# AI Technology in Nowcasting

Wai Kin WONG
Hong Kong Observatory
Chair / ET-WIPPSDE

AI Nowcasting - Agenda Item 4.1

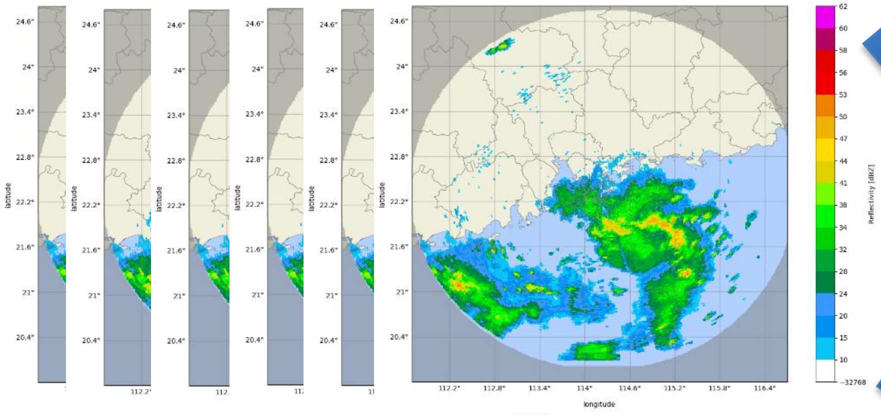
12 November 2024

WORLD METEOROLOGICAL ORGANIZATION

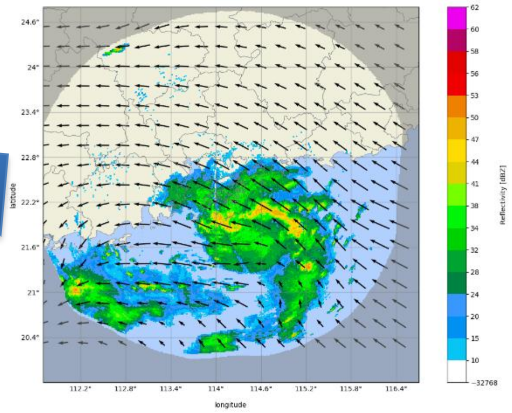
# AI in Precipitation Nowcasting

Sequence of actual radar images

$\{\mathbf{X}_{t-m}, \mathbf{X}_{t-m+1}, \ldots, \mathbf{X}_{t-1}, \mathbf{X}_{t}\}$

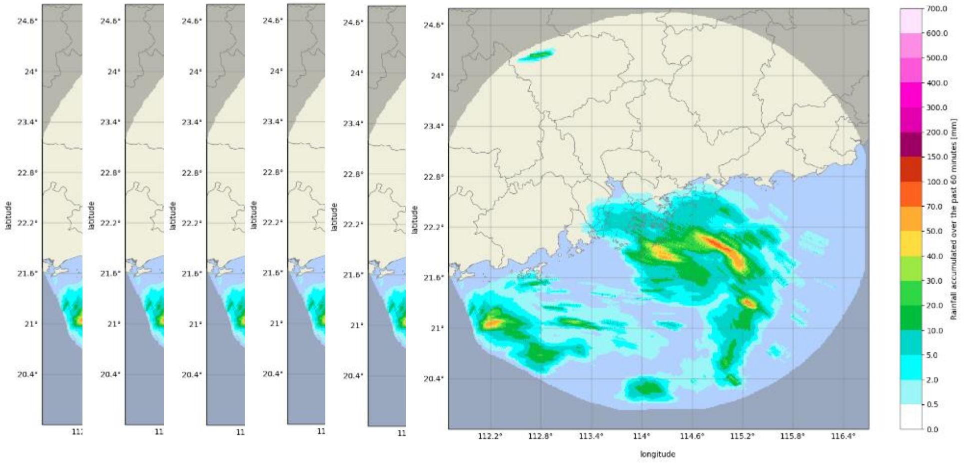


Motion of radar echoes from
computer vision tracking



Rainfall or significant convection nowcasts

$\{\mathbf{R}_{t+1}, \mathbf{R}_{t+2}, \ldots, \mathbf{R}_{t+n-1}, \mathbf{R}_{t+n}\}$



$$x_{t-m:t} \longrightarrow \boxed{\mathcal{M}} \longrightarrow x_{t+1:t+n}$$

**Deep Learning Model**

# Learning the "model" in multi-step forecasting problem (1)

## Iterative Multi-step Estimation

$$p(\mathbf{X}_{t+1:t+L} \mid \mathbf{X}_{t-J+1:t}) \quad = \quad \prod_{i=1}^{L} p(\mathbf{X}_{t+i} \mid \mathbf{X}_{t-J+1:t+i-1}; \boldsymbol{\theta})$$

Optimal parameter estimated from maximum likelihood

$$\boldsymbol{\theta}^{\star} = \arg\max_{\phi} \mathbb{E}_{\hat{p}_{\text{data}}} \left[ \sum_{i=1}^{L} \log p(\mathbf{X}_{t+i} \mid \mathbf{X}_{t-J+1:t+i-1}; \boldsymbol{\theta}) \right]$$

Advantages:
- easy to train because it only requires optimizing for the one-step-ahead forecasting error
- able to predict for an arbitrary horizons in the future by recursively applying the basic "forecaster"

Discrepancy:
- in training phase, we use the ground-truths from t + 1 to t + i - 1 to predict the regional rainfall at timestamp t + i, which is also known as "teacher-forcing"
- however, in the testing phase, we feed the model predictions instead of the ground-truths back to the forecaster. This makes the model prone to accumulative errors in the forecasting process

# Learning the "model" in multi-step forecasting problem (2)

## Direct Multi-step Estimation

Use a different parameter $\theta_i$ for each forecasting horizon *i*

$$p(\mathbf{X}_{t+i} \mid \mathbf{X}_{t-J+1:t}; \boldsymbol{\theta}_i)$$

$$\{\boldsymbol{\theta}_1^\star, \ldots, \boldsymbol{\theta}_L^\star\}$$ Set of optimal parameters

$$\boldsymbol{\theta}_1^\star, \ldots, \boldsymbol{\theta}_L^\star = \arg\max_{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_L} \mathbb{E}_{\hat{p}_{\text{data}}} \left[ \sum_{i=1}^{L} \log p(\mathbf{X}_{t+i} \mid \mathbf{X}_{t-J+1:t}; \boldsymbol{\theta}_i) \right]$$

**Note:**

1. usually more accurate predictions when model is (a) ill-specified, or (b) sequences are non-stationary, or (c) the training set is too small. However, this estimation is more computationally expensive than the iterative approach

2. When applied for precipitation nowcasting, DL models adopt learning strategies (or a mix of above) called scheduled sampling.

# Iterative / Recursive Approach

- Non-adversarial framework
  - Effective in capturing spatial-temporal structures or dependency
  - Examples: ConvLSTM, TrajGRU, PredRNN, Metnet
  - Limitation: increased blurriness with time
    - Augmented / guided by multi-sensor inputs (satellite) and dynamical model (NWP / EPS) to increase lead time (e.g. Metnet)
    - Increased sharpness by enhancing loss function terms (e.g. regularization and cross-entropy in DB-RNN) or segmented into various types (intensity) of precipitation

- Adversarial framework
  - Generate prediction with sharpness
  - Examples: GAN, TS-RainGAN, DGMR
  - Limitation: model instability, collapse in modes / features

# Multistep Approach and Model Frameworks:

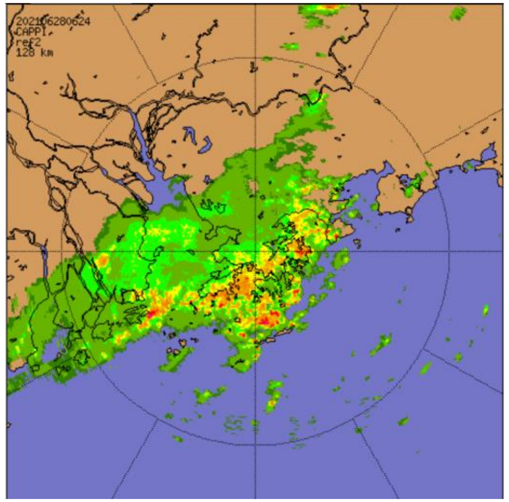
- UNet-based model
  - Effective in multivariate forecasting
  - Enhance design of loss function or model framework to represent temporal dependency (NowcastNet, MSSTNet, …)

