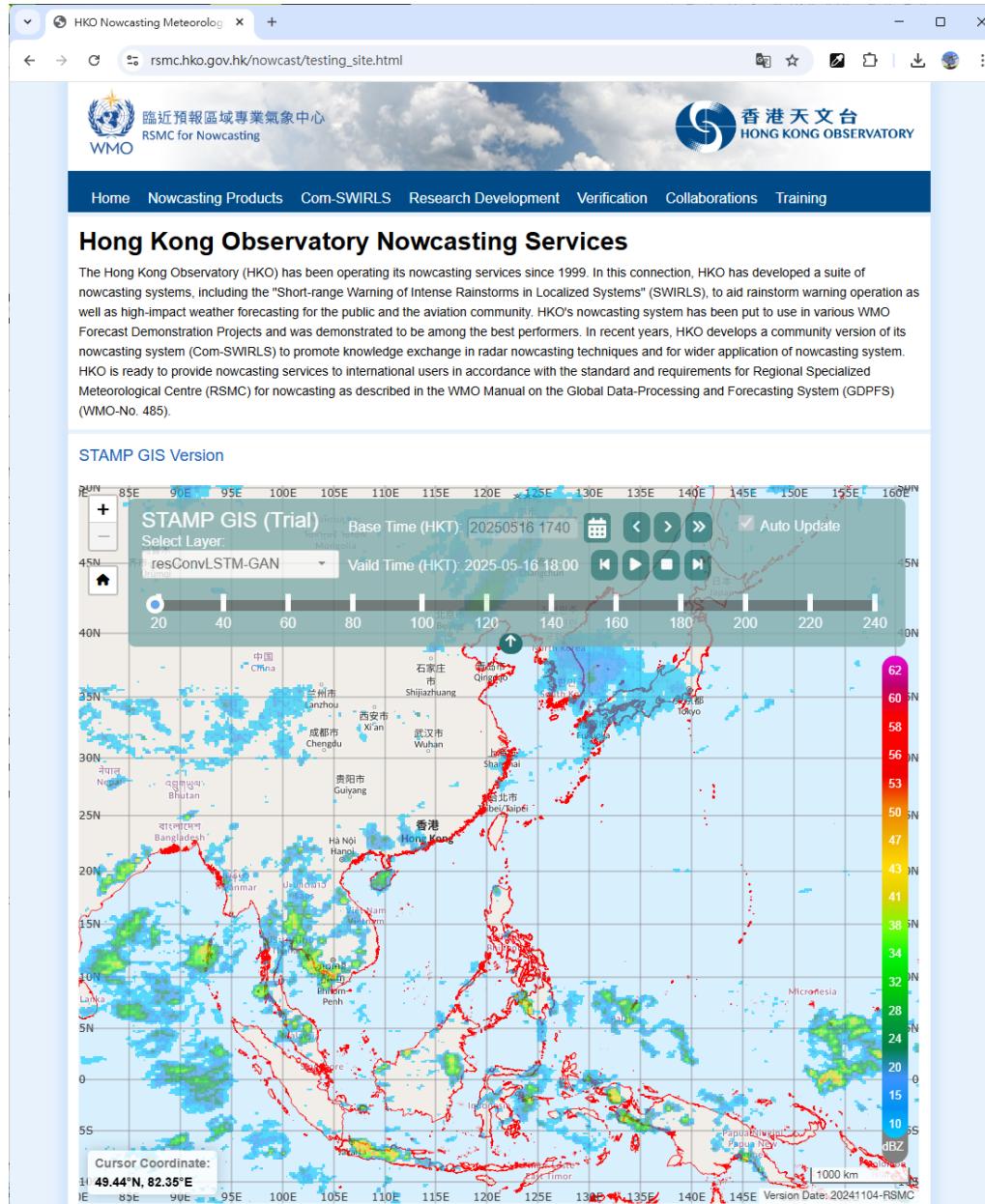
ResConvLSTM-GAN



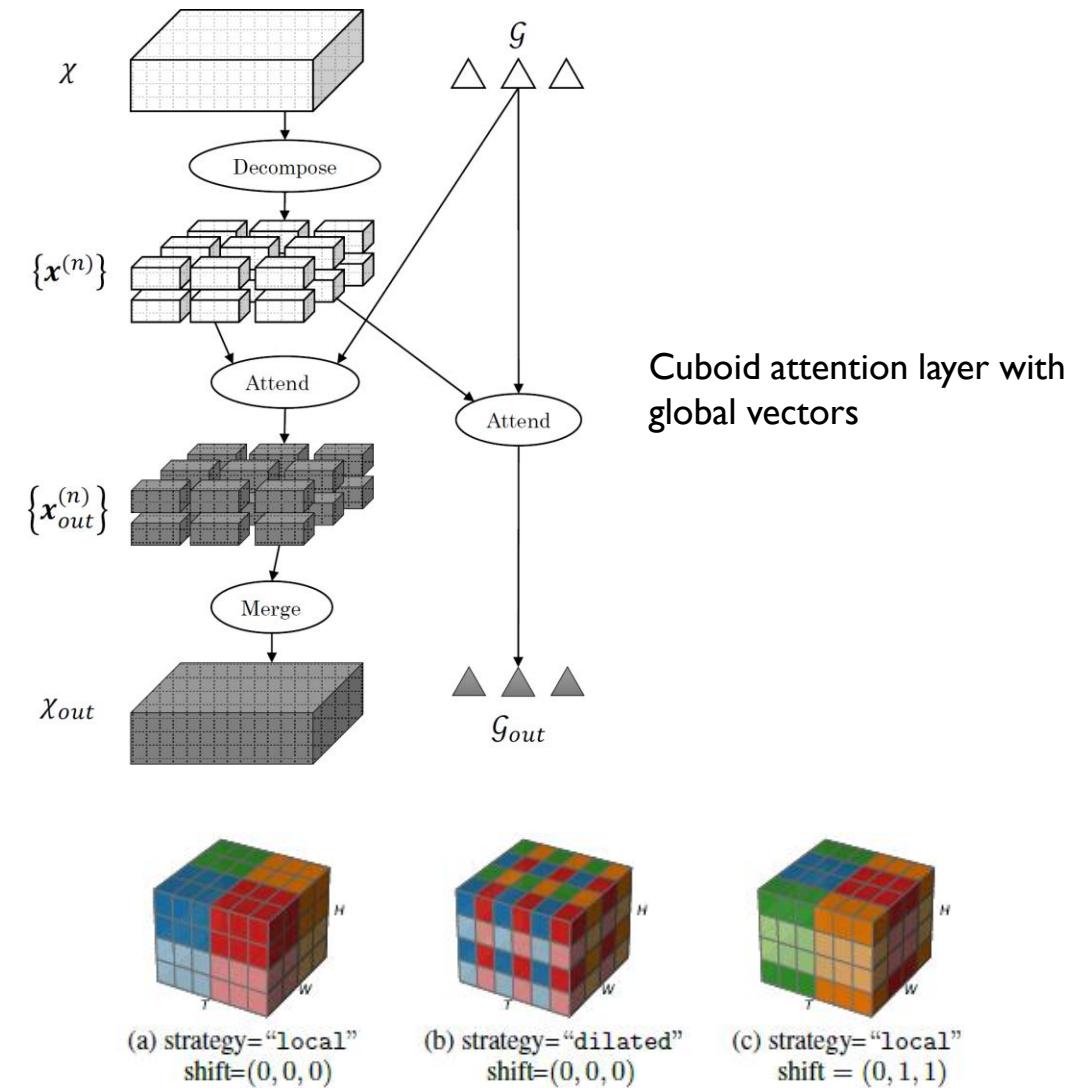
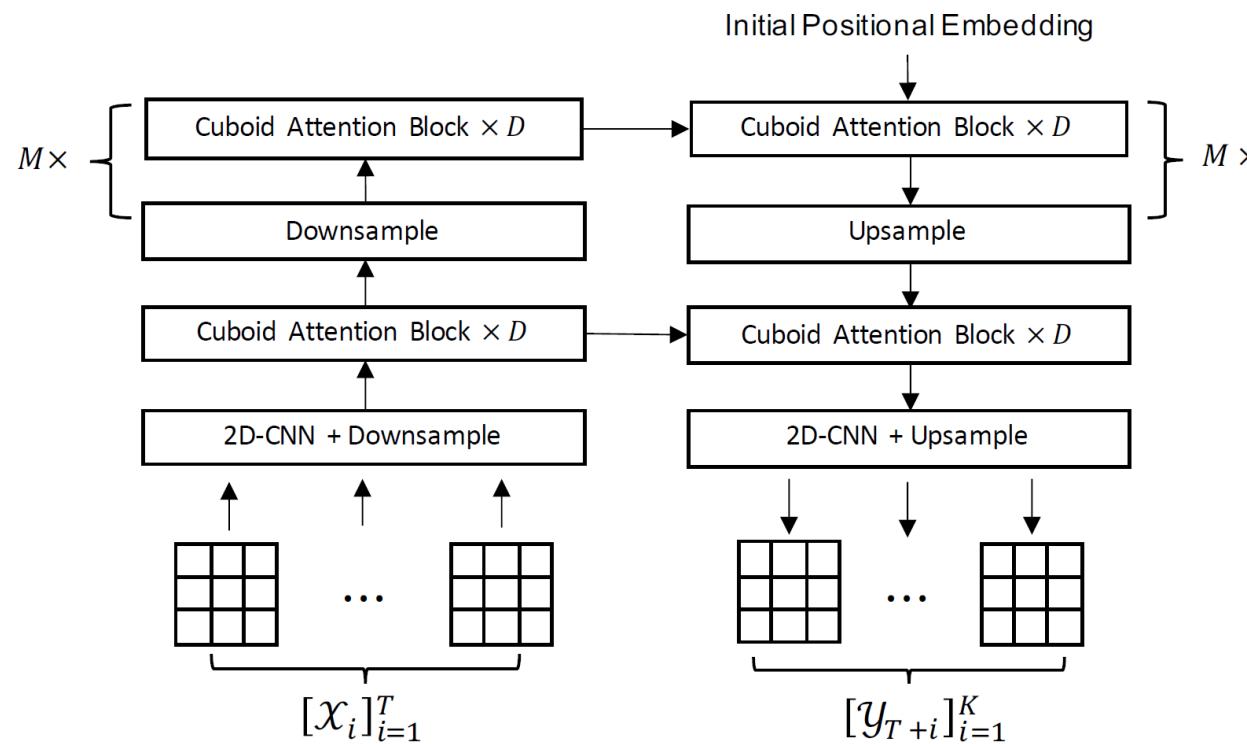


Web-GIS tool showing AI/ML nowcast product and interactive visualization
(upcoming version of RSMC for Nowcasting website)



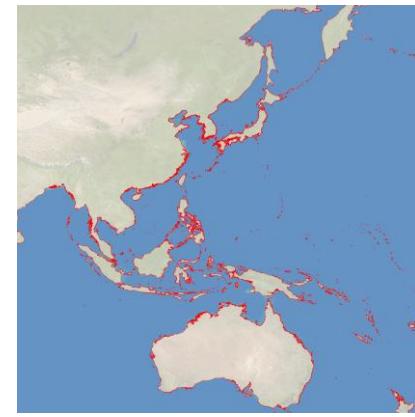


Earthformer

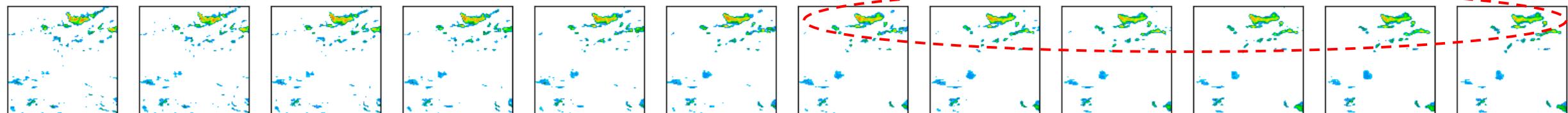
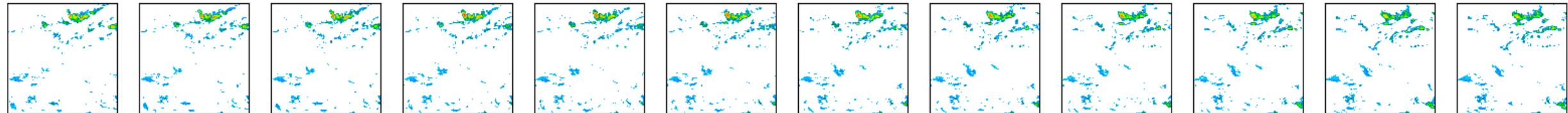
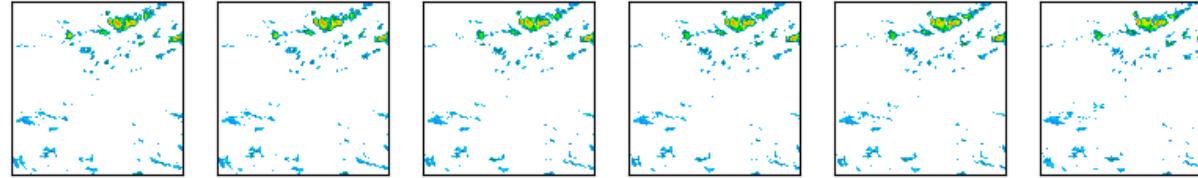


- **Space-time Transformer** model based on Cuboid Attention
- Spatiotemporal data are divided into non-overlapping cuboids, and **self-attention** is applied locally within each cuboid
- **Global vectors** are used to connect the cuboids for capturing the overall pattern

Extending satellite retrieved reflectivity nowcast using Earthformer



Input
Observation
Prediction



T+20m T+60m T+100m

.....

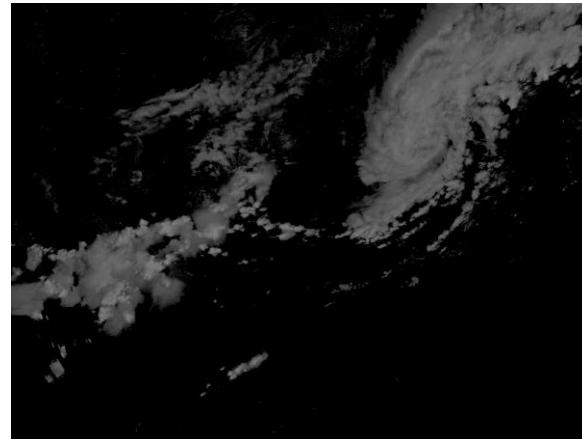
T+420m T+460m

Preserving intensity in T+ 4 to T+8 hours

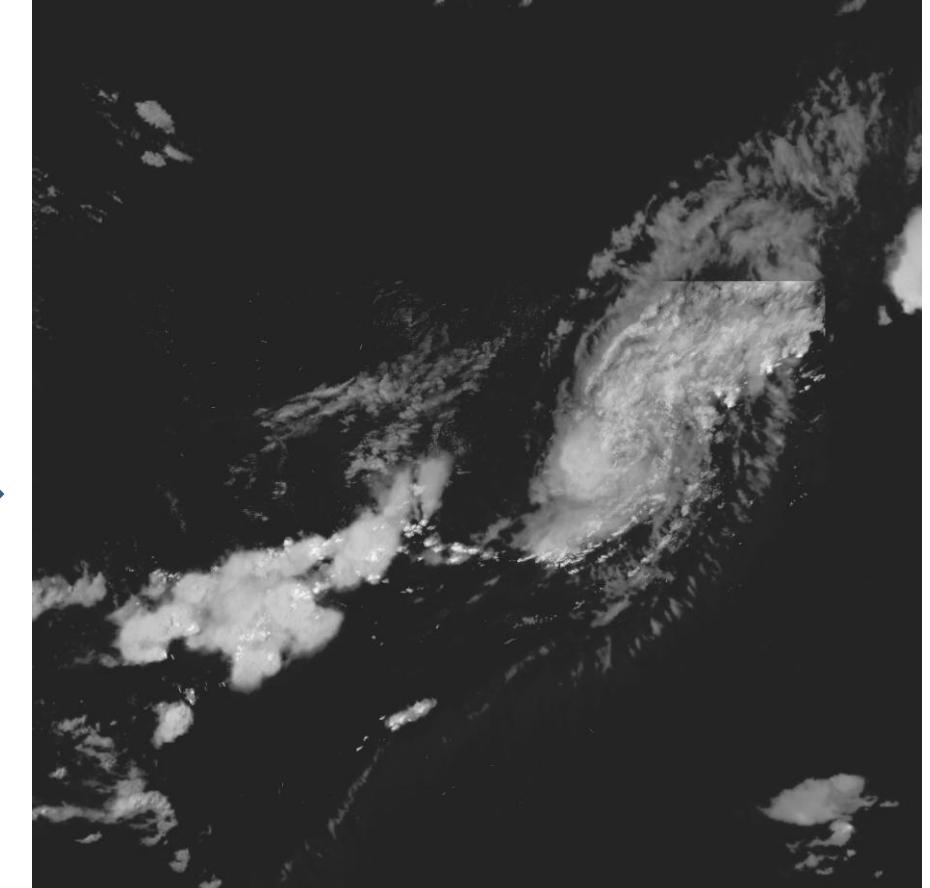
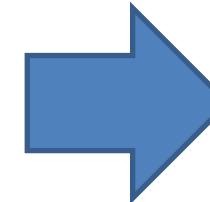
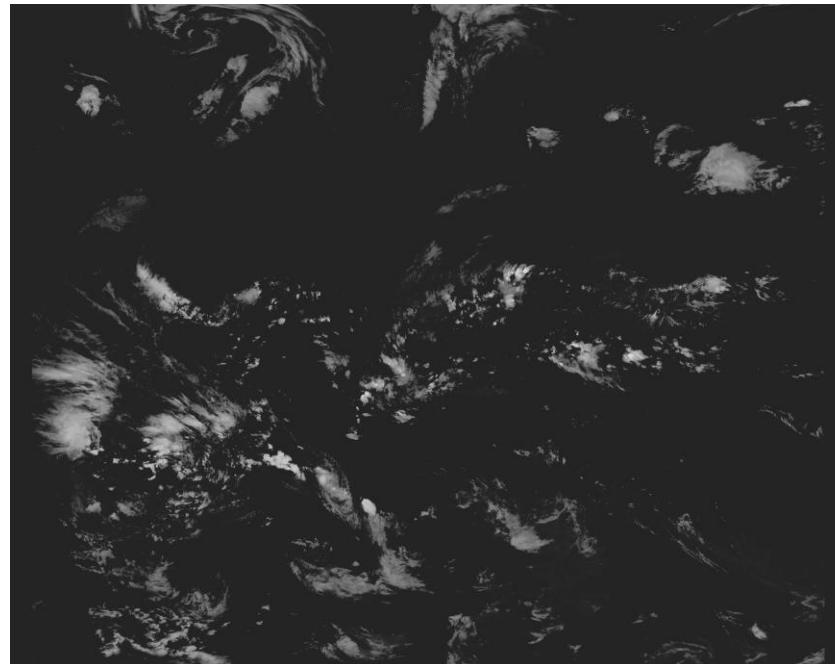
Earthformer for CMA radar composite with H9/GK2A satellite simulated reflectivity in AINPP intercomparison project

CMA radar composite for
AINPP intercomparison
experiments

18-30N 104-120E



Simulated
reflectivity using
H9/GK2A
 imagers
 composite



Mosaic reflectivity imagery
16.00-36.48N, 101.76-122.24E @ 0.01 deg
(2048 x 2048 pixels)

Satellite-Radar Composite Reflectivity

Date range

Training / Validation set:
Mar 2023 to Oct 2024

Domain coverage

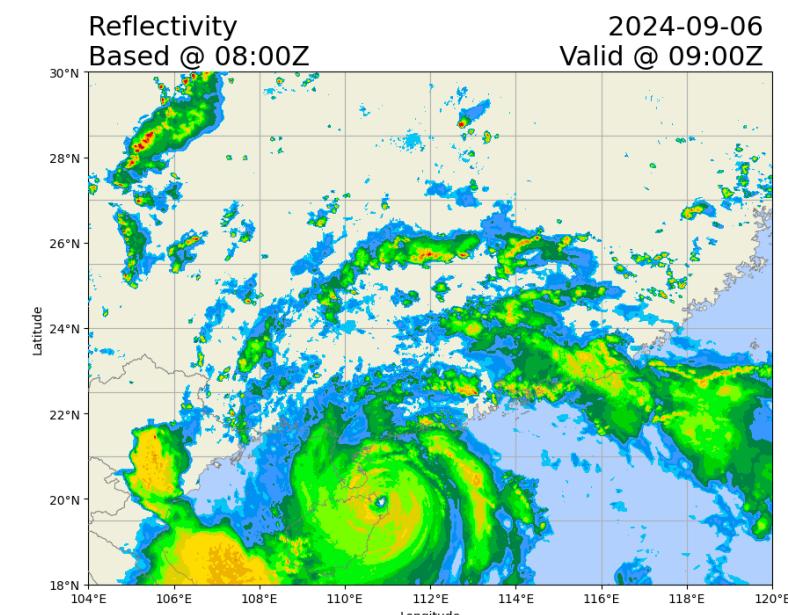
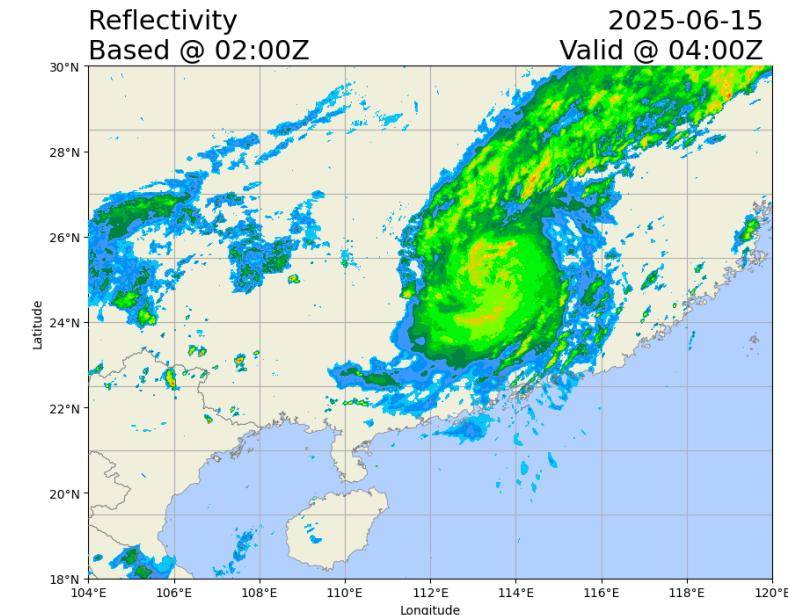
2048 x 2048 grid from
16 to 36.48N; 101.76 to 122.24E
(0.01x 0.01 deg)

Input data

6 maps with 20 minutes interval
(2 hours of observations)

Output data

24 maps with 20 minutes interval
(8 hours regional reflectivity nowcast)



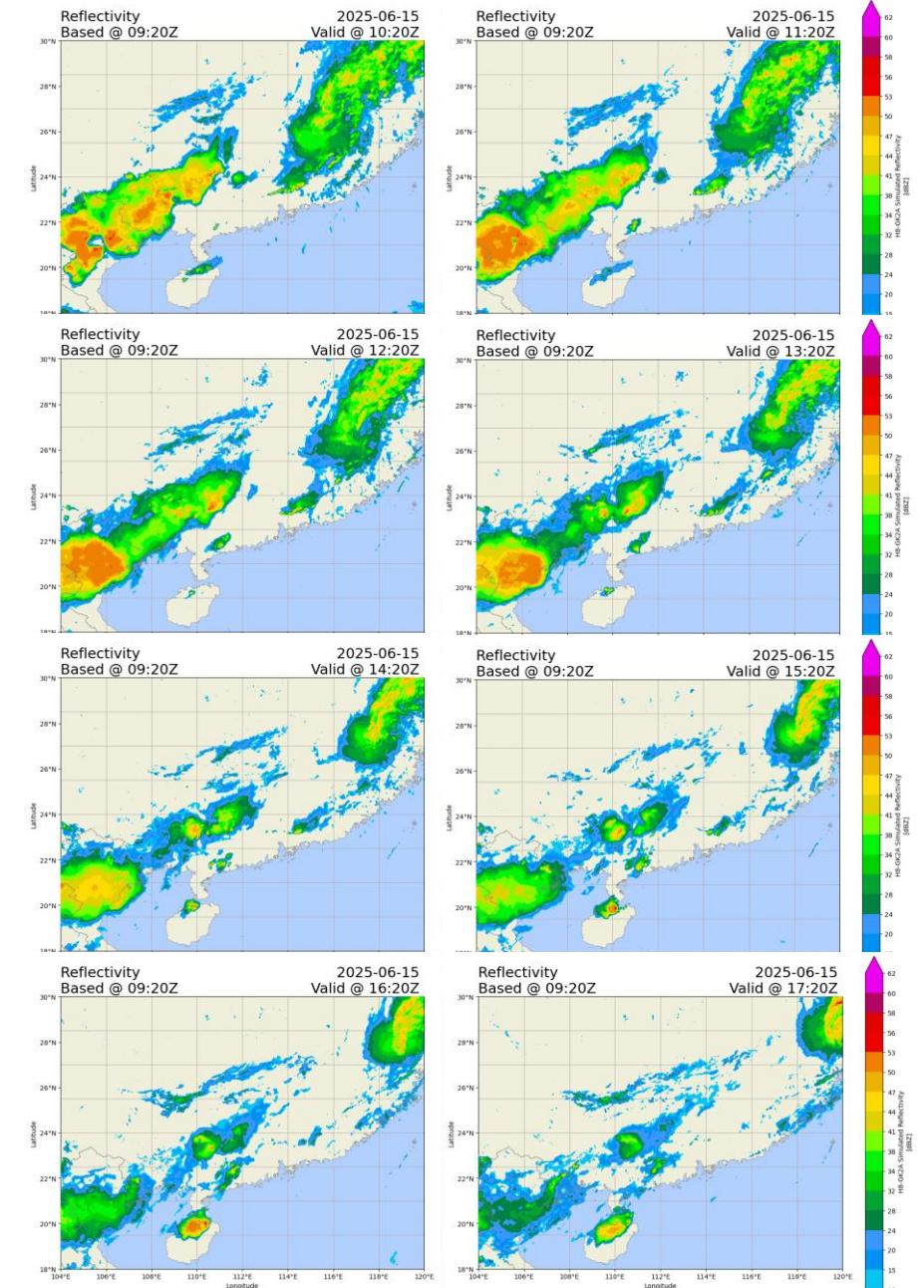
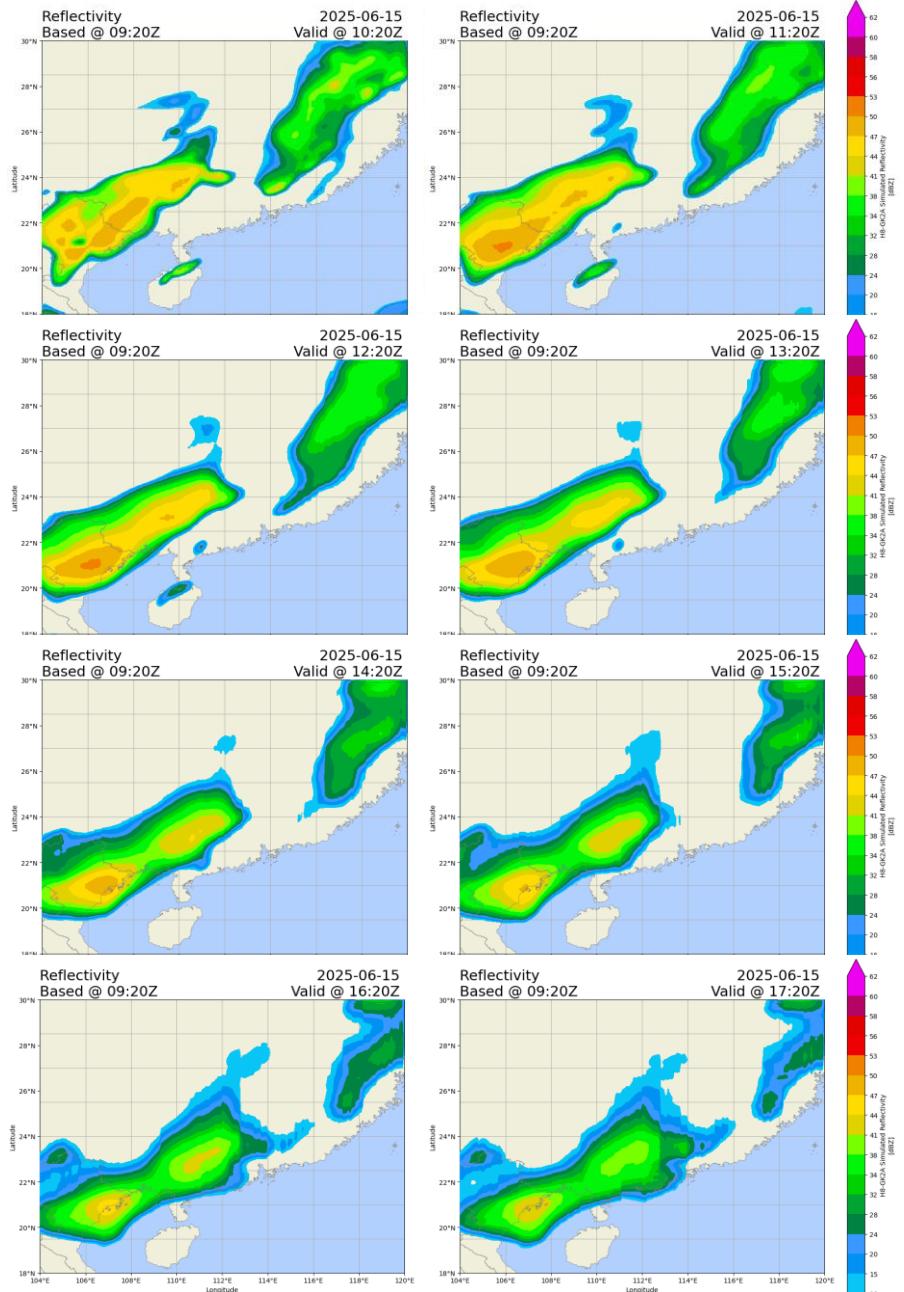
Earthformer T+1 to T+8h Nowcasts

2025-06-15 09:20Z



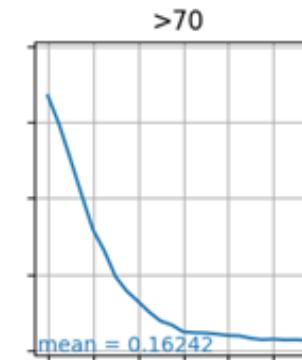
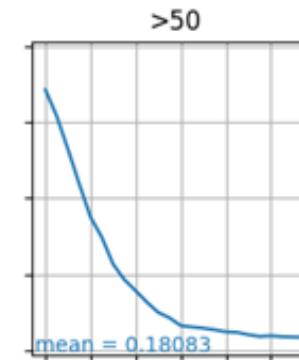
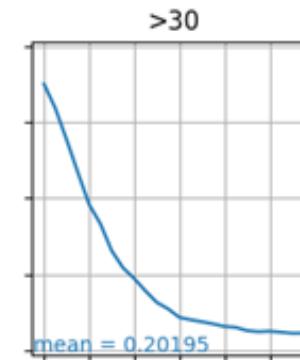
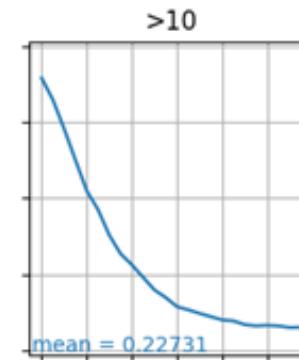
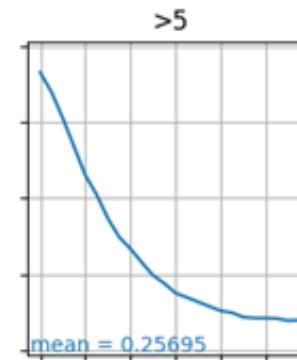
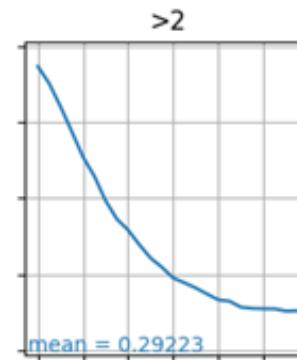
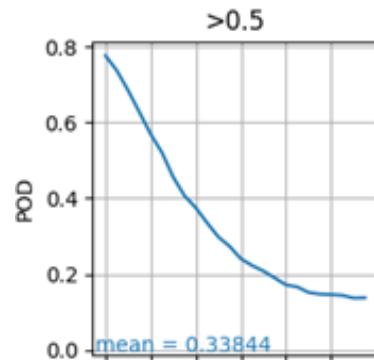
臨近預報區域專業氣象中心
RSMC for Nowcasting

Actual reflectivity composite

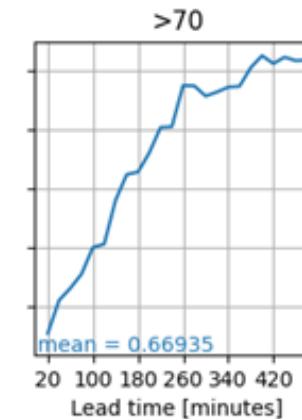
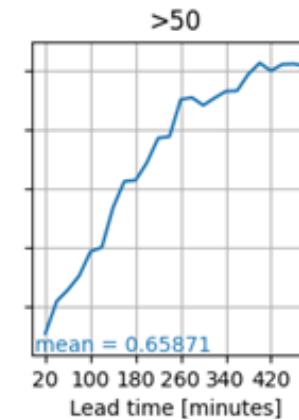
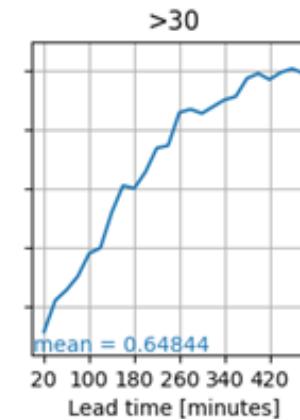
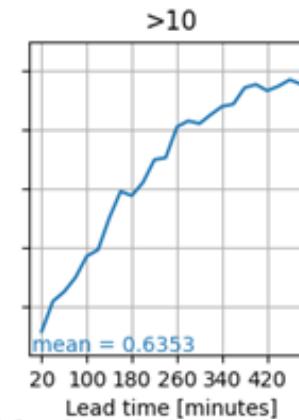
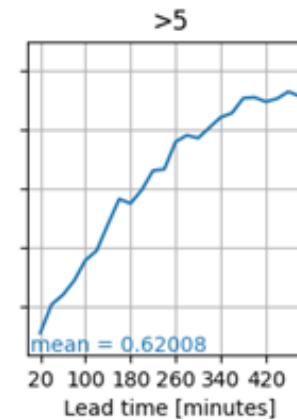
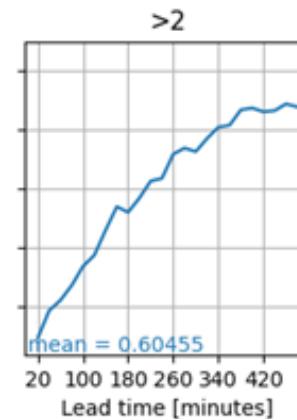
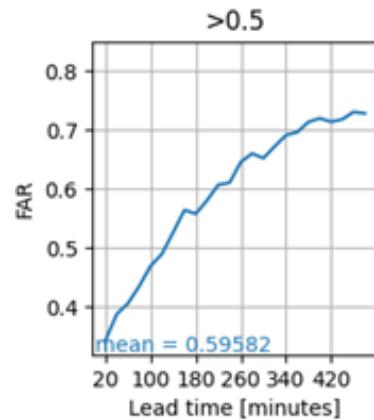