- Transformer-based model
  - Physics-informed approach and consider temporal dependency at longer lead time (Earthformer, Rainformer …)
  - More computational intensive

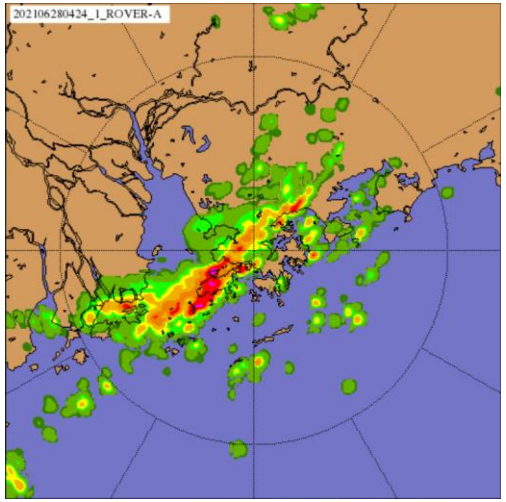
- Diffusion method
  - Generate prediction with sharpness (though sometimes too aggressive) and reliability (to quantify uncertainty)
  - Computational intensive as denoising inputs sequentially over multiple time steps is necessary
  - Physics-informed design (Prediff, Diffcast, CasCast …)

# Two-hour radar reflectivity nowcasts from optical flow and DL models

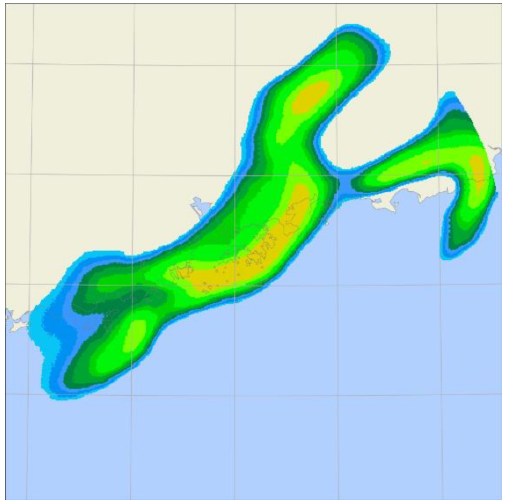
Actual



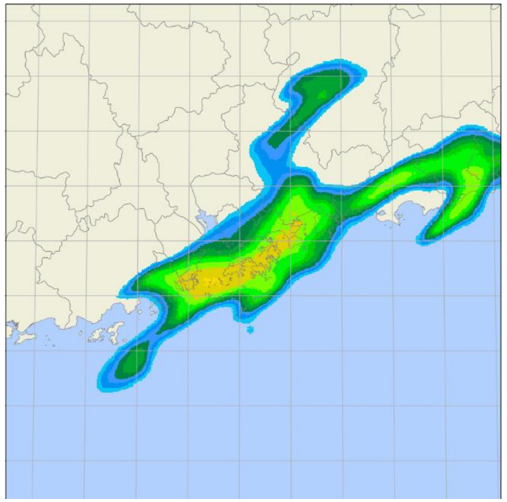
Optical flow




ConvLSTM



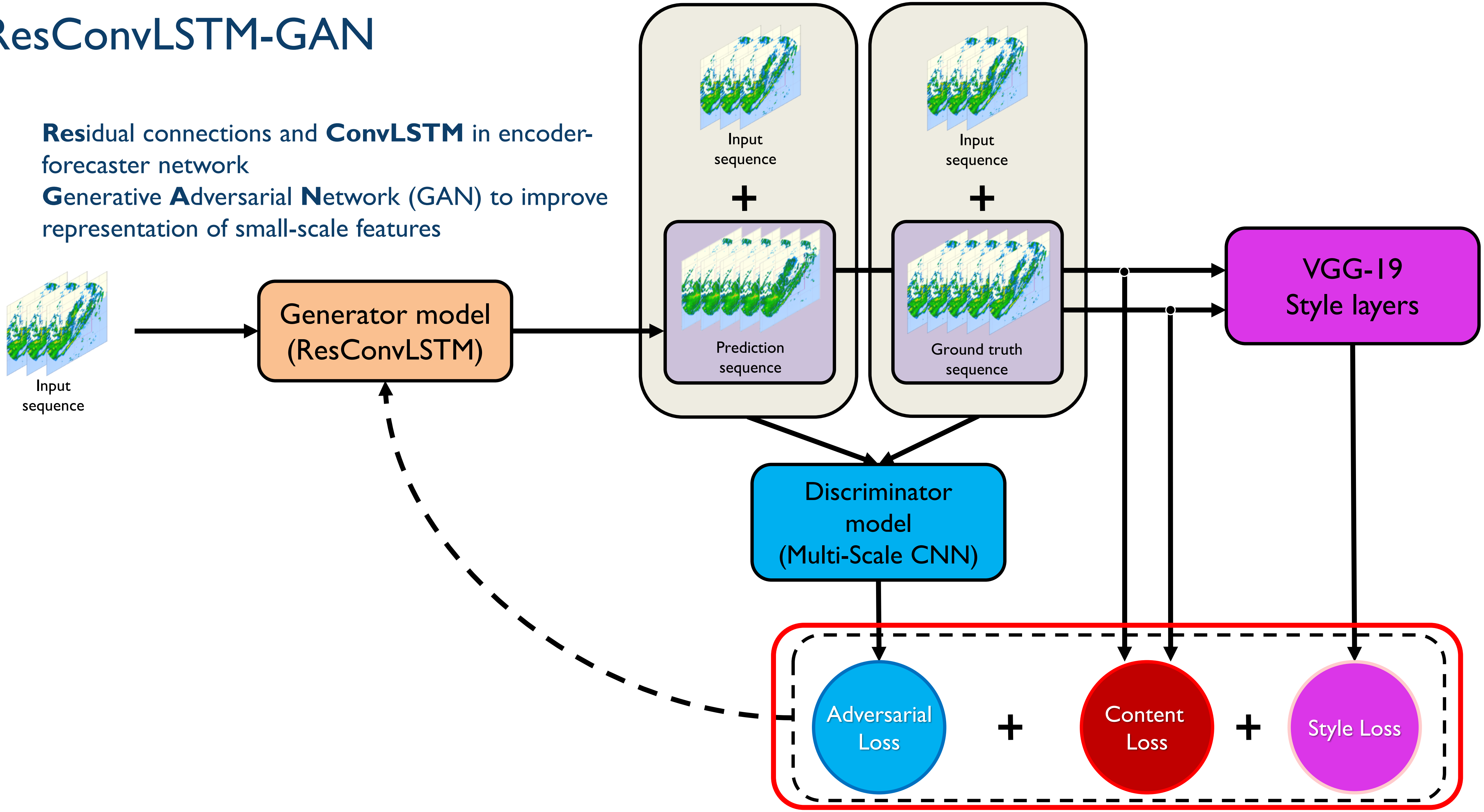
TrajGRU



- Better at predicting the movement and intensity evolution radar reflectivity than extrapolation using gridded optical flow field
- Generally, sequence-to-sequence video prediction model trained with simple loss function (MSE / MAE) faces the blurring problem.

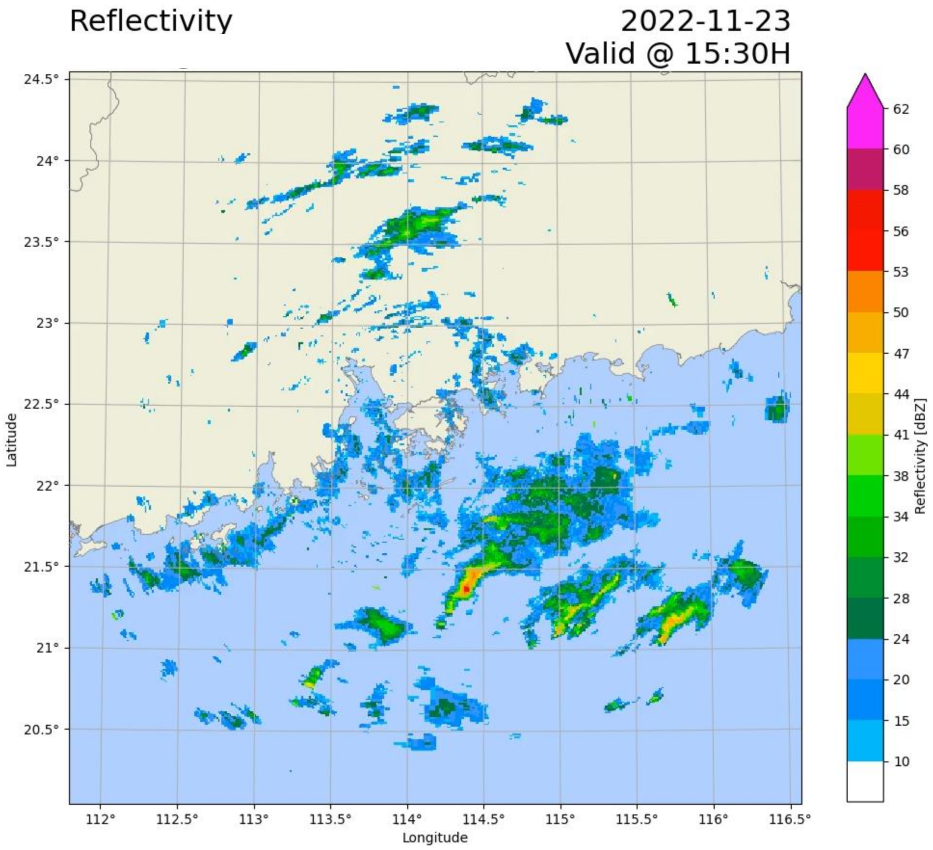
# ResConvLSTM-GAN

- **Res**idual connections and **ConvLSTM** in encoder-forecaster network
- **G**enerative **A**dversarial **N**etwork (GAN) to improve representation of small-scale features