Verification – POD, FAR and HSS

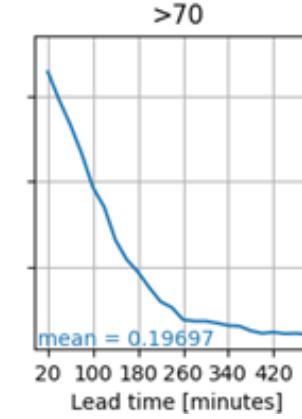
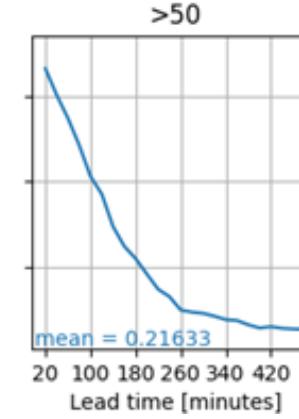
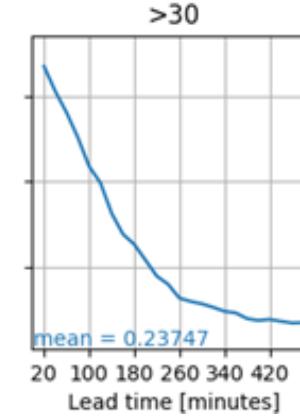
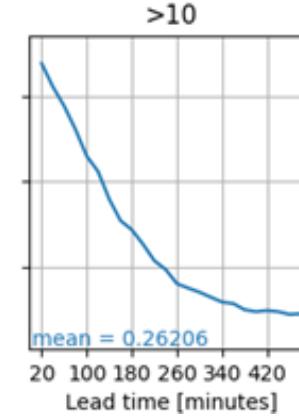
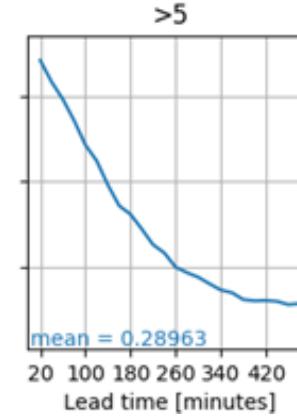
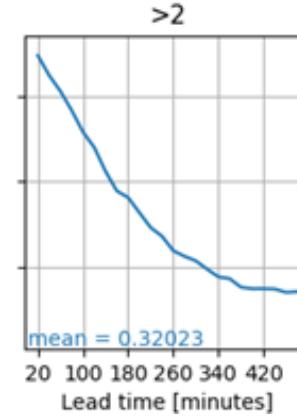
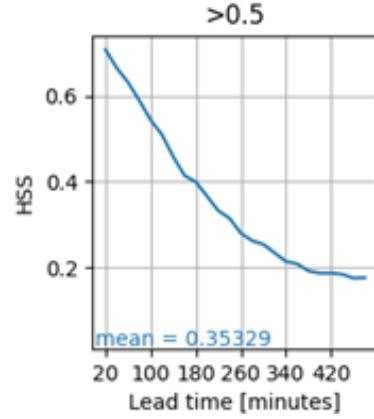
POD - Lead time plot



FAR - Lead time plot



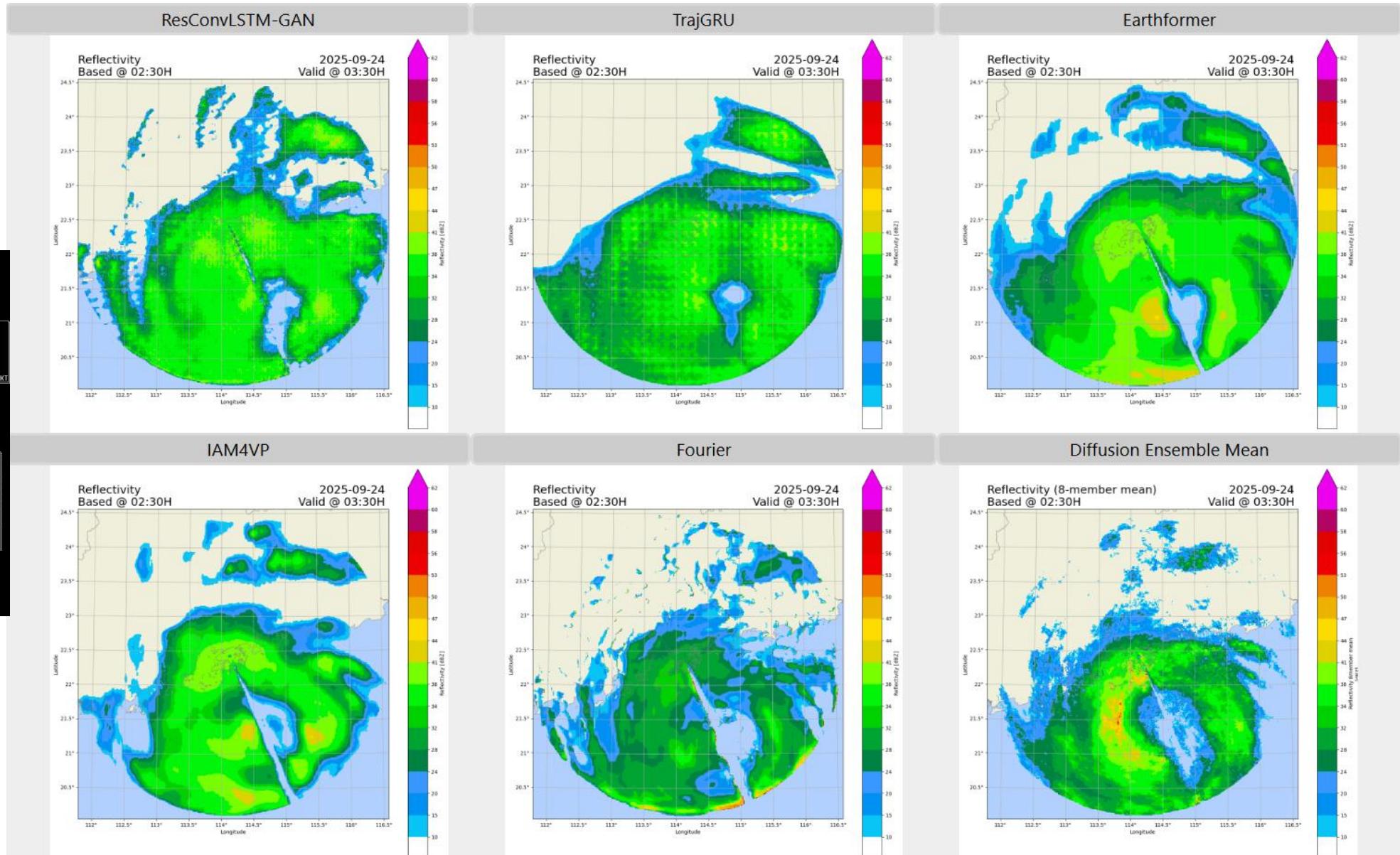
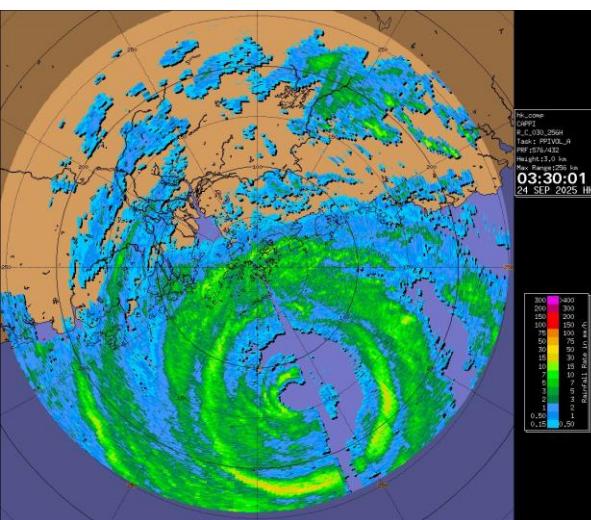
HSS - Lead time plot



Ensemble Nowcast and AI-based Guidance in Extreme Precipitation Prediction

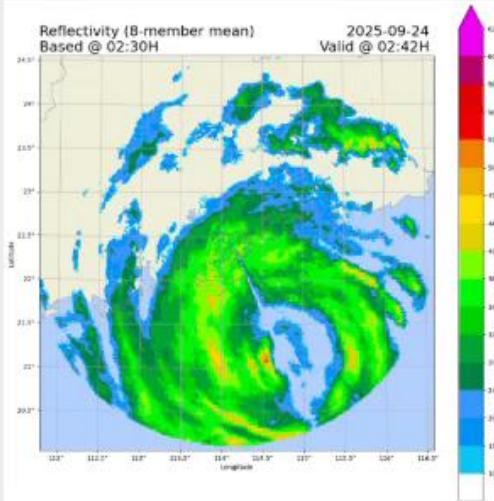
Super Typhoon Ragasa (2025-09-24)

T+1h nowcast from
6 deep learning
algorithms

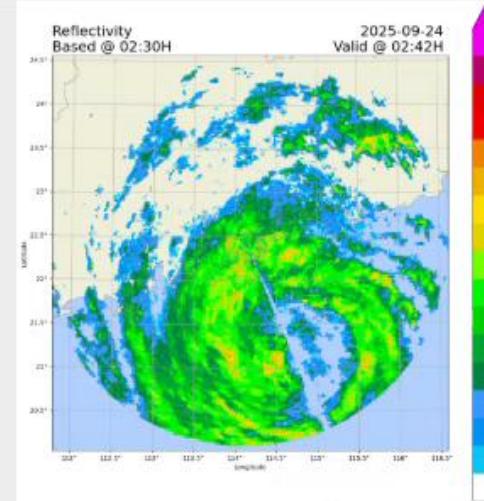


Ensemble Mean and 8 Diffusion Ensemble Members

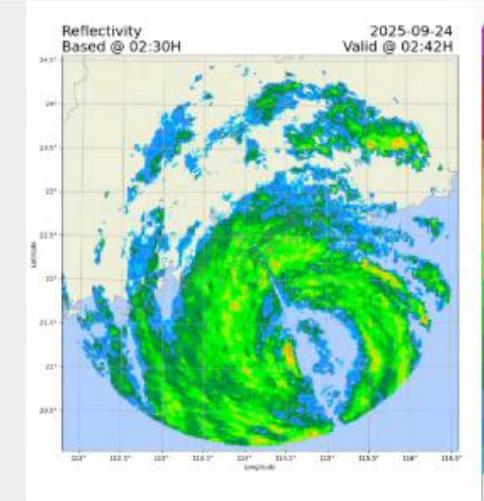
Ensemble Mean



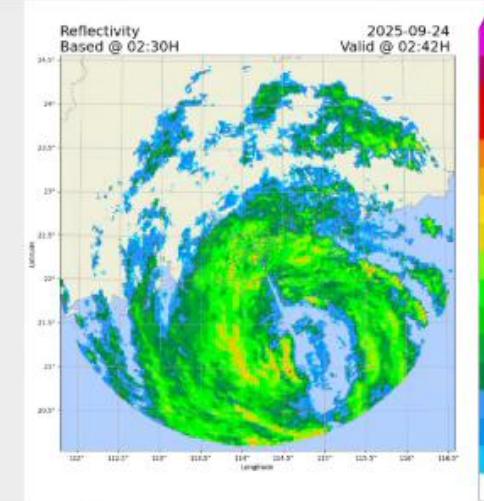
Member 1



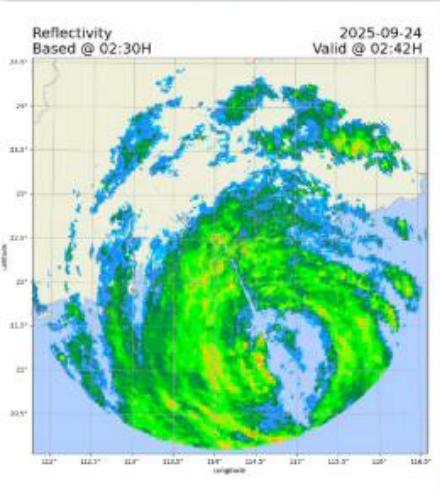
Member 2



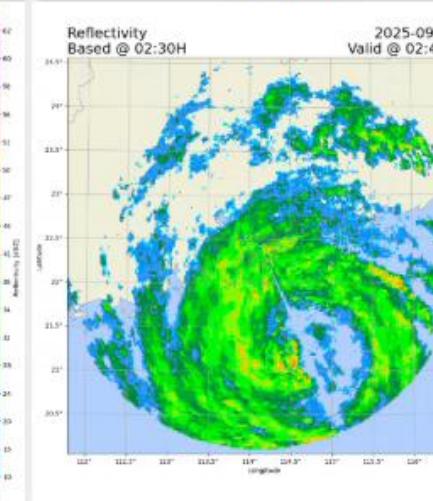
Member 3



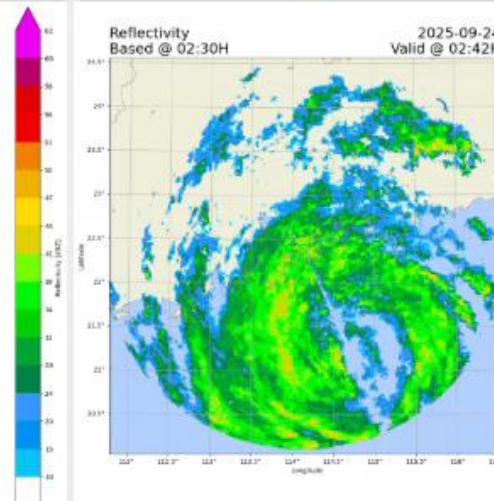
Member 4



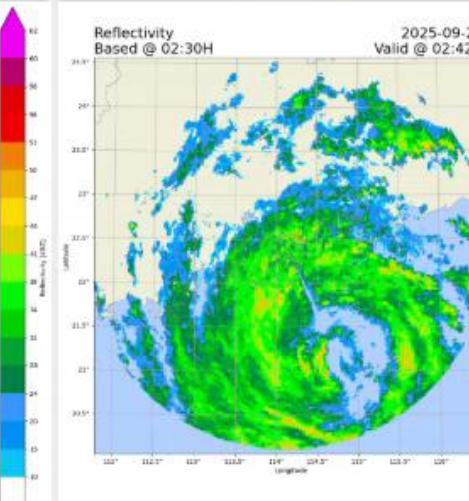
Member 5



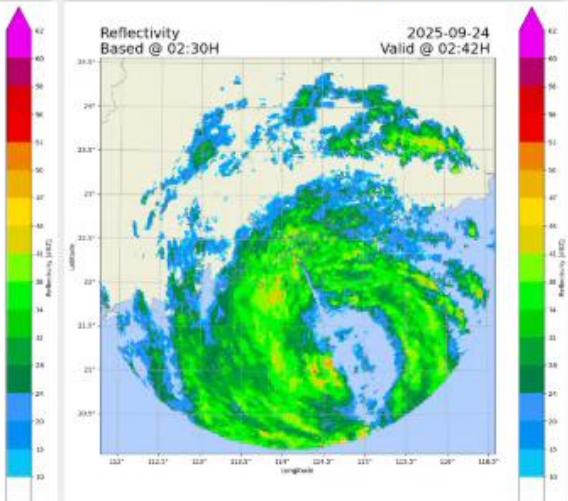
Member 6



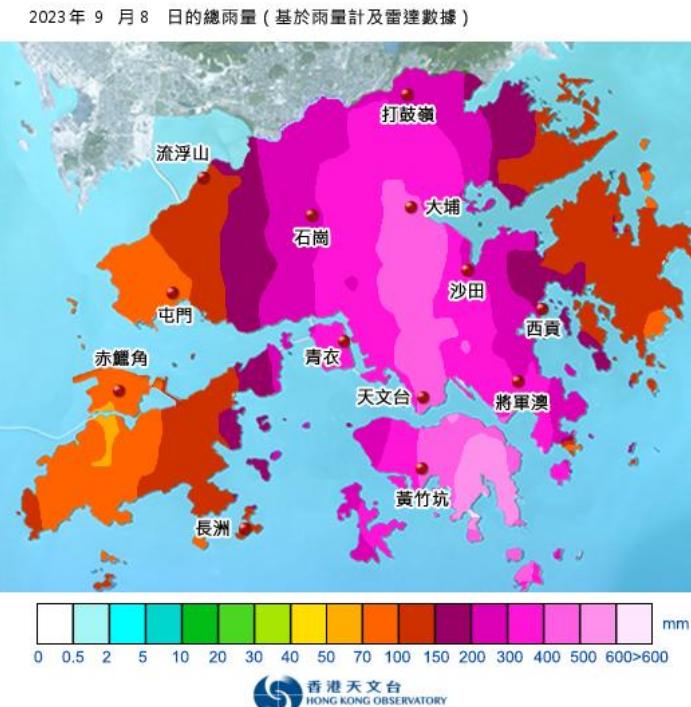
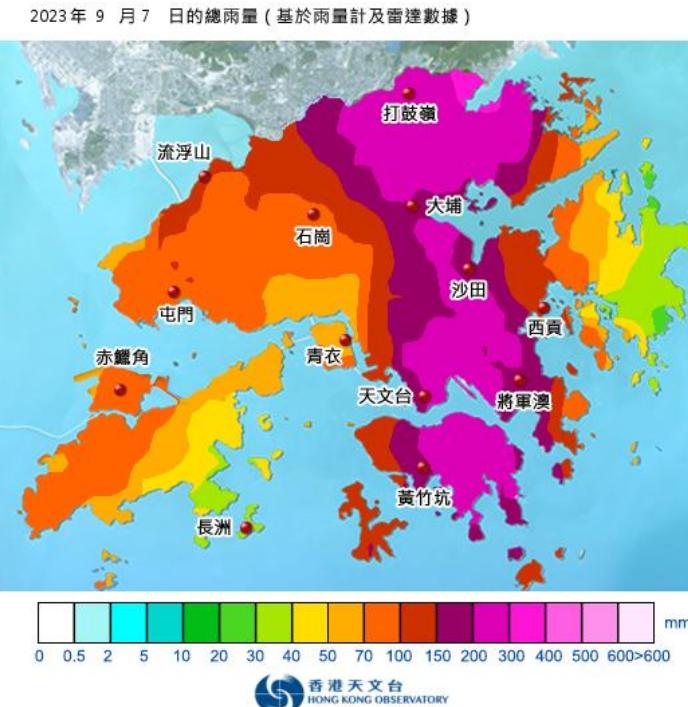
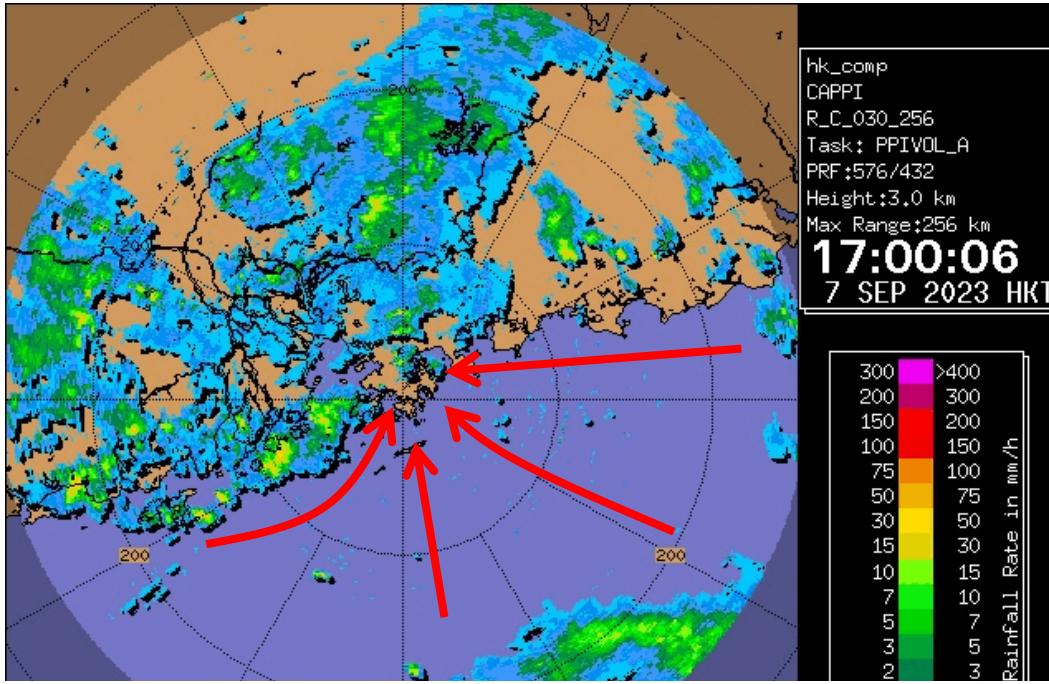
Member 7



Member 8



Exceptionally severe rainstorm on 7-8 September 2023



100 Highest 1-Hour Total Rainfall (mm) at the Hong Kong Observatory for All Months

No	Rank	Total Rainfall	Date	Ending Hour (HKT)
1	1	158.1	2023.09.07	24
2	2	145.5	2008.06.07	09
3	3	115.1	2006.07.16	03
4	4	109.9	1992.05.08	07
5	5	108.2	1966.06.12	07
6	6	105.9	2023.09.08	08
7	7	104.8	1989.05.02	13
8	8	100.7	1926.07.19	04
9	9	100.0	1968.06.13	03
10	10	98.7	1972.06.18	12

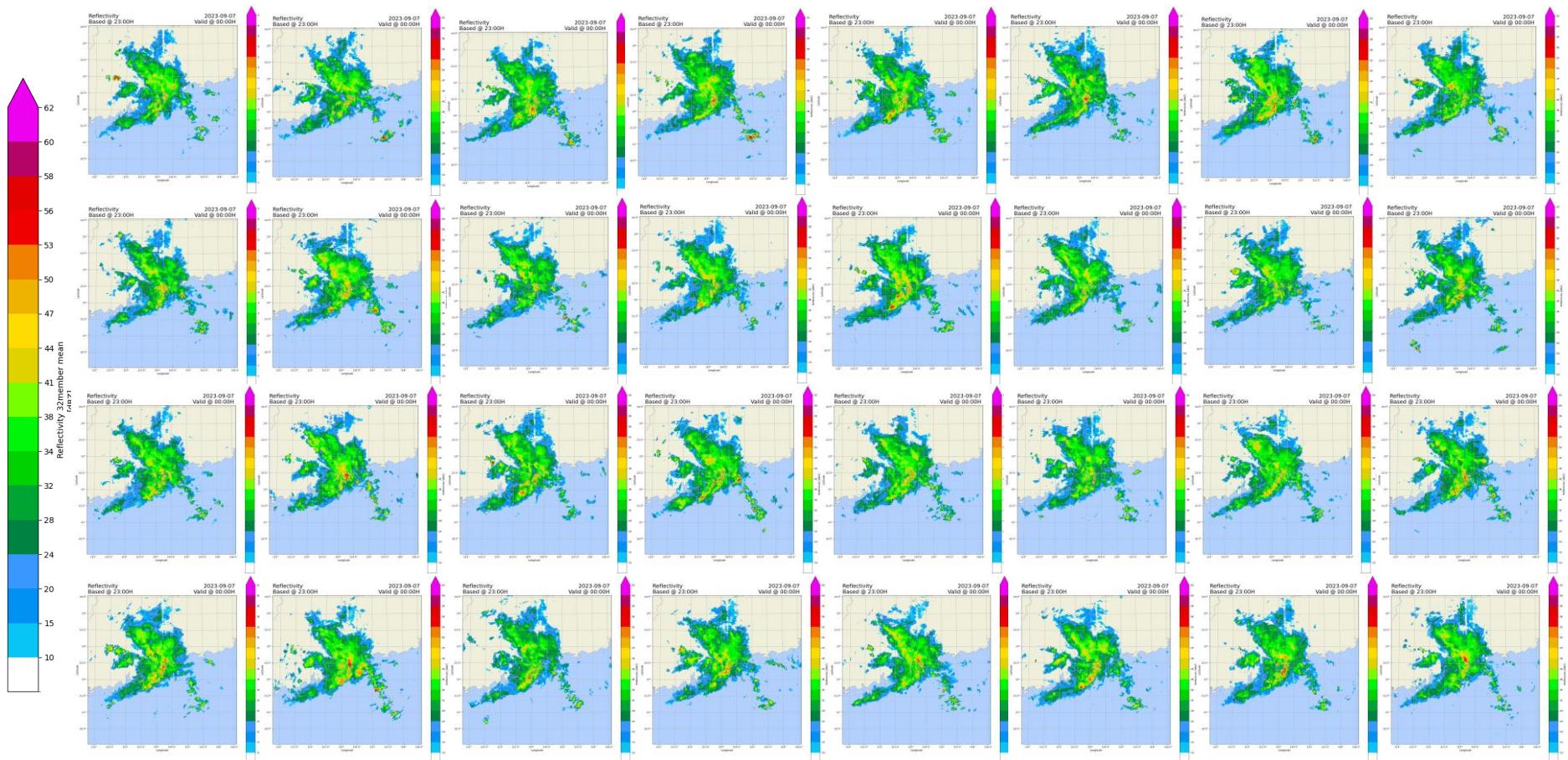
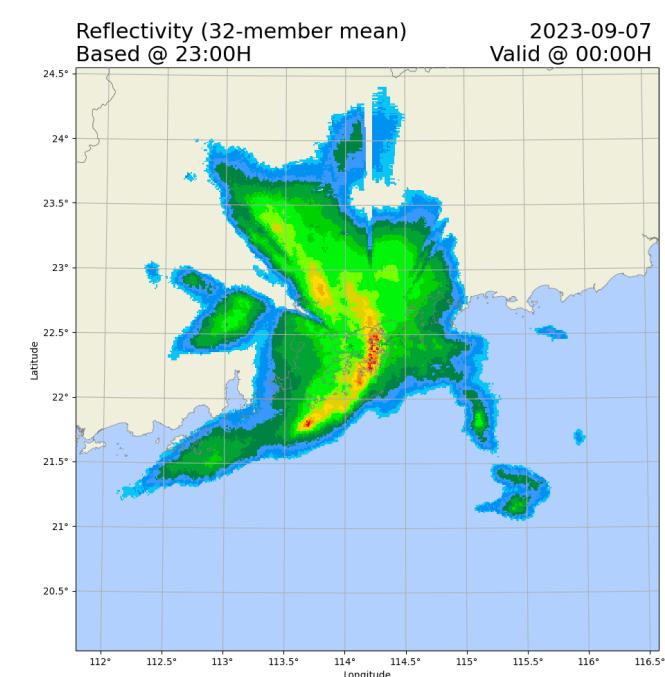
Highest Daily Total Rainfall (mm) at Hong Kong Observatory for All Months

Record Since: 1884.03.01, exclude 1940-1946

No	Rank	Total	Date of record
1	1	534.1	1926.07.19
2	2	520.6	1889.05.30
3	3	425	2023.09.08
4	4	411.3	1998.06.09
5	5	382.6	1966.06.12
6	6	346.7	1983.06.17
7	7	342.3	1886.07.15
8	8	334.2	1982.08.16
9	9	329.7	2021.10.08
10	10	325.5	1965.09.27

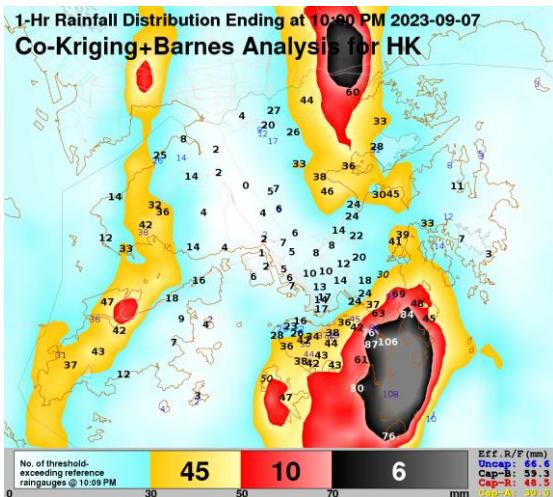
32-member DDPM ensemble

→ Dynamic z-R relationship based on CAPPI reflectivity and raingauges in HK ($z = a R^b$) to generate hourly rainfall

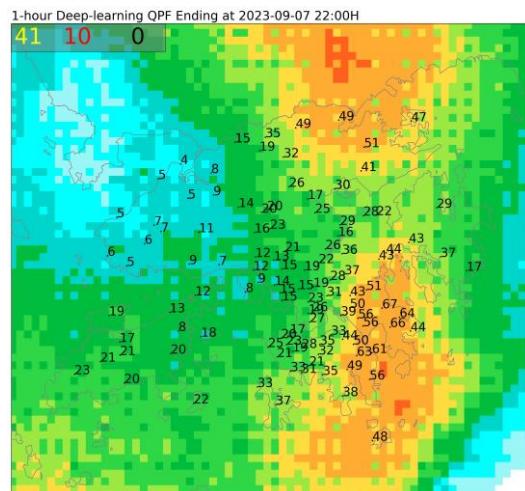


1hr QPF Base time: 2100H

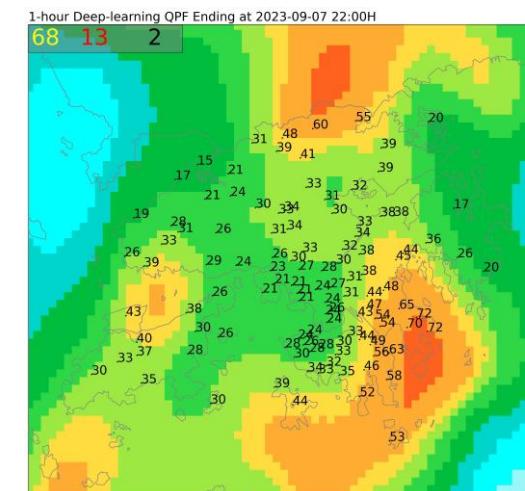
QPE



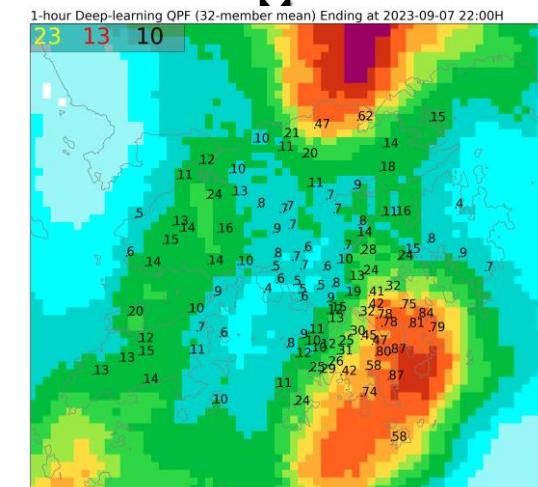
ResConvLSTM-GAN



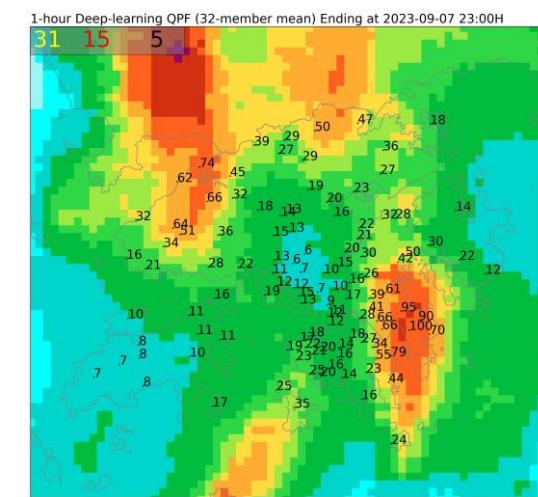
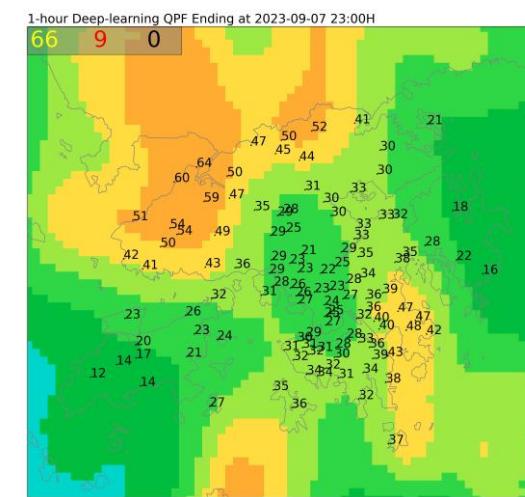
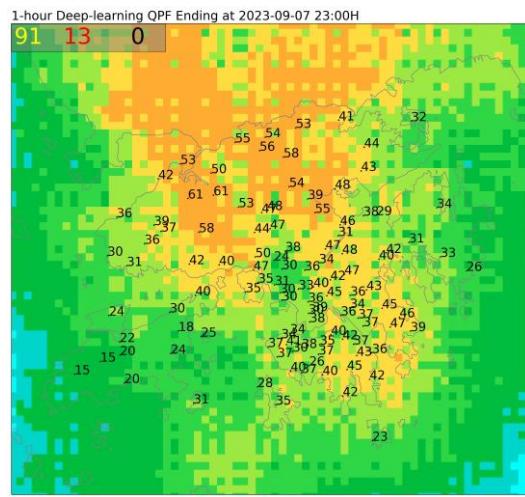
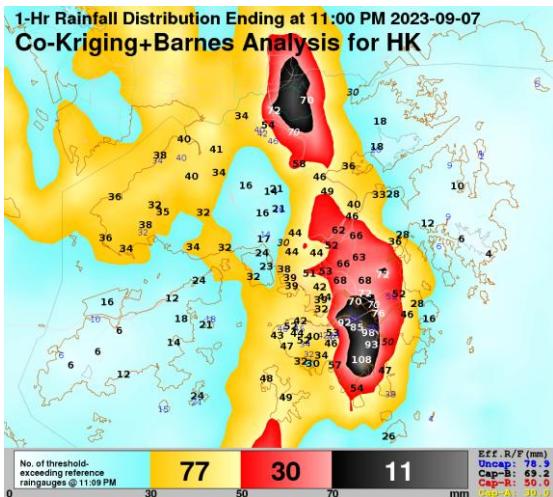
Earthformer