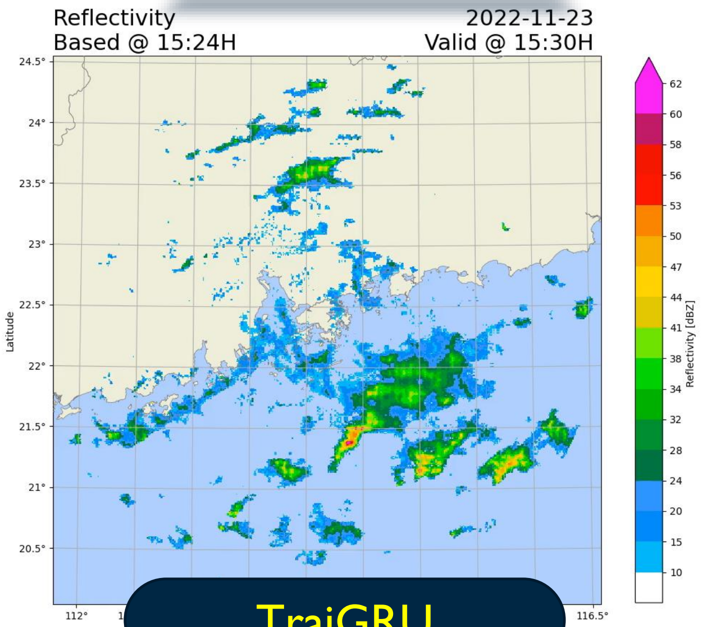


# Application of ResConvLSTM-GAN

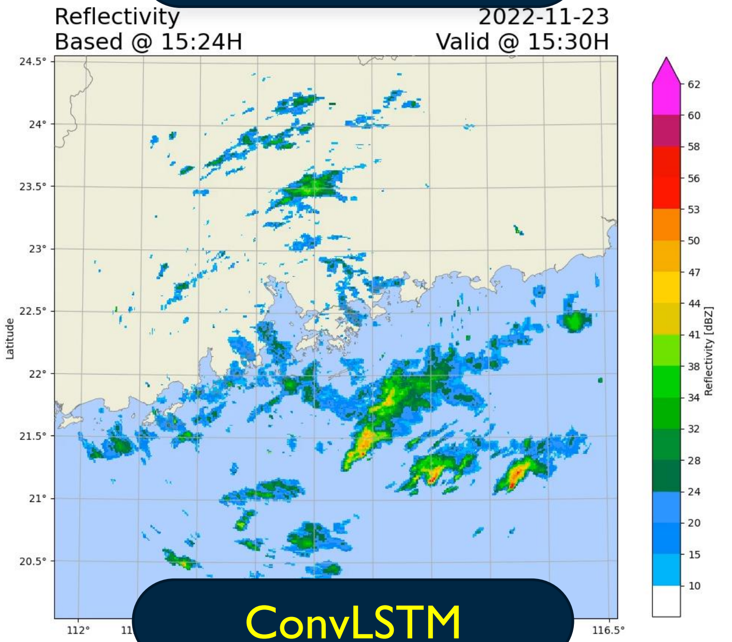
## Actual Observations



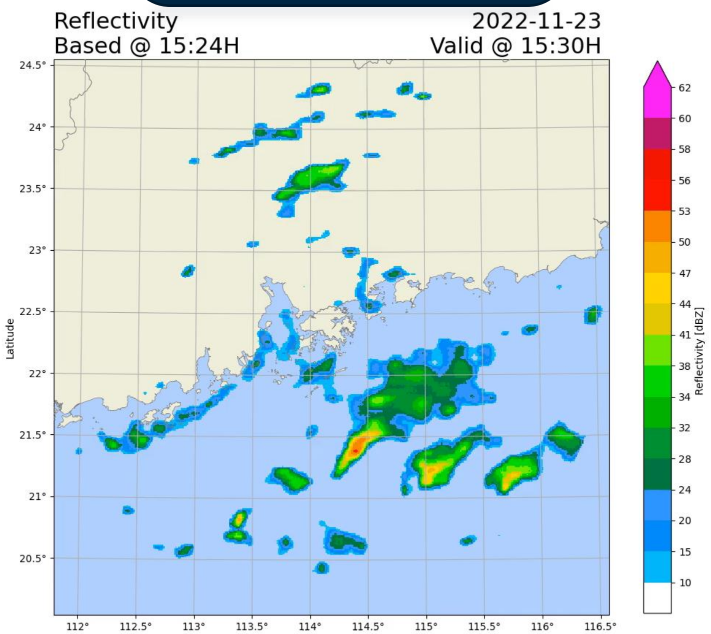

## ResConvLSTM-GAN



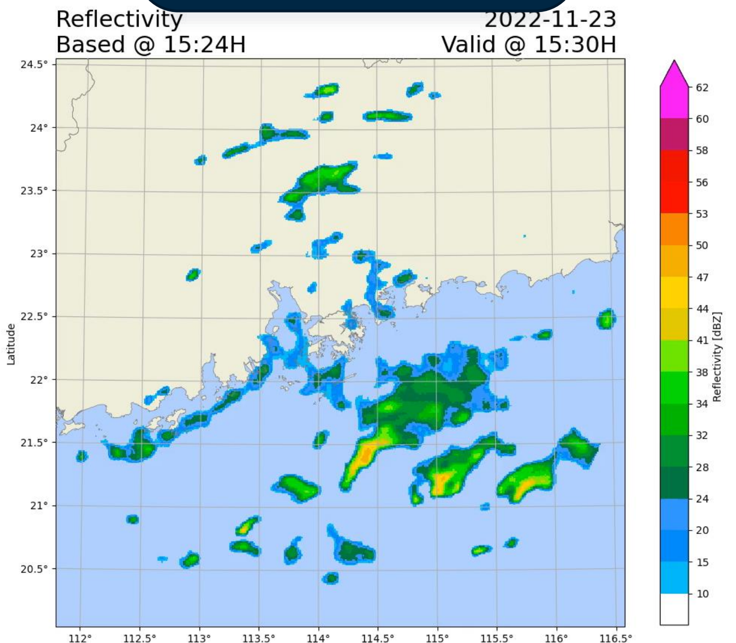

## Optical Flow



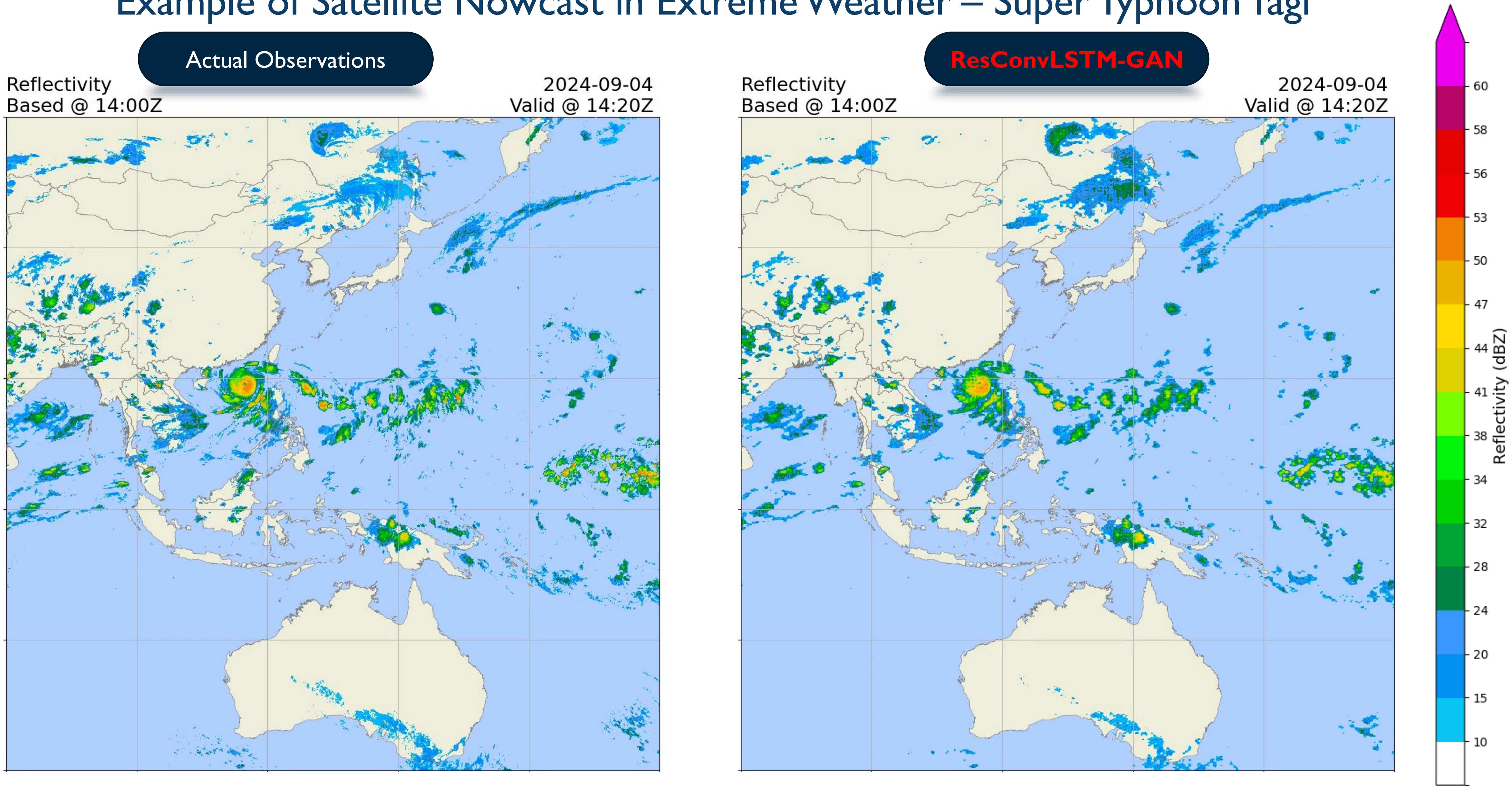

## TrajGRU




## ConvLSTM



# Example of Satellite Nowcast in Extreme Weather – Super Typhoon Yagi



- Denoising Diffusion Probabilistic Model (DDPM) aims to capture the data distribution by training a neural network to undo a Markov noising process that gradually distorts the data

A two-phase training approach on large-scale feature (deterministic model) and residual components

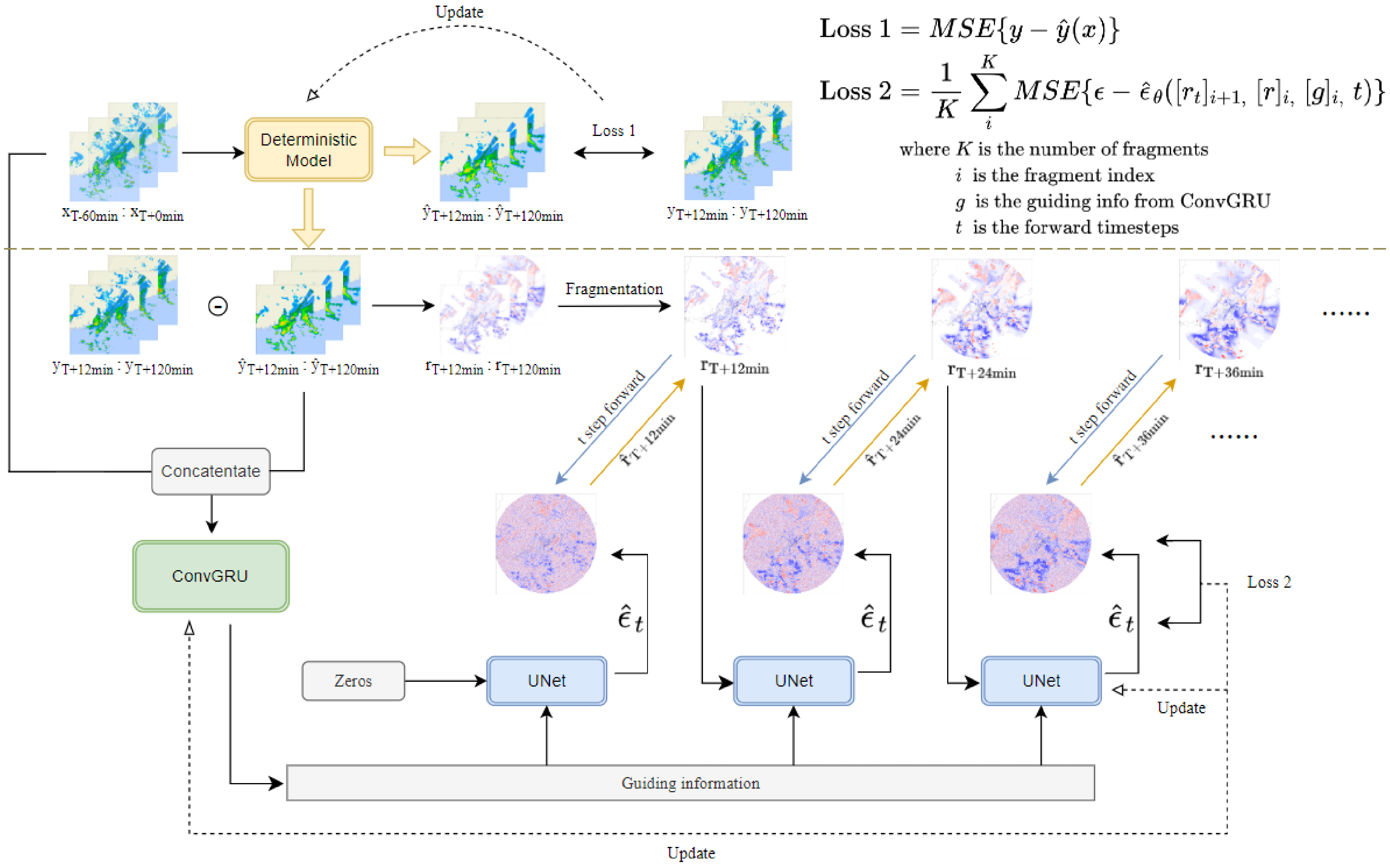
First train the deterministic model

Once its training is completed, freeze its network parameters and proceed to train the residual components

$$\text{Loss 1} = MSE\{y - \hat{y}(x)\}$$

$$\text{Loss 2} = \frac{1}{K}\sum_{i}^{K} MSE\{\epsilon - \hat{\epsilon}_\theta([r_t]_{i+1},\ [r]_i,\ [g]_i,\ t)\}$$