Diffusion Ensemble

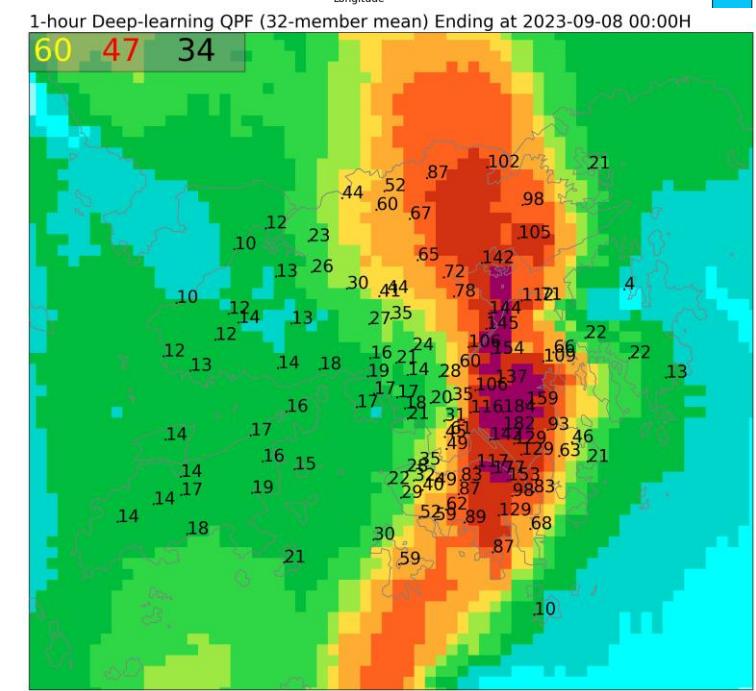
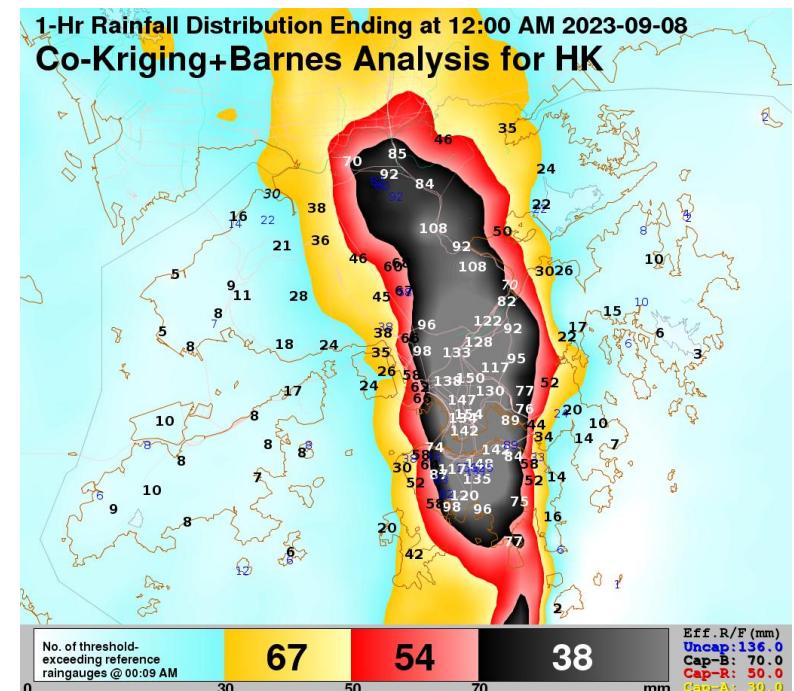
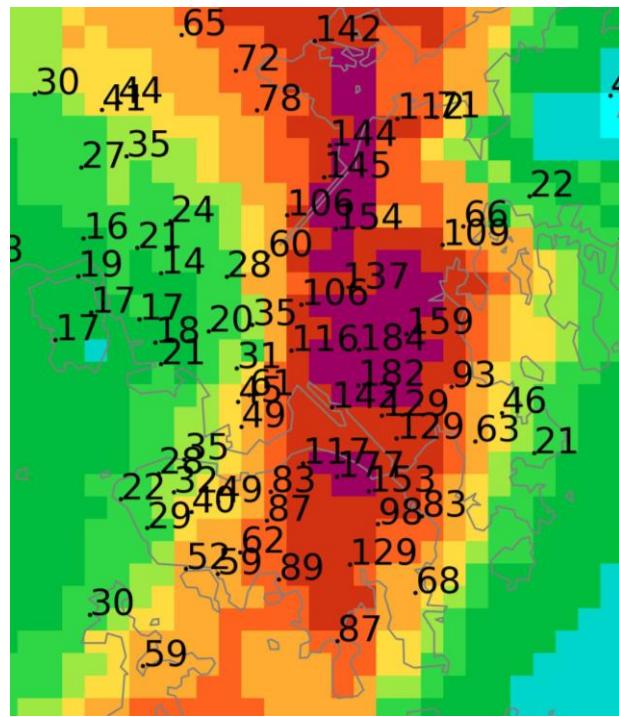
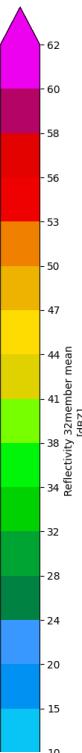
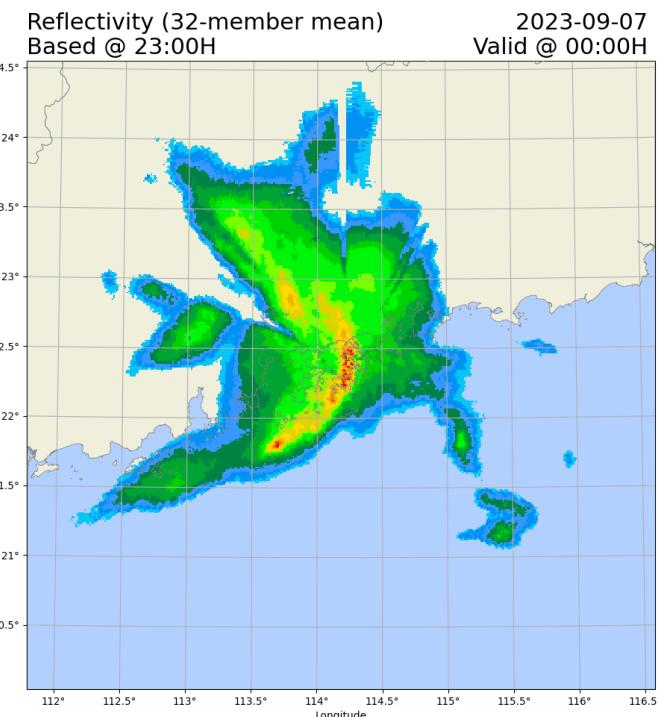
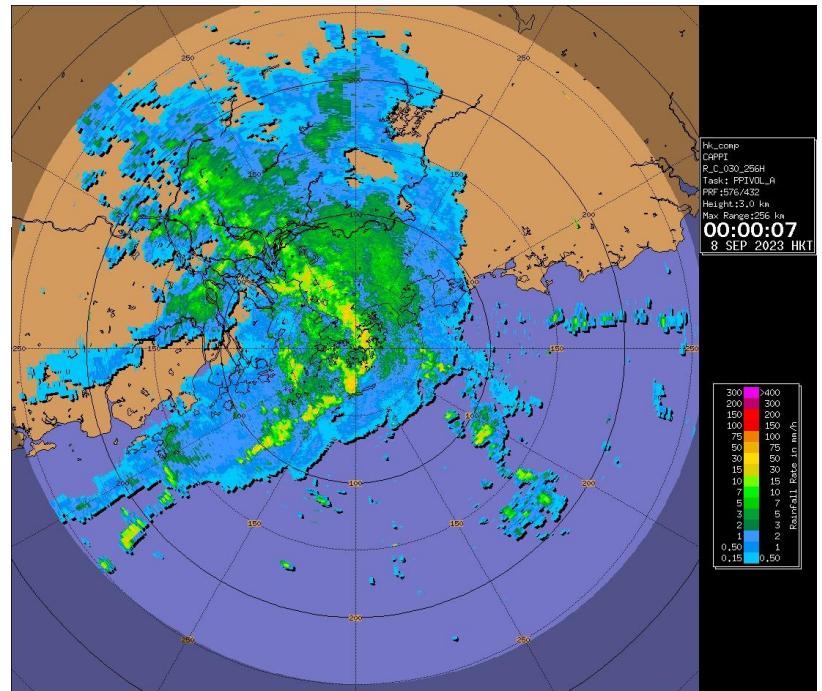


1hr QPF Base time: 2200H



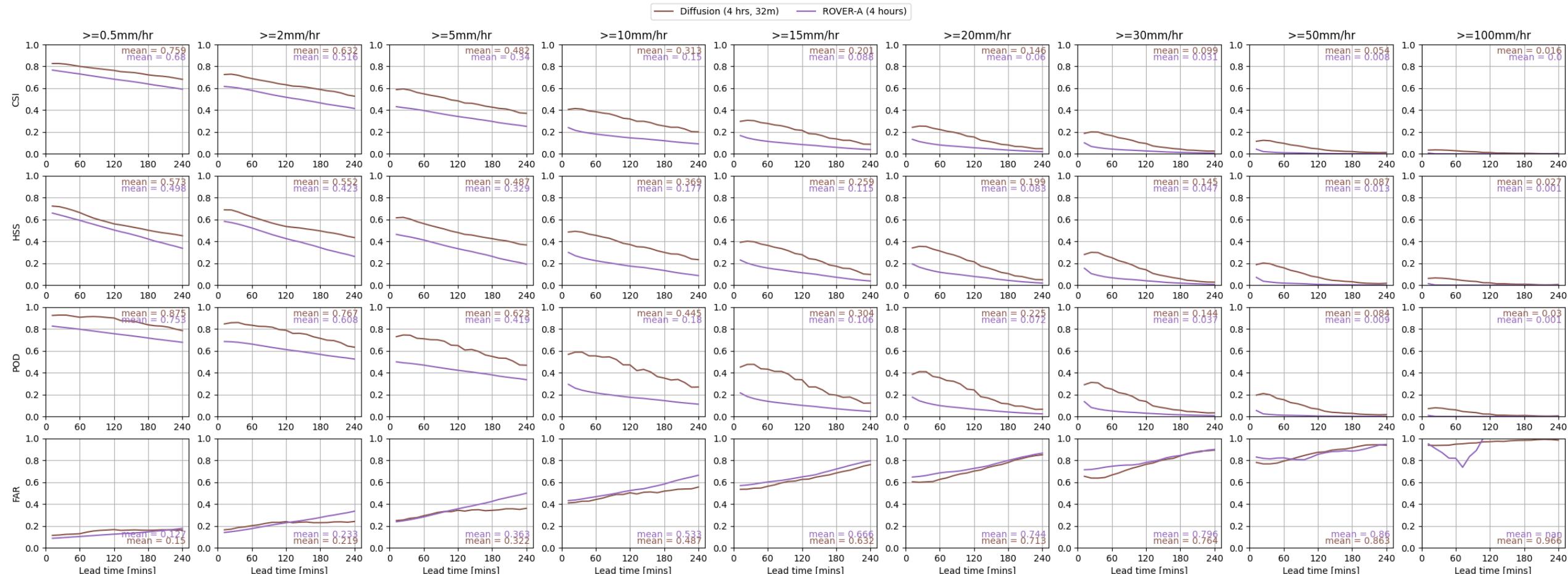
100 Highest 1-Hour Total Rainfall (mm) at the Hong Kong Observatory for All Months

No	Rank	Total Rainfall	Date	Ending Hour (HKT)
1	1	158.1	2023.09.07	24
2	2	145.5	2008.06.07	09
3	3	115.1	2006.07.16	03
4	4	109.9	1992.05.08	07
5	5	108.2	1966.06.12	07
6	6	105.9	2023.09.08	08
7	7	104.8	1989.05.02	13
8	8	100.7	1926.07.19	04
9	9	100.0	1968.06.13	03
10	10	98.7	1972.06.18	12



Diffusion (ensemble mean) vs optical flow extrapolation (CSI, HSS, POD, FAR)

Deep-Learning 480x480 from 202309071800 to 202309081200 every 6min



0.5
mm/hr

2.0
mm/hr

5.0
mm/hr

10.0
mm/hr

15.0
mm/hr

20.0
mm/hr

30.0
mm/hr

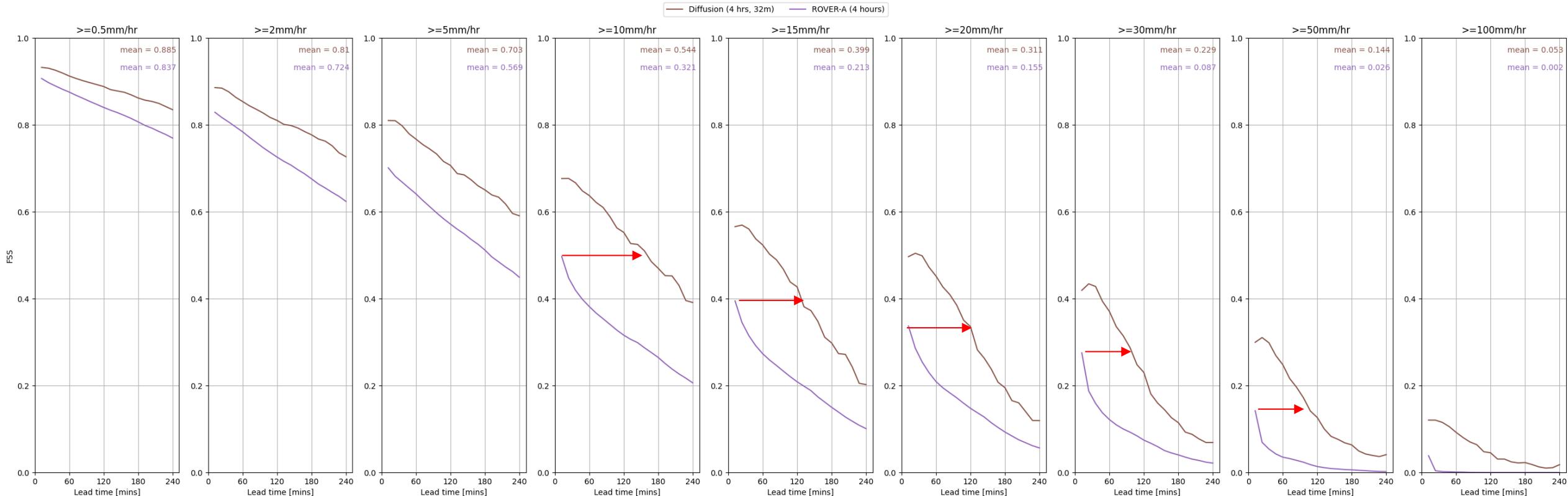
50.0
mm/hr

100.0
mm/hr

T+12 – T+240 minutes

Diffusion (ensemble mean) vs optical flow extrapolation (FSS 3x3 window size)

FSS Score win3_stride2 Deep-Learning 480x480 from 202309071800 to 202309081200 every 6min



0.5
mm/hr

2.0
mm/hr

5.0
mm/hr

10.0
mm/hr

15.0
mm/hr

20.0
mm/hr

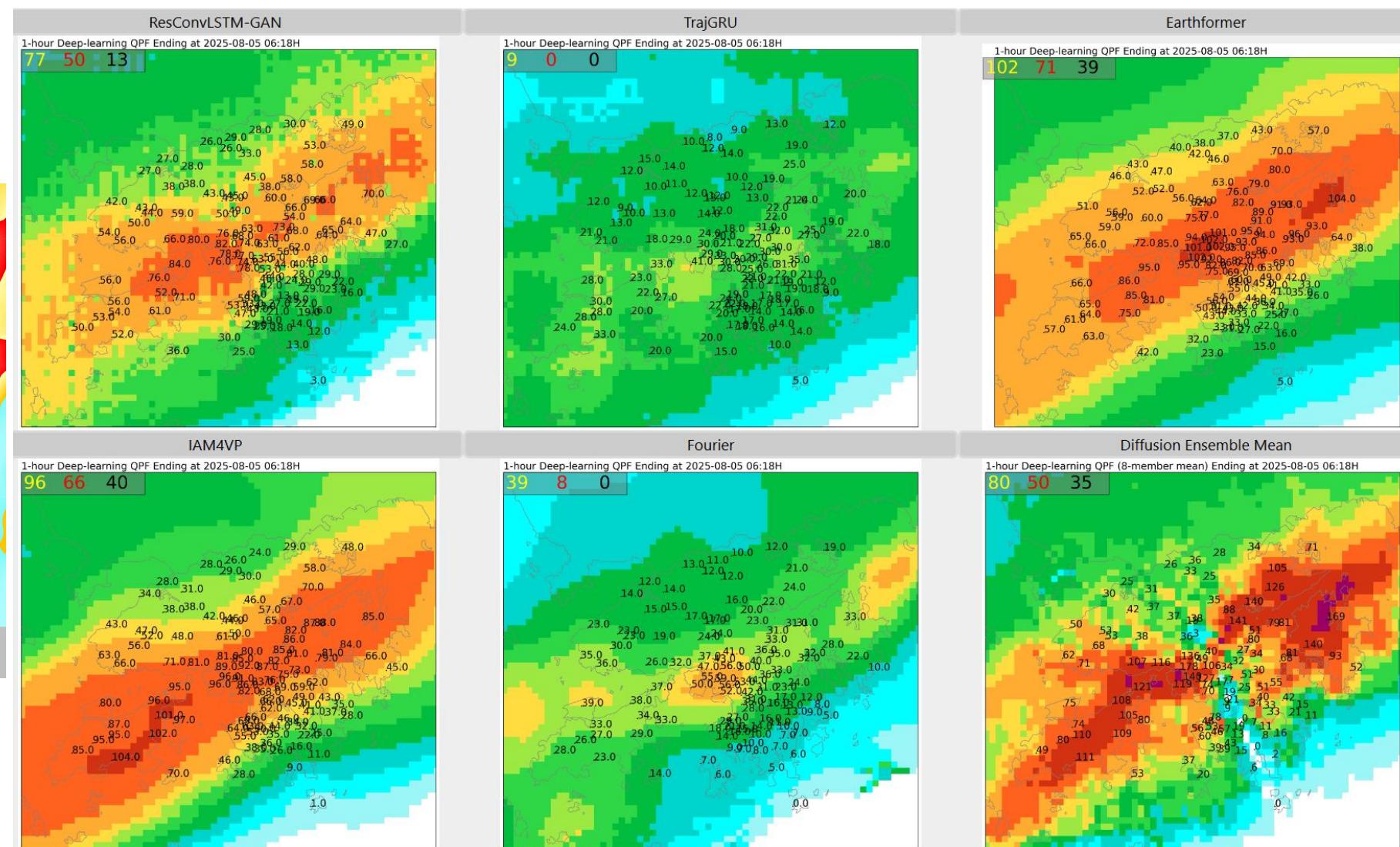
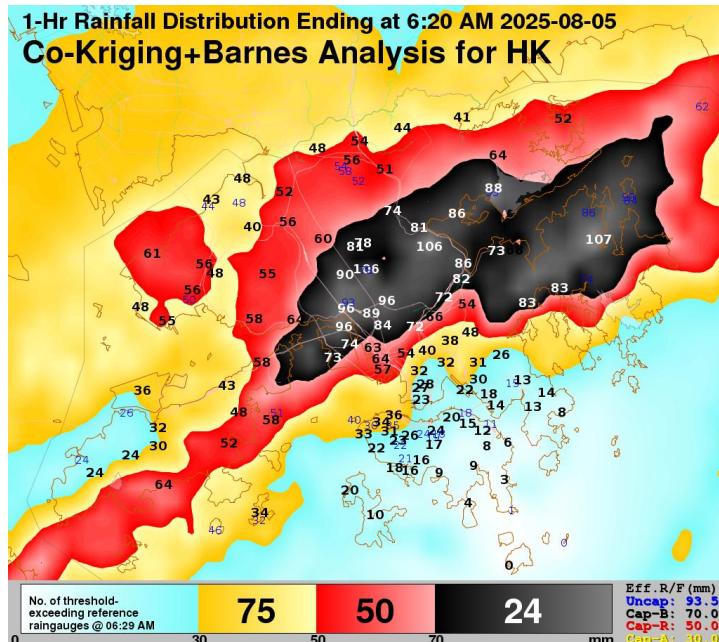
30.0
mm/hr

50.0
mm/hr

100.0
mm/hr

Black Rainstorm on 5 August 2025

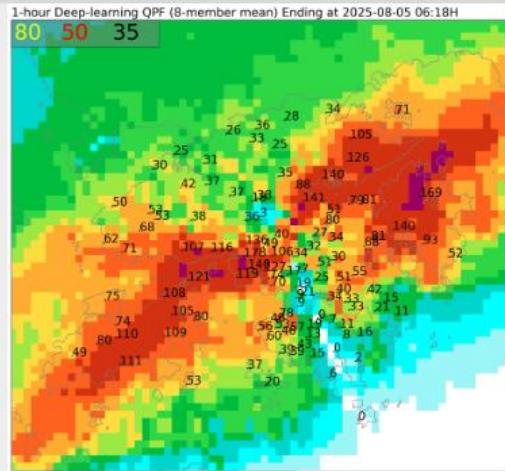
2025-08-05
06:20H



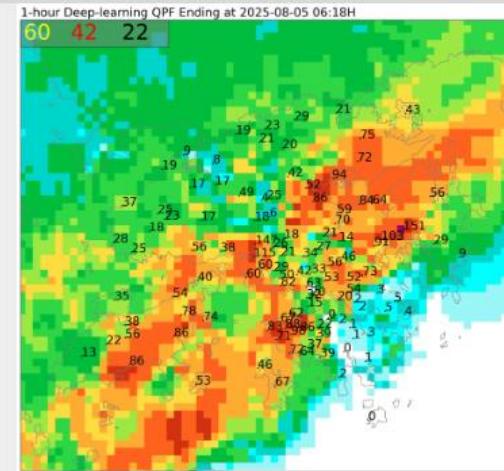
T+1h nowcast from 6 deep learning algorithms