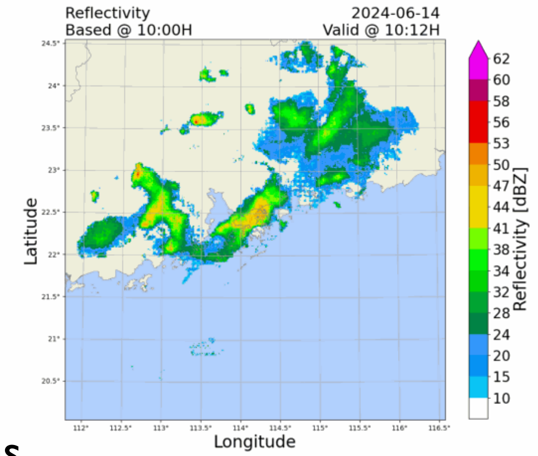
where $K$ is the number of fragments
$i$ is the fragment index
$g$ is the guiding info from ConvGRU
$t$ is the forward timesteps



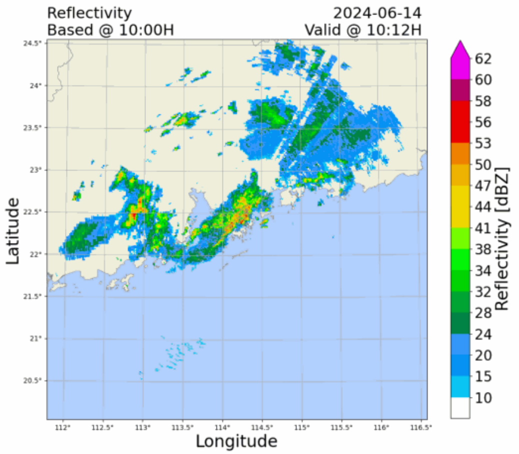
Amber and Red Rainstorms

2024-06-14

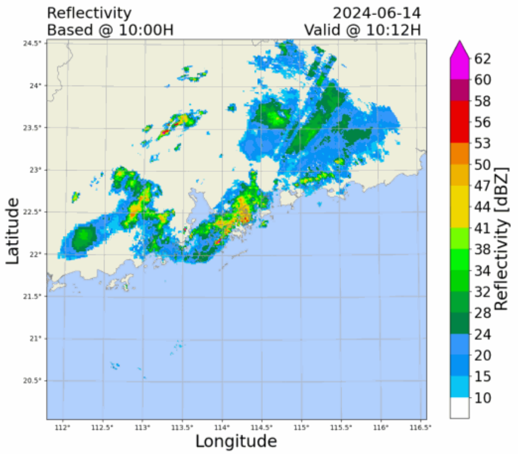
## ResConvLSTM-GAN



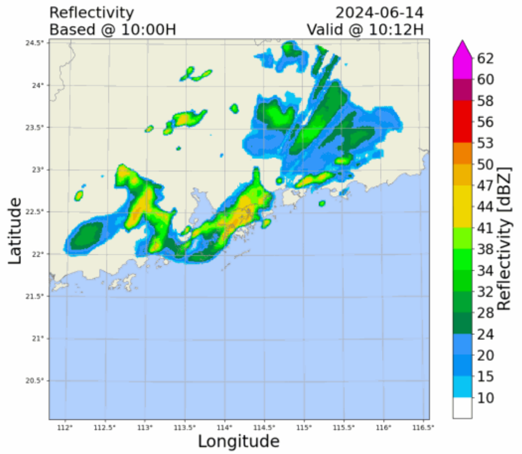

## Target



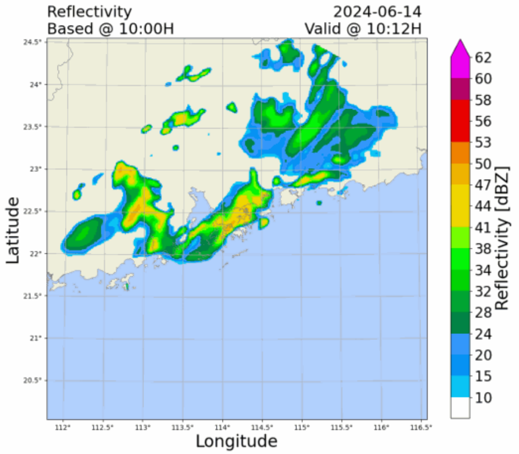

## Diffusion



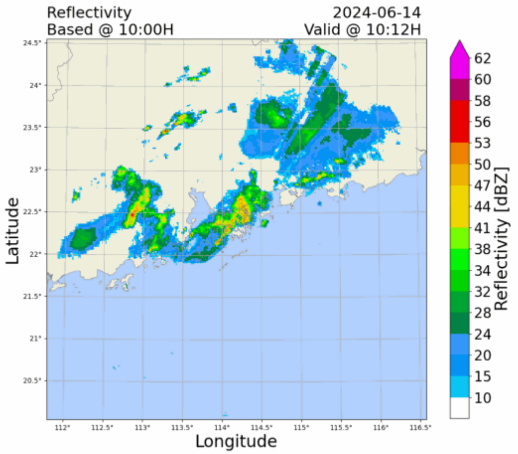

## VMRNN




## Earthformer




## Diffusion_RMS



香港天文台
HONG KONG OBSERVATORY

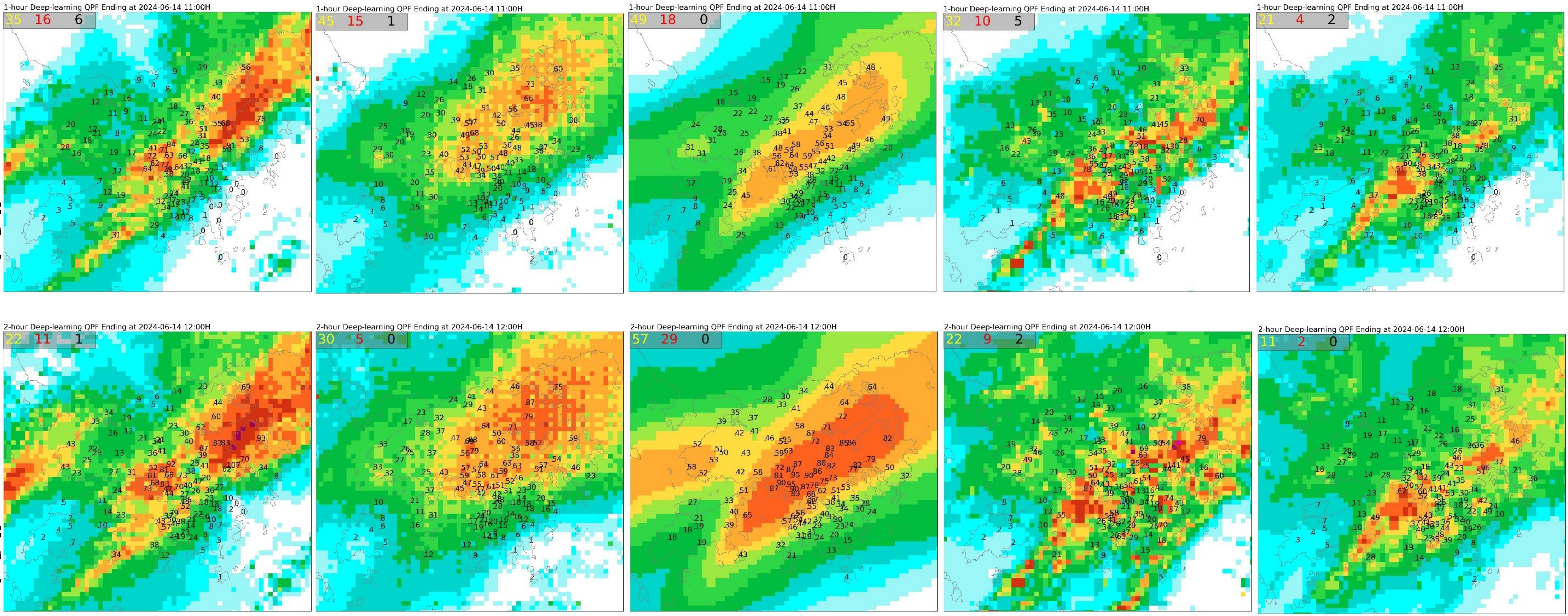
臨近預報區域專業氣象中心
RSMC for Nowcasting
WMO

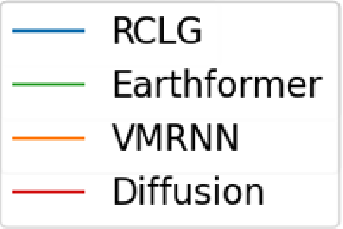
# Radar QPF

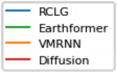
**Ground Truth** | **ResConvLSTM-GAN** | **Earthformer** | **Diffusion** | **Diffusion-RMS**

1-hr accumulated

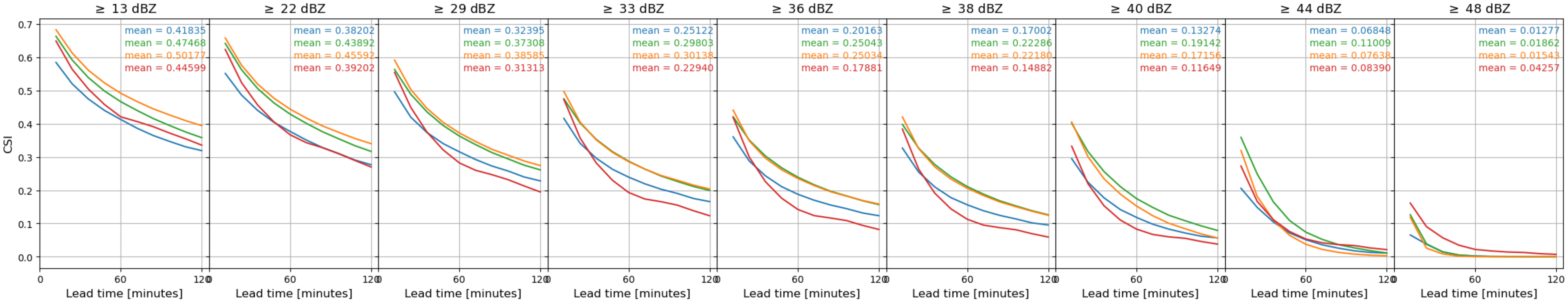
2-hr accumulated





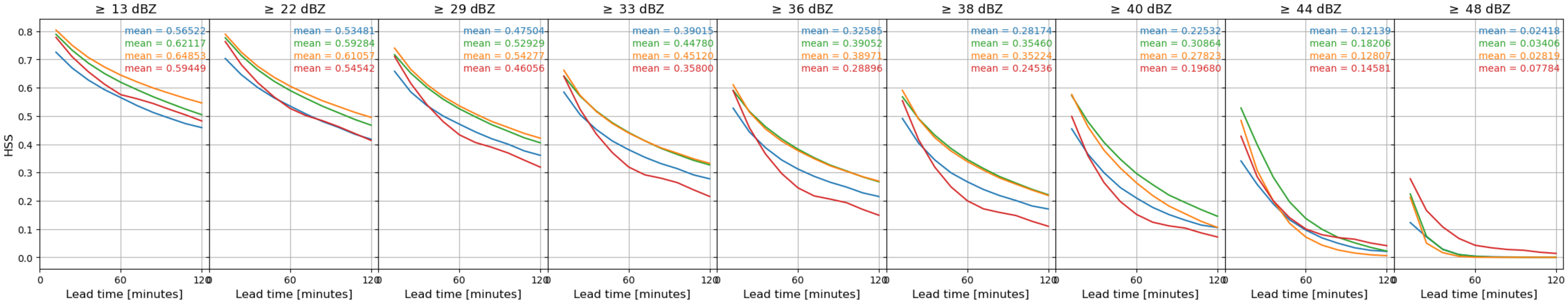

# CSI



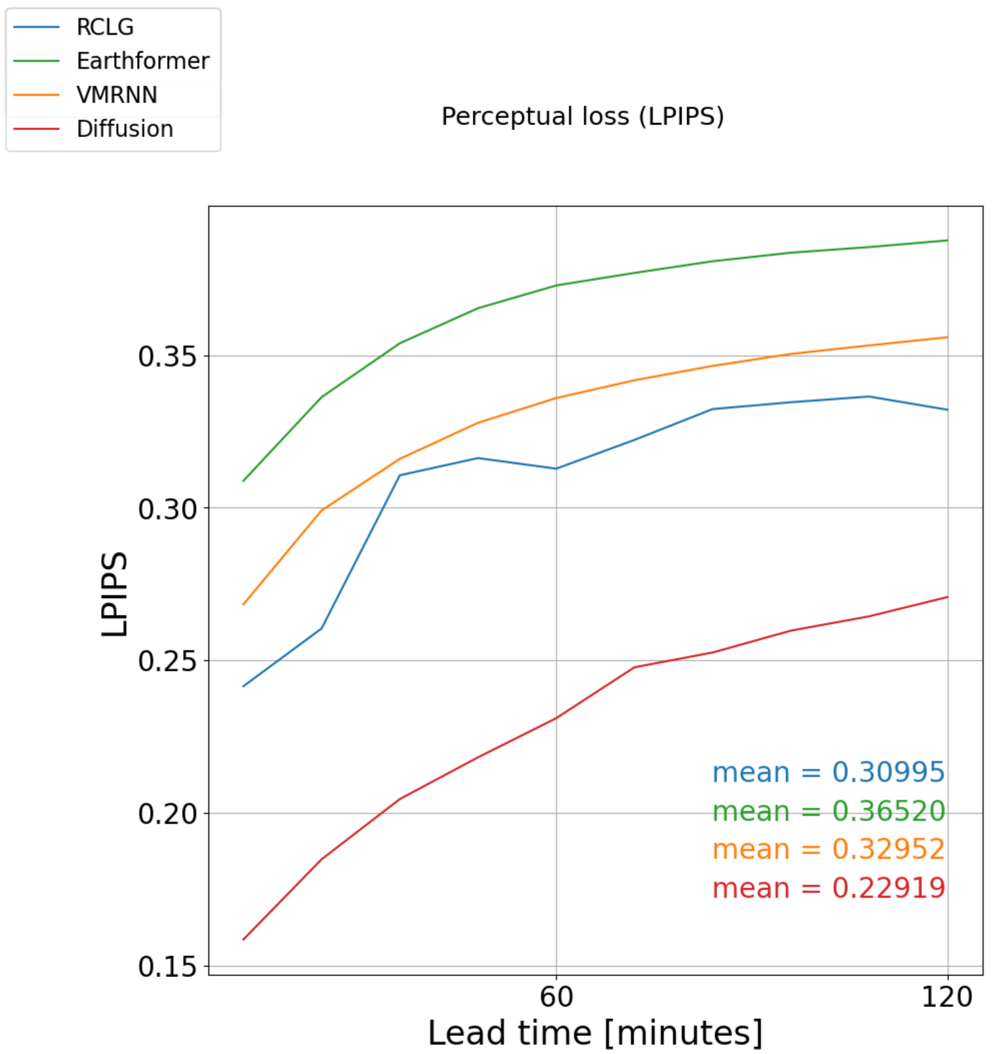


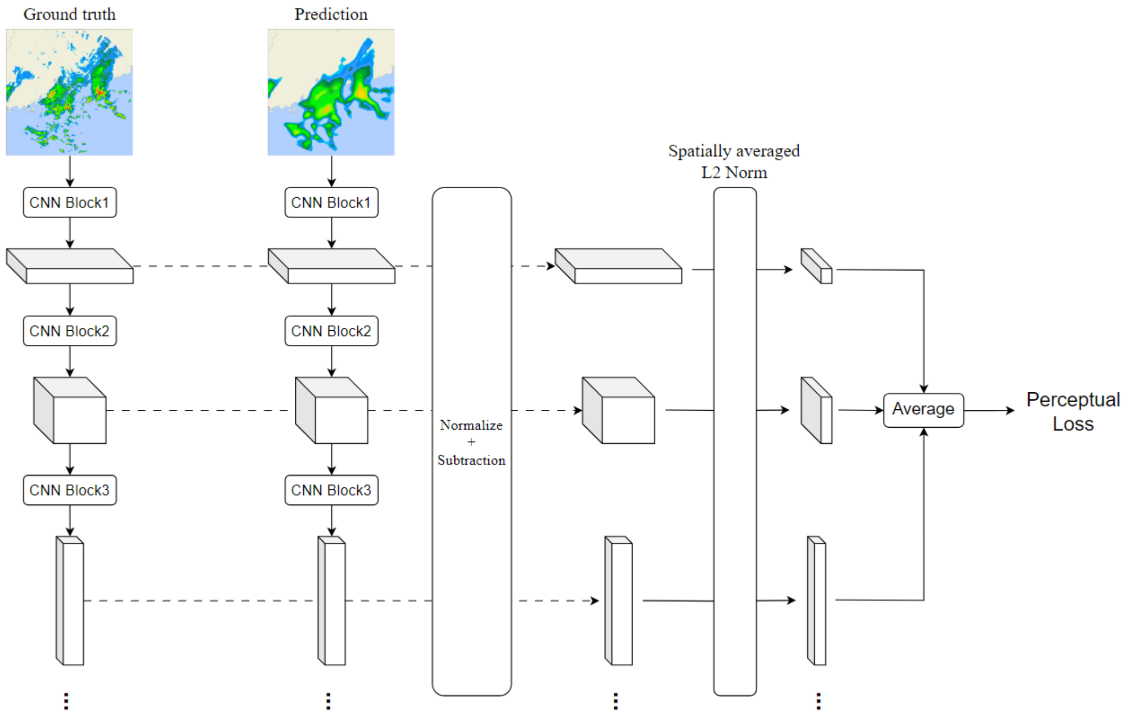
# HSS

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



# Verification (Jan 2022 – Dec 2023)

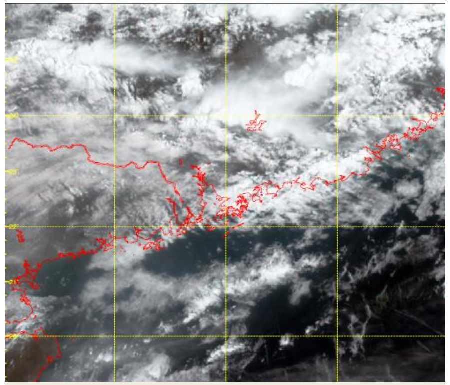



## Learned Perceptual Patch Similarity (LPIPS)

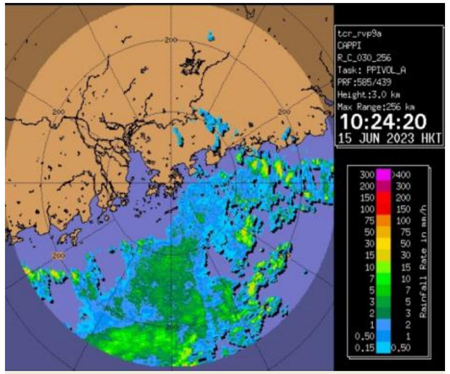
- a.k.a. Perceptual Loss
- a neural network (NN) based metric to match human perception ➔ lower the better
- measure the difference of activations in a pre-defined NN between 2 images
- pre-defined NN: AlexNet