T+1h nowcast from diffusion ensemble mean and 8 members

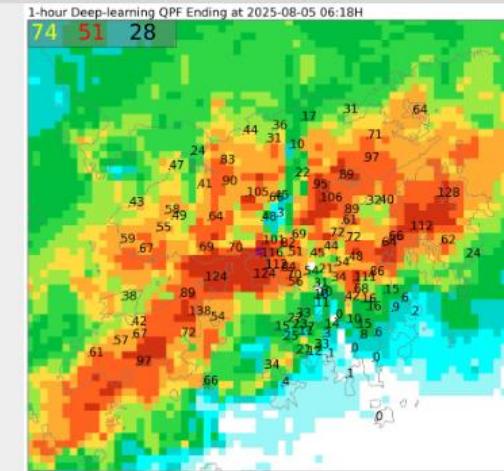
Ensemble Mean



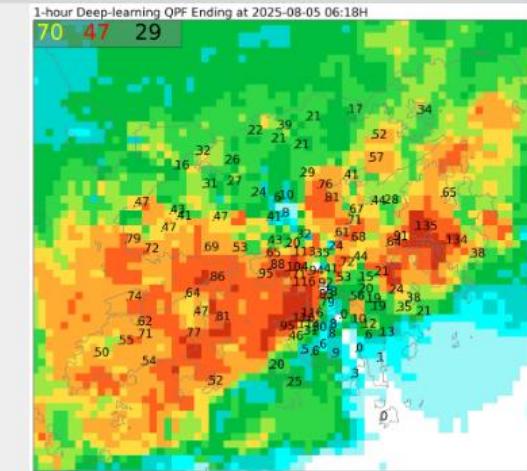
Member 1



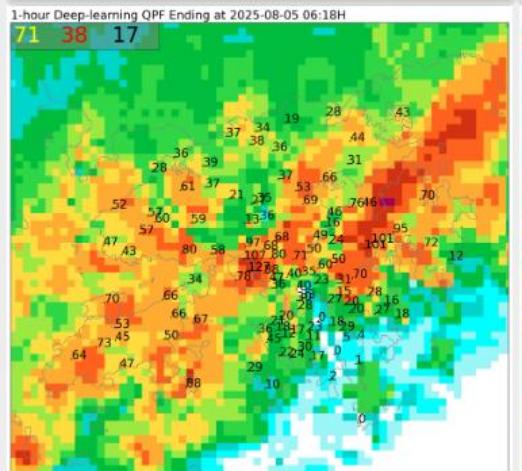
Member 2



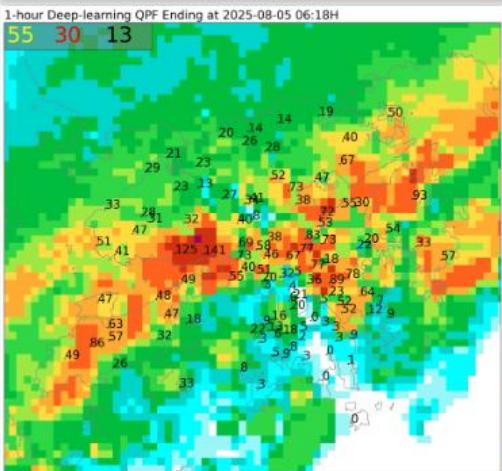
Member 3



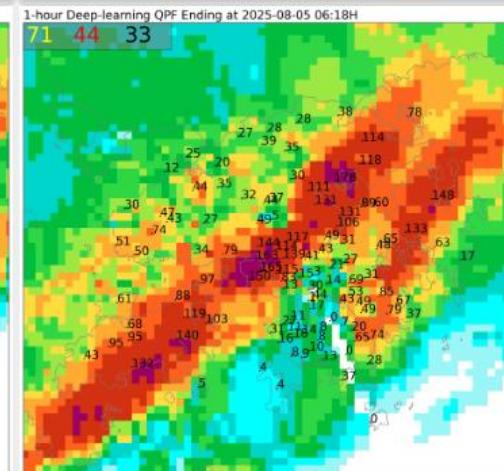
Member 4



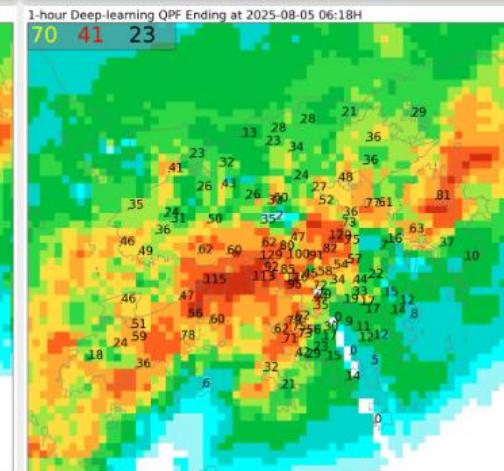
Member 5



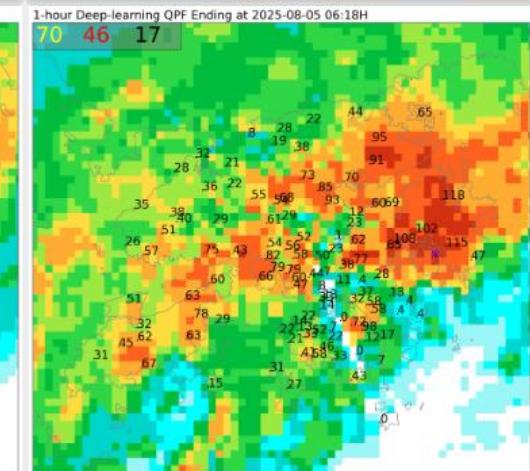
Member 6



Member 7

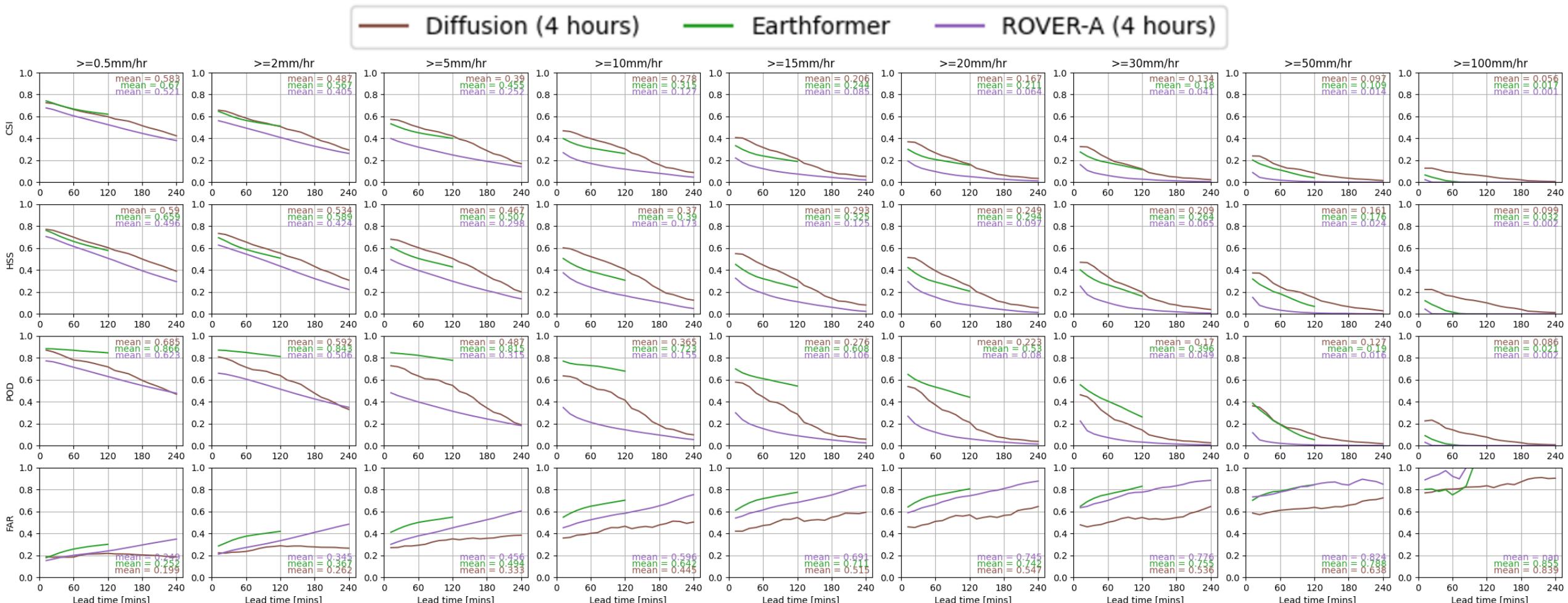


Member 8



Verification (4 – 5 Aug 2025) – CSI, HSS, POD, FAR

optical flow vs Earthformer (2 hrs) vs Diffusion ensemble mean



0.5
mm/hr

2.0
mm/hr

5.0
mm/hr

10.0
mm/hr

15.0
mm/hr

20.0
mm/hr

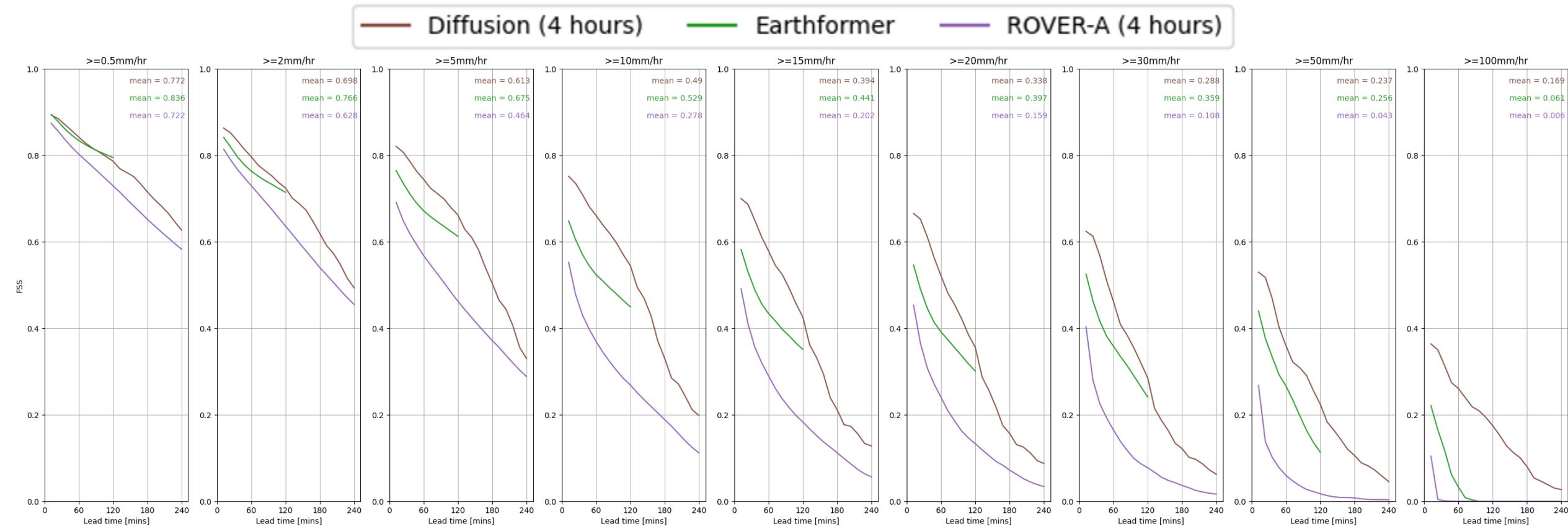
30.0
mm/hr

50.0
mm/hr

100.0
mm/hr

Verification – FSS

optical flow vs Earthformer (2 hrs) vs Diffusion ensemble mean



0.5
mm/hr

2.0
mm/hr

5.0
mm/hr

10.0
mm/hr

15.0
mm/hr

20.0
mm/hr

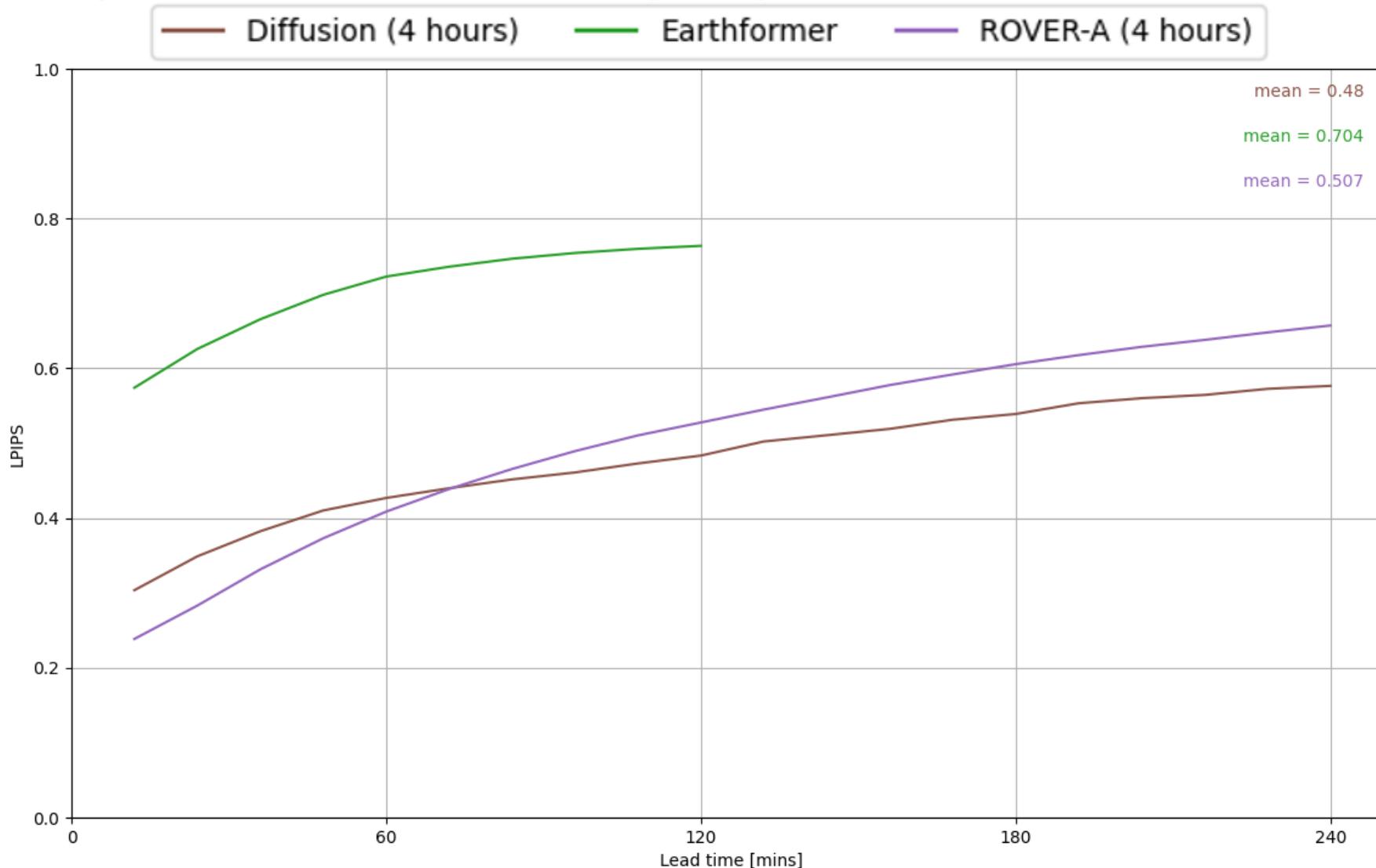
30.0
mm/hr

50.0
mm/hr

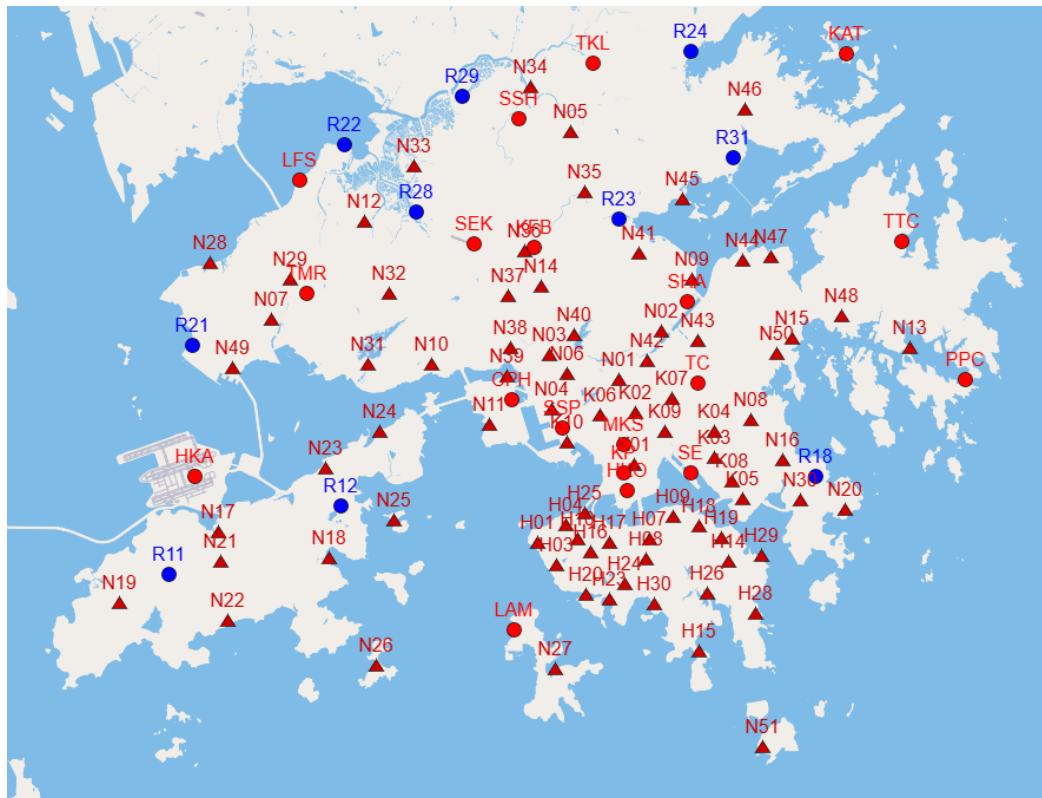
100.0
mm/hr

Verification – LPIPS

optical flow vs Earthformer (2 hrs) vs Diffusion ensemble mean



“AI of AI” in improving rainstorm prediction

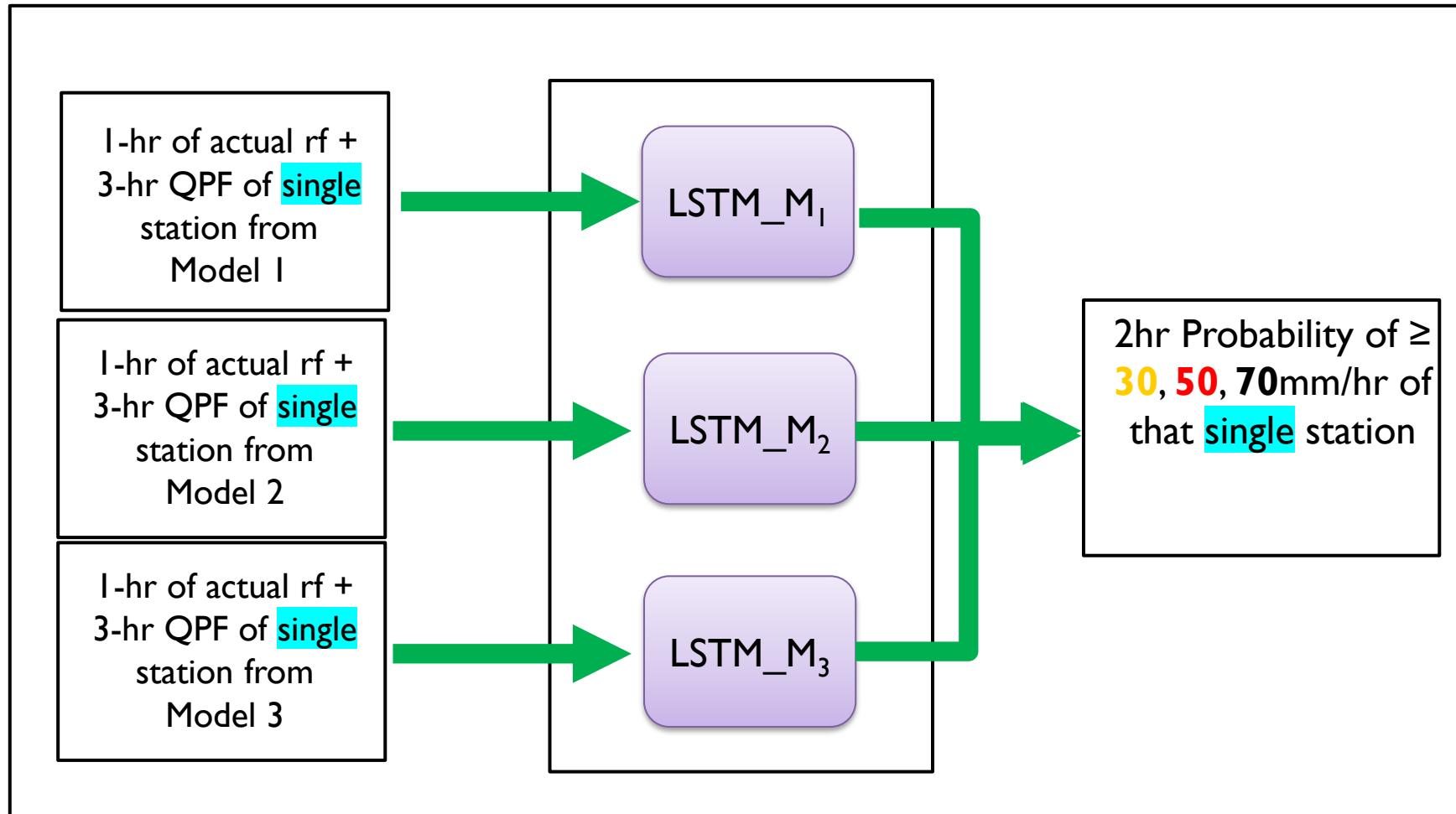


- Raingauge observations and site-specific QPFs are crucial for assessment and timeliness of issuing rainstorm warnings (Amber / Red / Black)
- SWIRLS has been providing a suite of nowcast and “post-processed” rule-based guidance based on optical-flow extrapolation, ensemble based on perturbed optical-flow field, and now into AI models and AI-ensemble being in trial
 - Note they are updated based on 6-minutely radar scan or 1-minute rapid-scan, but available time vary
- Any effective and robust way to integrate all these guidance as a one-stop assessment, rapid-update with minutely rainfall observations?

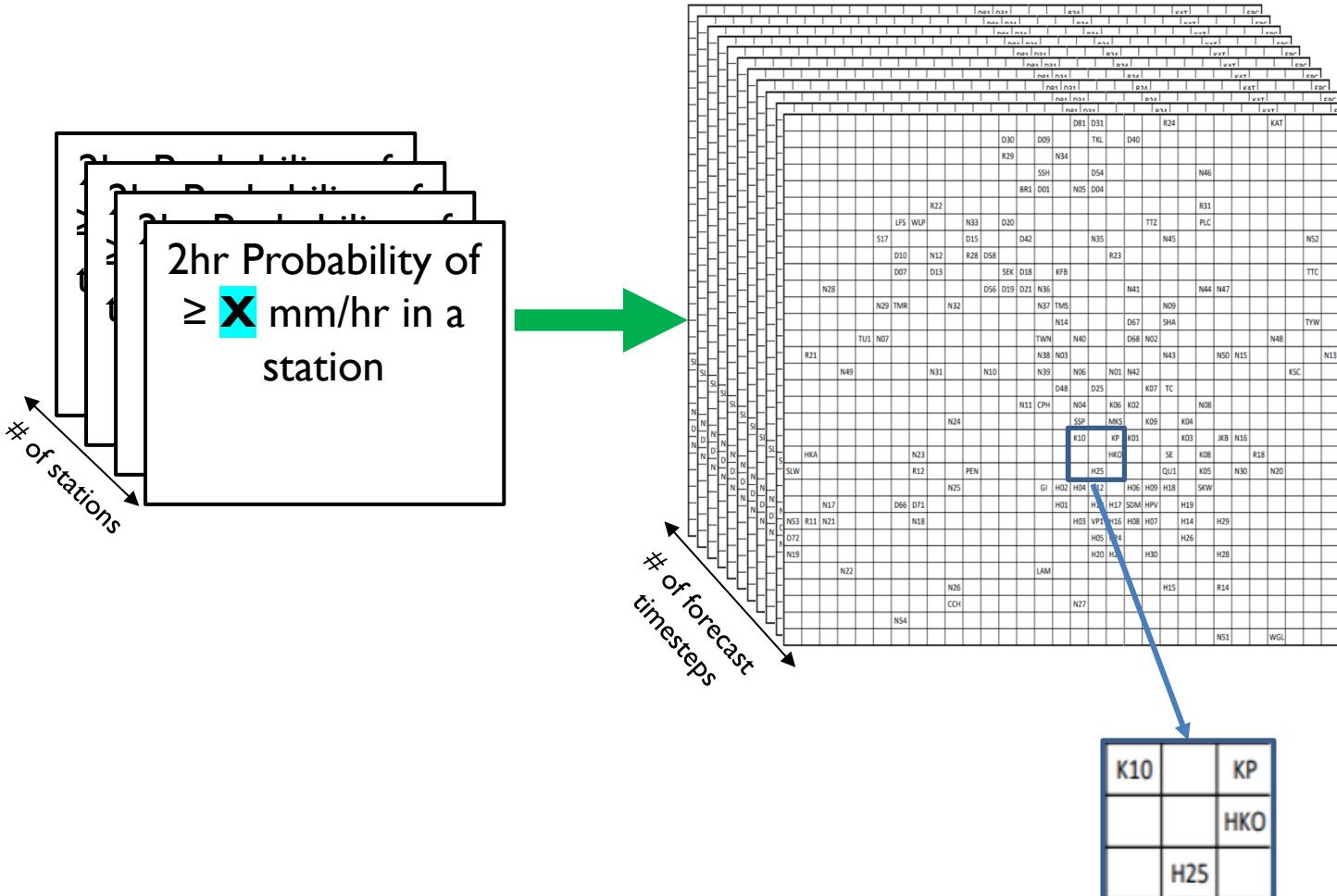
→ AI-RAINVIS (Artificial Intelligence Rainstorm Analysis and INtegrated Visuals of SWIRLS)

AI-RAINVIS Model Structure (i)

Universal Station-based Probabilistic Models



AI-RAINVIS Model Structure (ii)



**Universal Rainstorm
Probabilistic Prediction
Model**

VideoMAE
(Transformer)

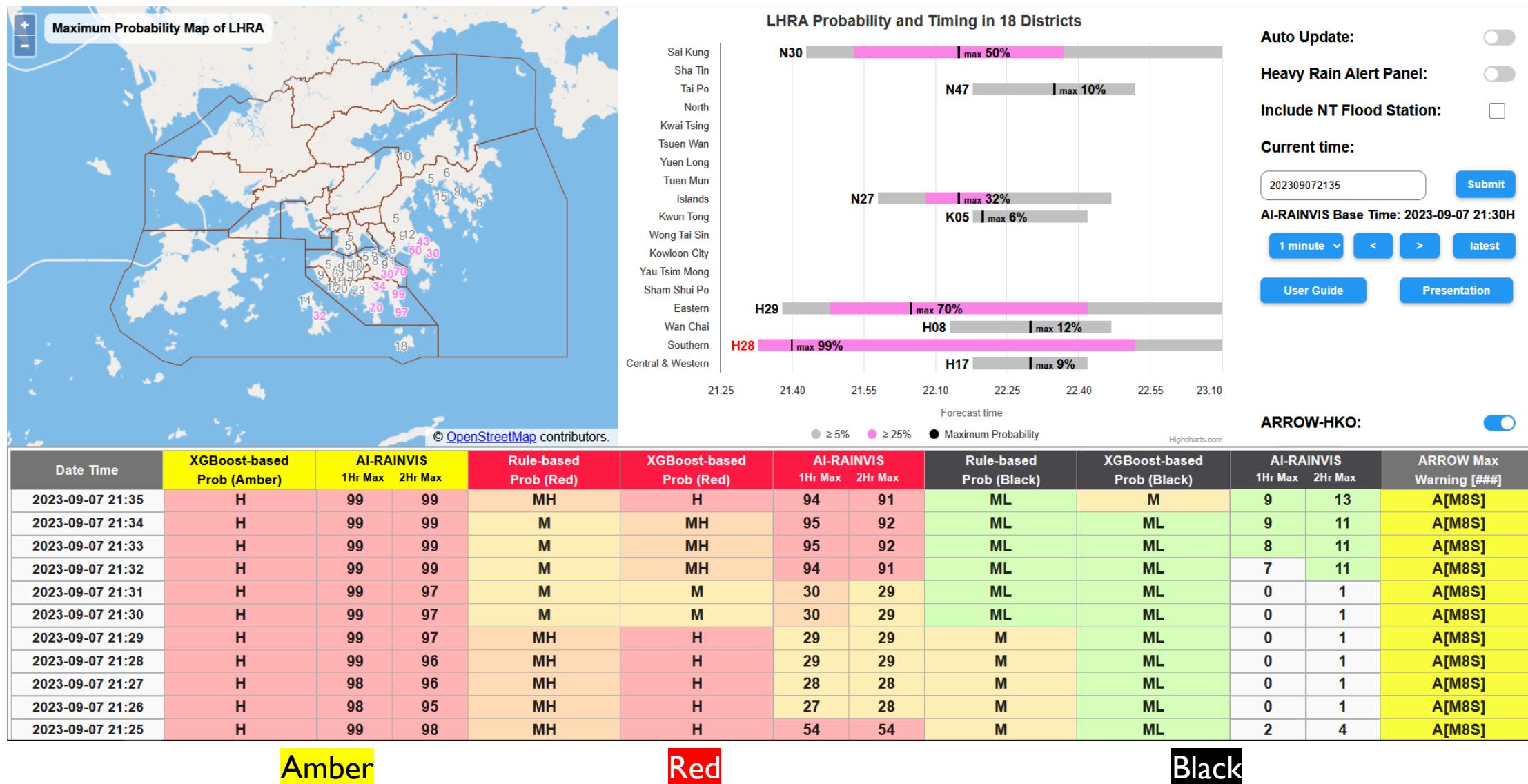
Same model structure
for each category of
rainstorm warnings X

2hr Probability of
 $\geq N$ stations
reach X mm/hr

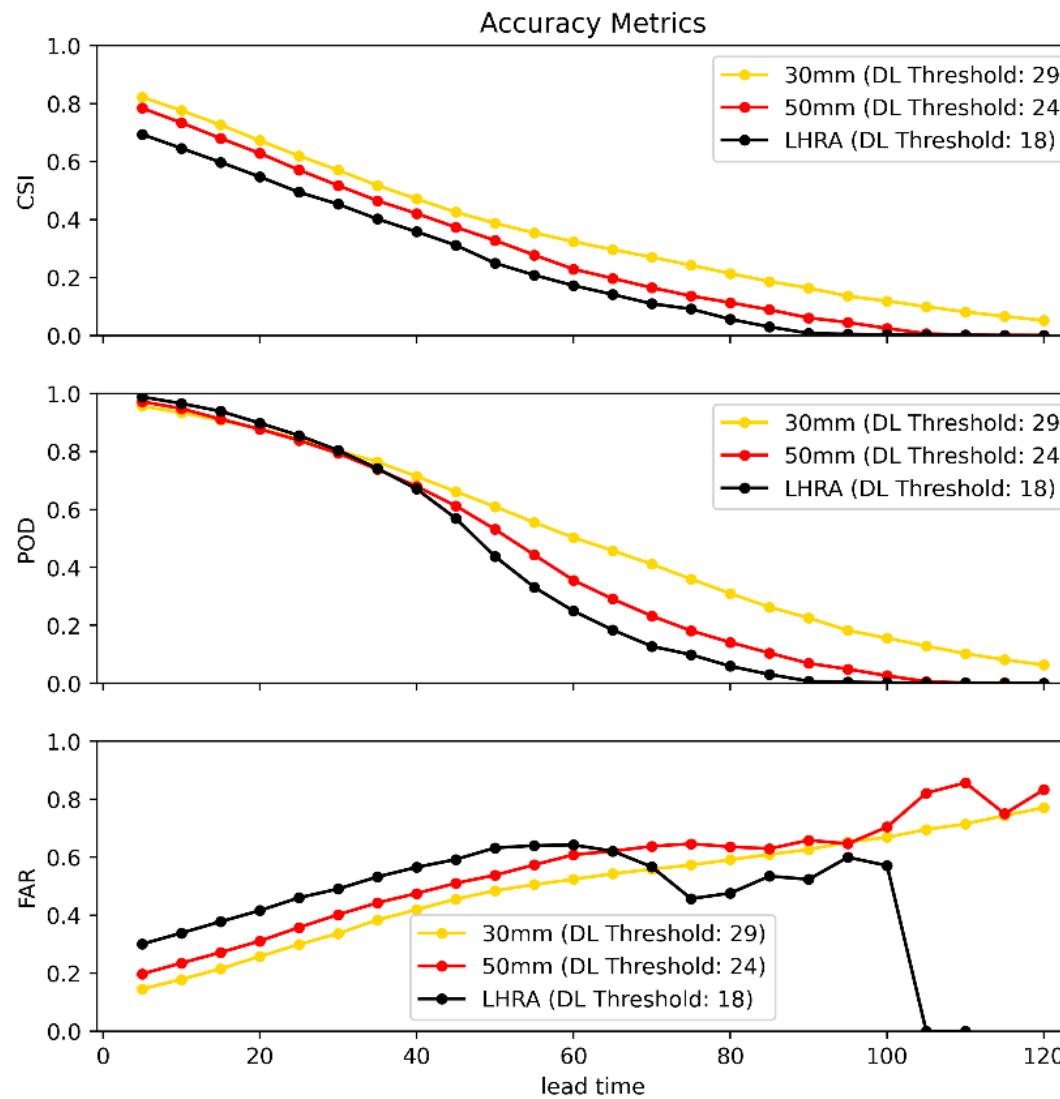
Value of X	Warning
30	Amber
50	Red
70	Black

AI-RAINVIS Web-interface

(Artificial Intelligence Rainstorm Analysis and INtegrated Visuals of SWIRLS)



AI-RAINVIS Verification



Prob	CSI	Lead time [25/50/75p in min]		
Amber (MH)	0.257	22	44	65
Red (MH)	0.537	39	61	81
Black (MH)	0.450	53	68	84
Black (H)	0.521	0	3	25

Verification against actual # of raingauges reaching Amber/Red/Black level when respective probability category is met

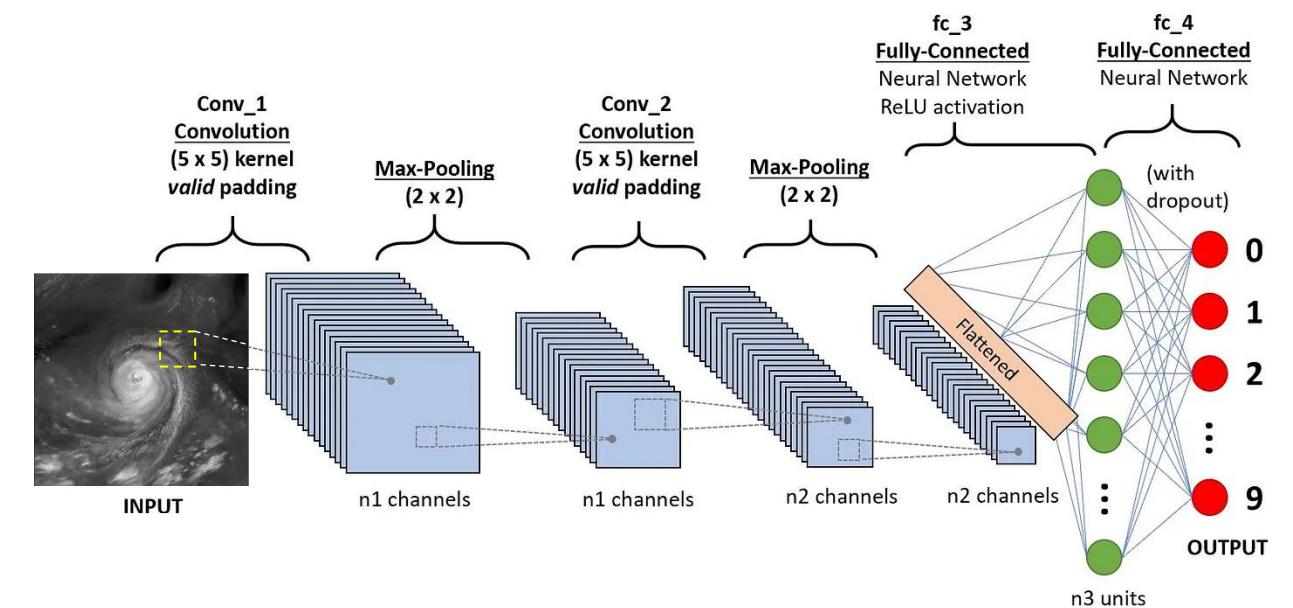
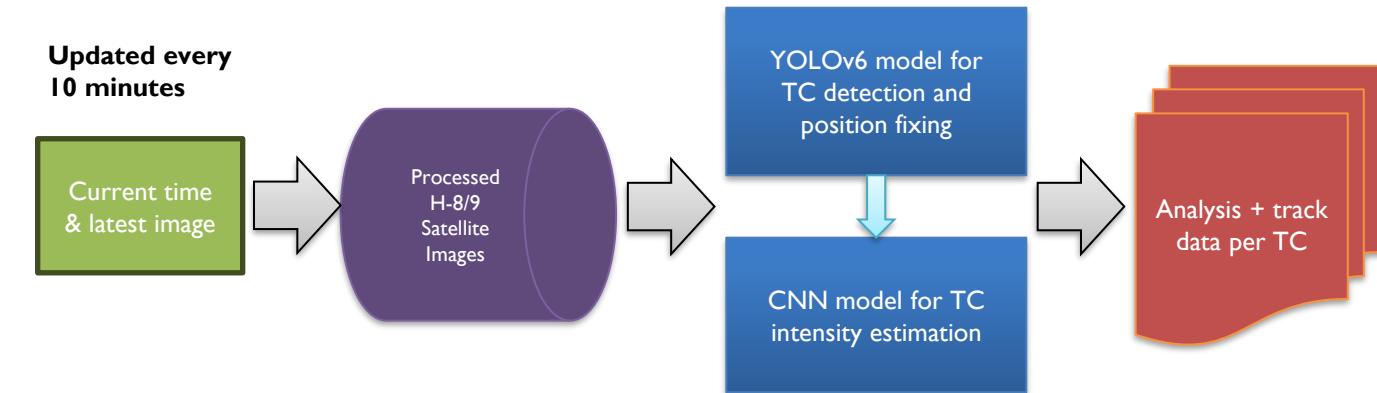
Way forward on AI Nowcast Support

- Deep learning nowcast algorithms
 - More efficient diffusion based technique (spatio-temporal latent diffusion) to speed up training and inference
 - Enhancement of satellite-based reflectivity and rainfall nowcasts (Earthformer, diffusion)
 - Physics constraints and inputs from NWP / AIWP
- “AI of AI” approach to add values
 - Extend ML technique to LLM (or SLM/VLM for specific high impact convective weather or site-specific assessments)
- **AI-based automatic tropical cyclone detection and intensity analysis**
- **Regional collaboration on AI nowcasting techniques**

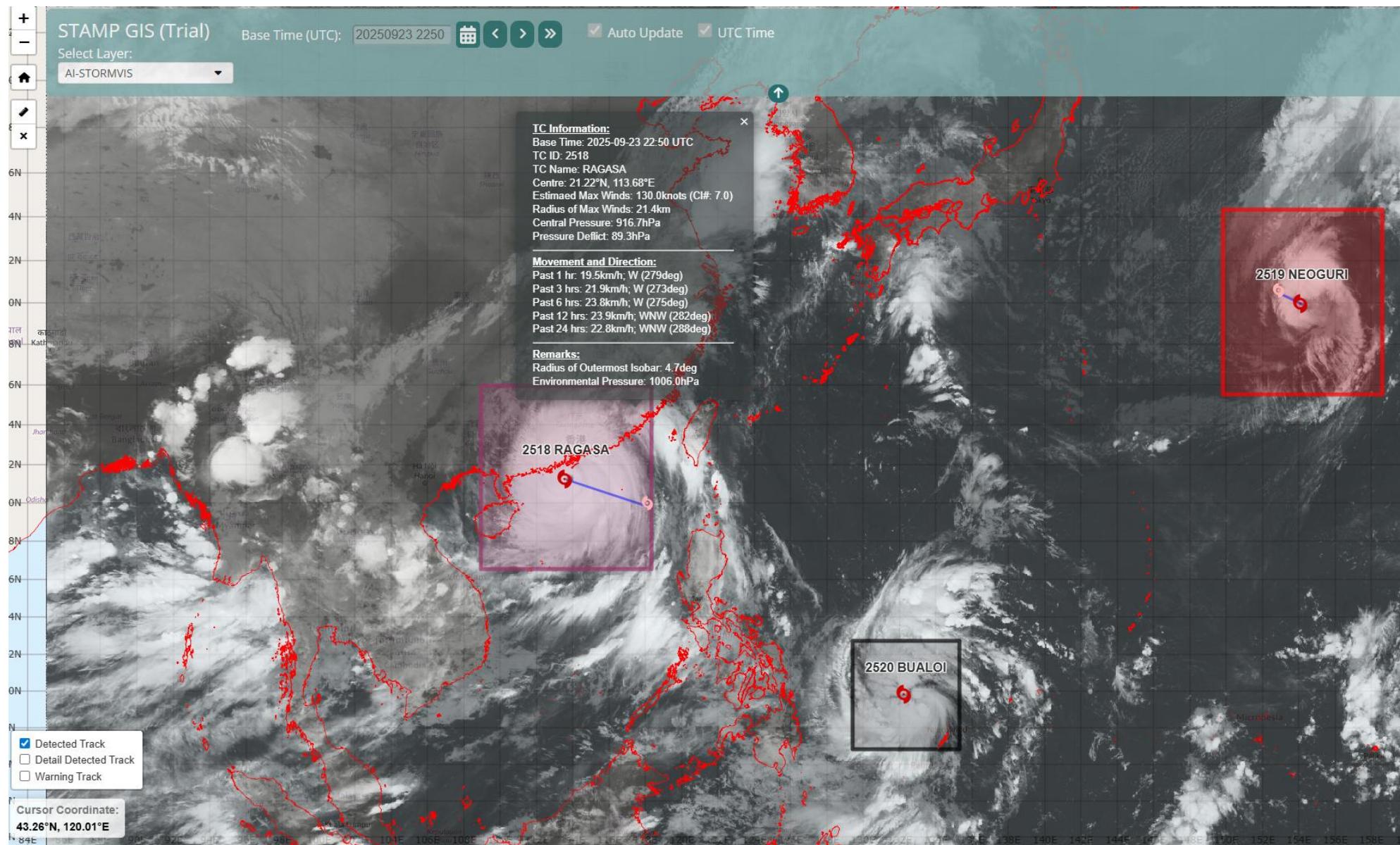
AI in Tropical Cyclone Analysis and Nowcast

• AI-STORMVIS

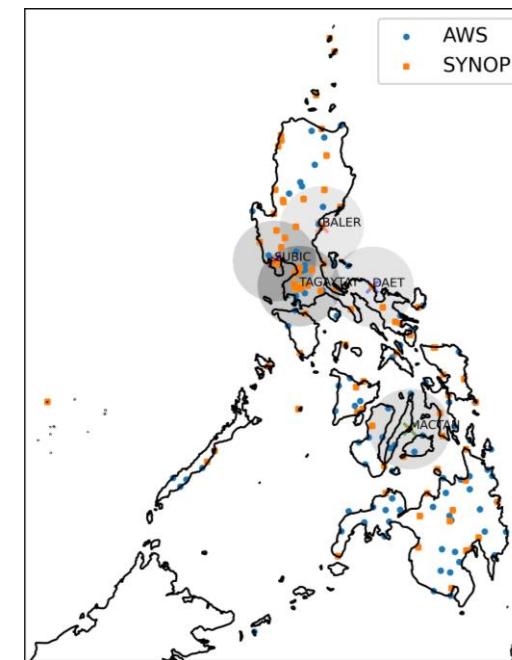
- AI-driven **S**atellite-based **T**C **O**bject **R**ecognition, **M**otion **V**isualisation and **I**ntensity estimation **S**ystem
- TC detection and position fixing
 - **Y**ou **O**nly **L**ook **O**nce (YOLO) small object detection model
 - Ensemble detection models
- TC intensity estimation
 - Convolutional neural network (CNN)
- Visualisation web-based platform



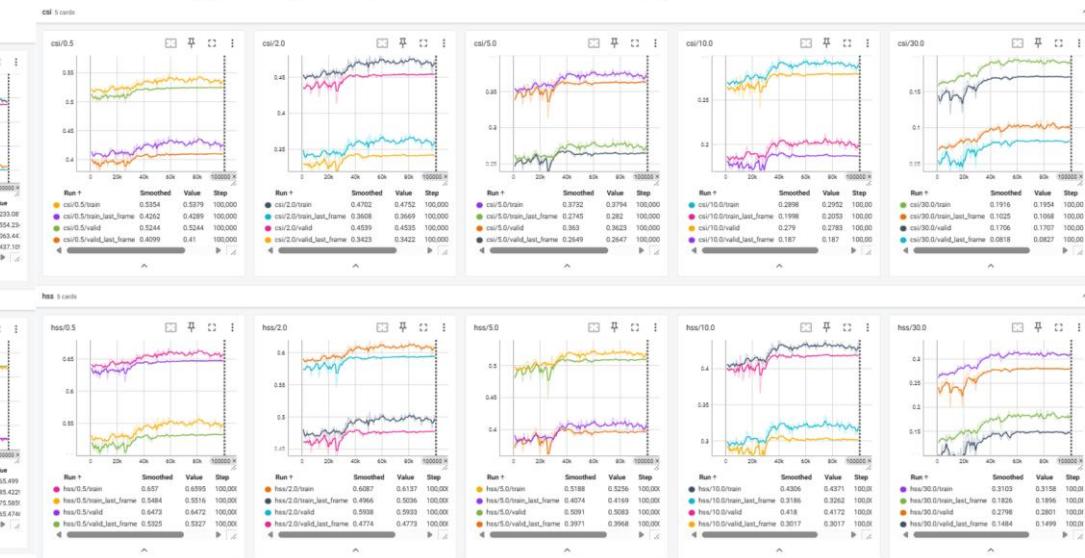
AI-STORMVIS on RSMC website (under trial)



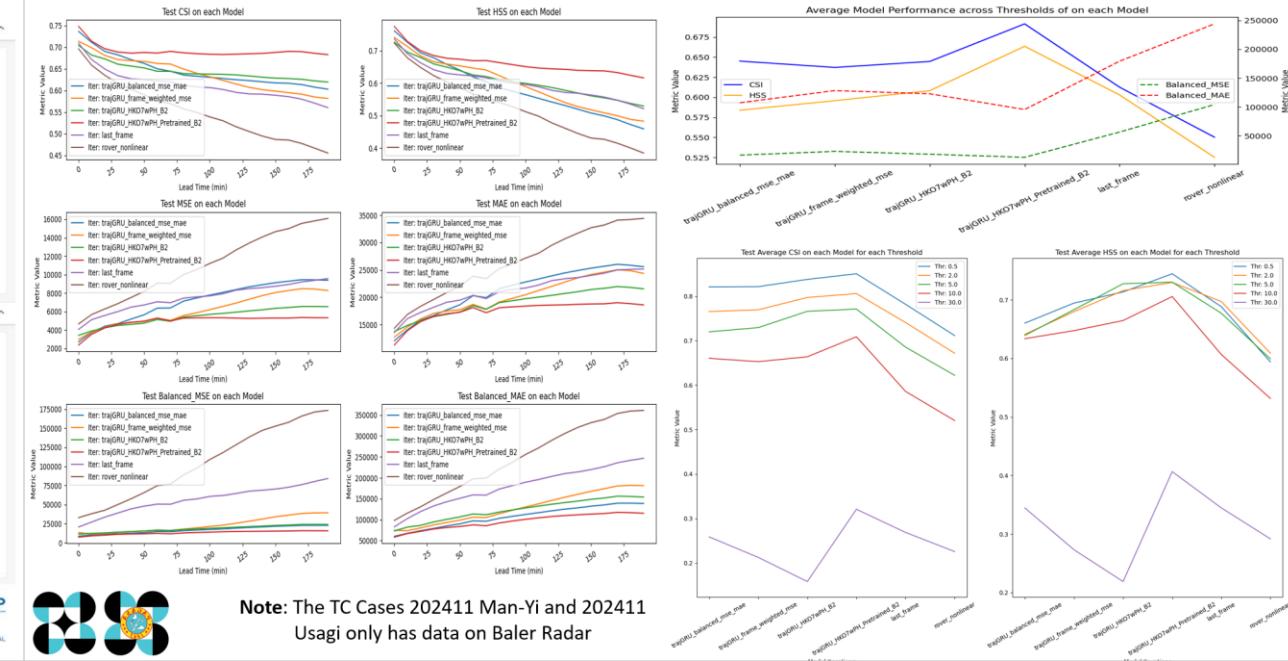
Collaboration with **PAGASA** and **TMD** in knowledge transfer and implementing Com-SWIRLS and deep learning nowcast models



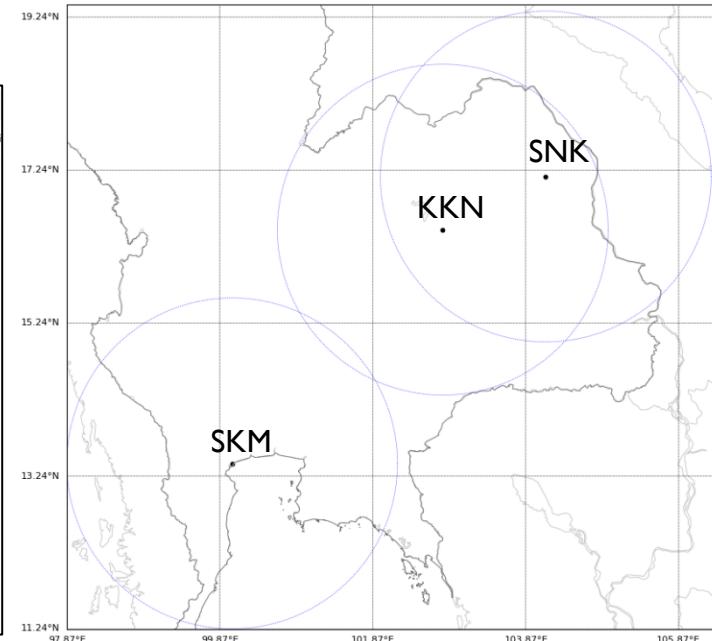
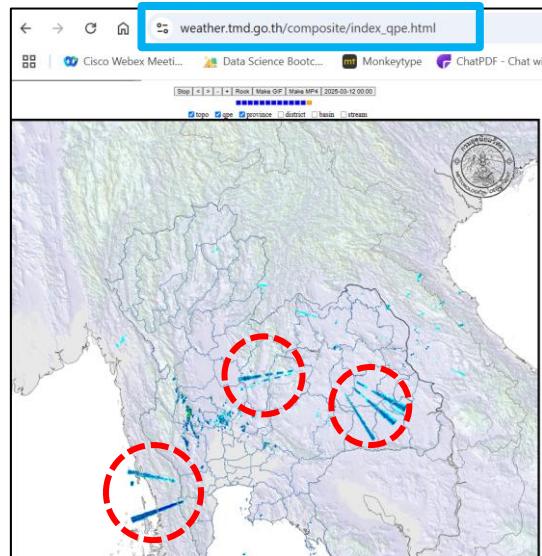
Model Training: TrajGRU with Pretraining, Individual Radars



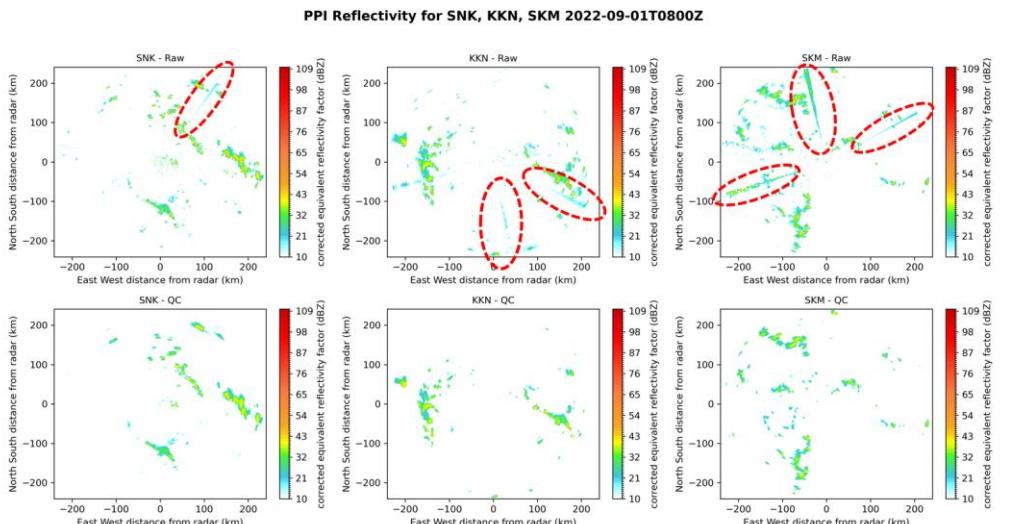
Model Evaluation – Test TC Data



Collaboration with **PAGASA** and **TMD** in knowledge transfer and implementing Com-SWIRLS and deep learning nowcast models



Radar Composite in PPI Lowest Elevation Technique



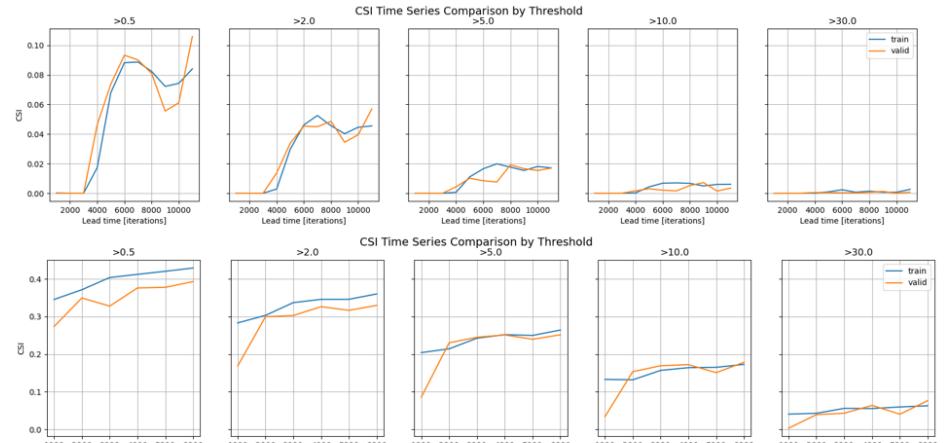
TH only

TH + HKO7

Xingjian Shi, Zhihan Gao, Leonard Lausen, Hao Wang, Dit-Yan Yeung, Wai-kin Wong, and Wang-chun Woo, 2017:
Deep learning for precipitation nowcasting: A benchmark and a new model

Worked for TrajGRU Model: Result (TH only VS TH + HKO7 – KKN site)

(After revising mask file and frequency sequence time)





香港天文台
HONG KONG OBSERVATORY

Thank you very much