# UNet (Multistep) Framework on Lightning Initiation and Nowcast



Satellite

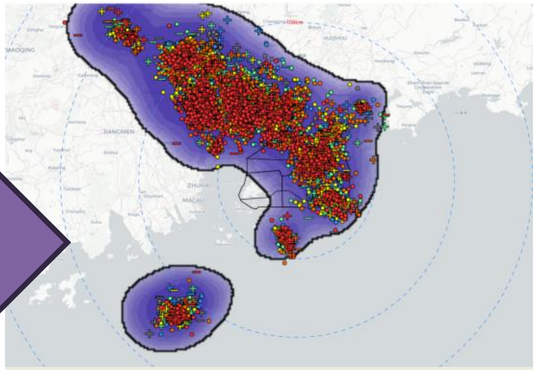
Himawari-8/9
IR bands: B13 and B15



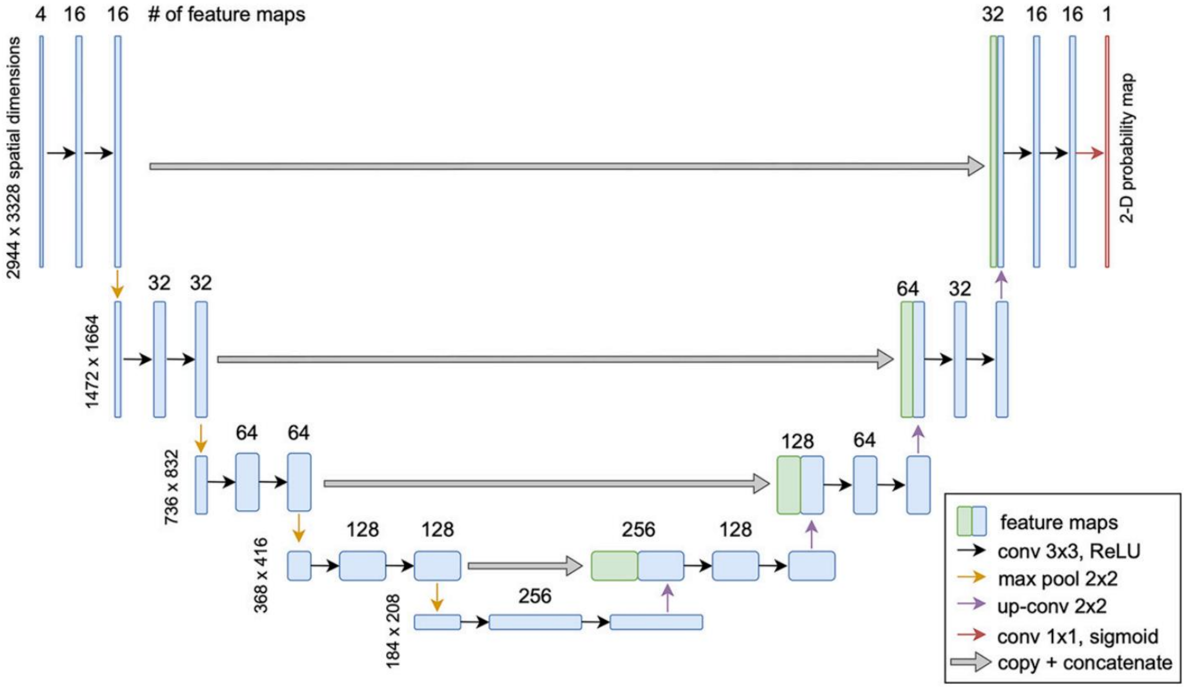

Radar

3km CAPPI REF, TOP, VIL

Predict



Lightning probability



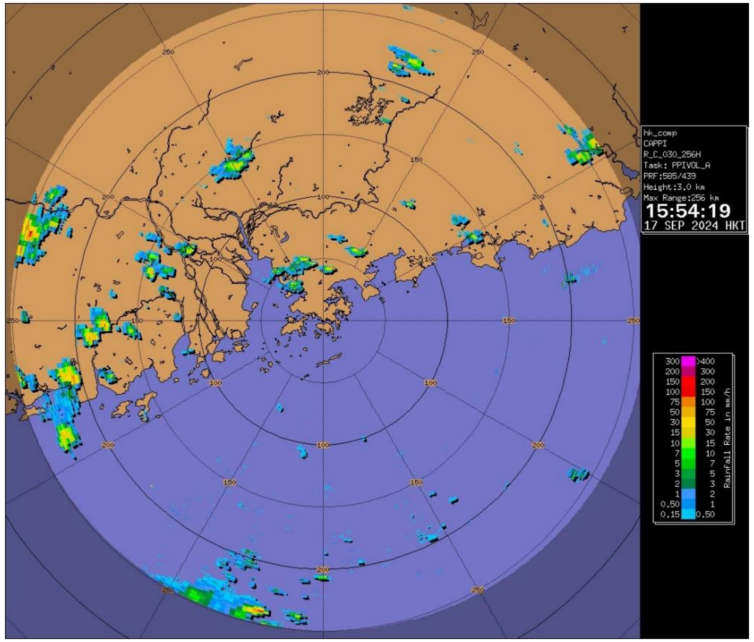

T – Δt min | T | T + Δt min

Input Data

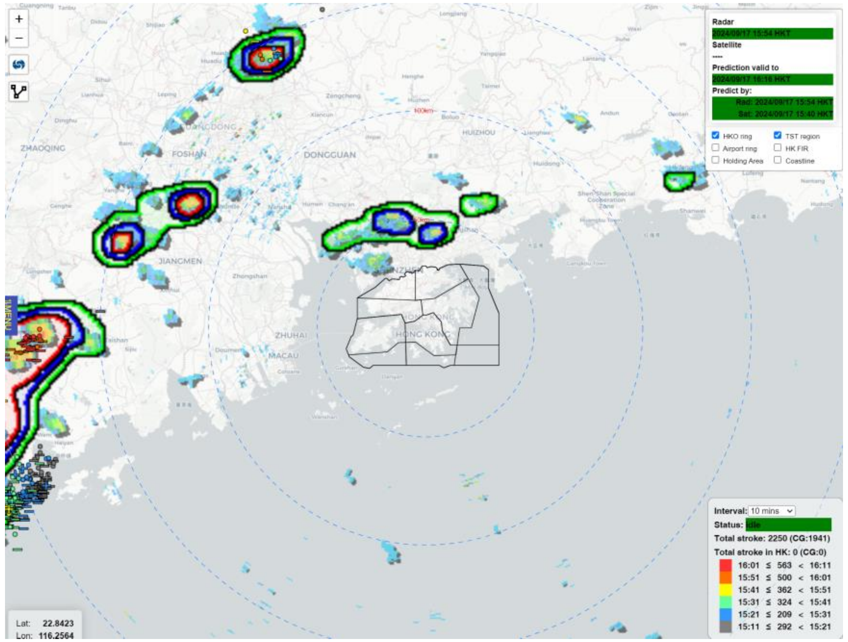
Prediction period

- Enhanced with adding stacked autoregressive framework to learn the temporal evolution of lightning activities and input radar features

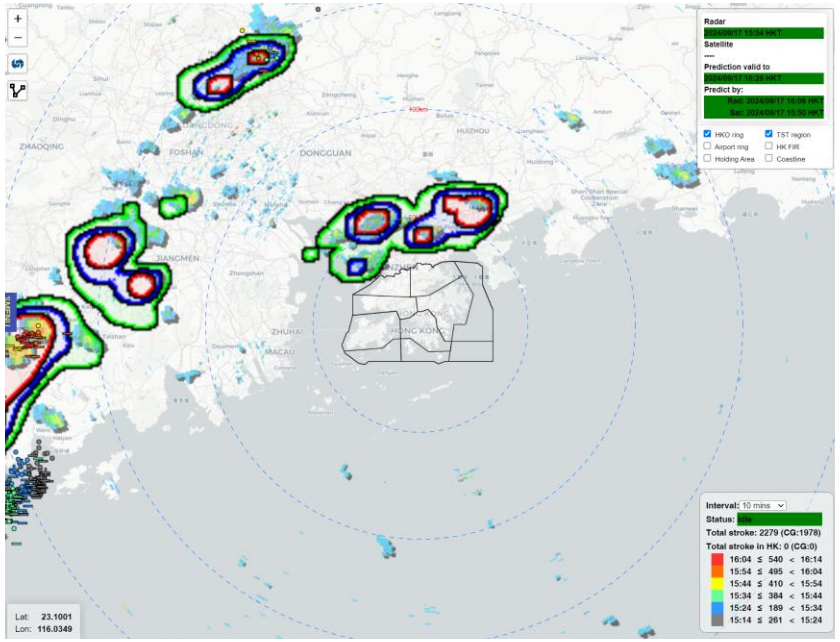
香港天文台
HONG KONG OBSERVATORY

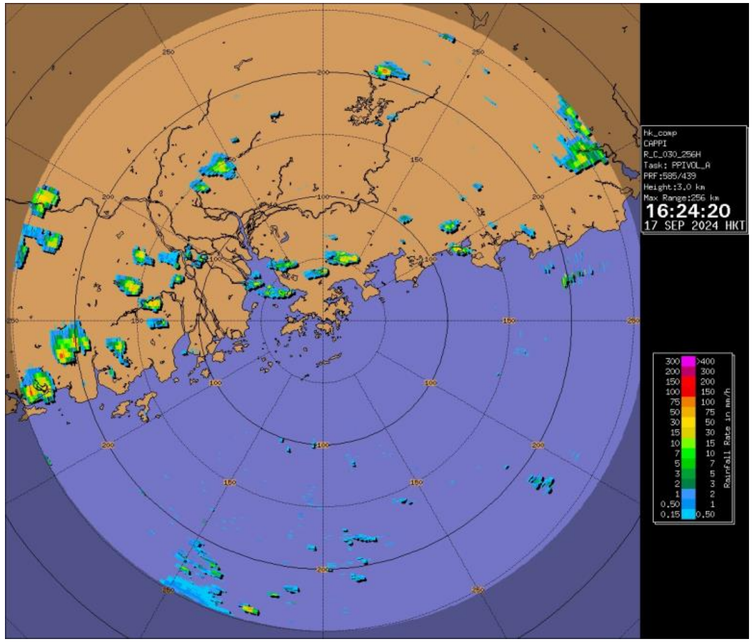


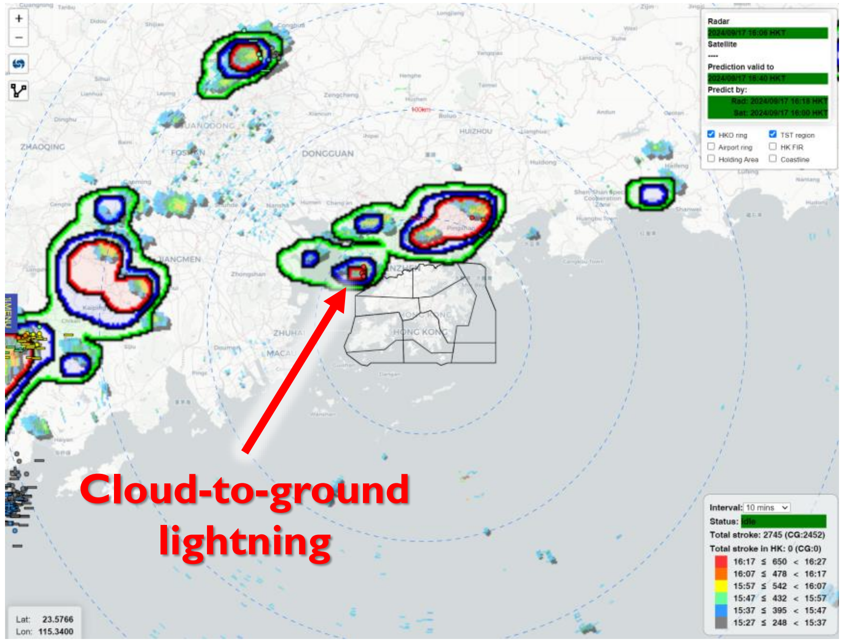

Lightning probability based on radar @1554H valid to 1616H




Lightning probability based on radar @1554H valid to 1626H



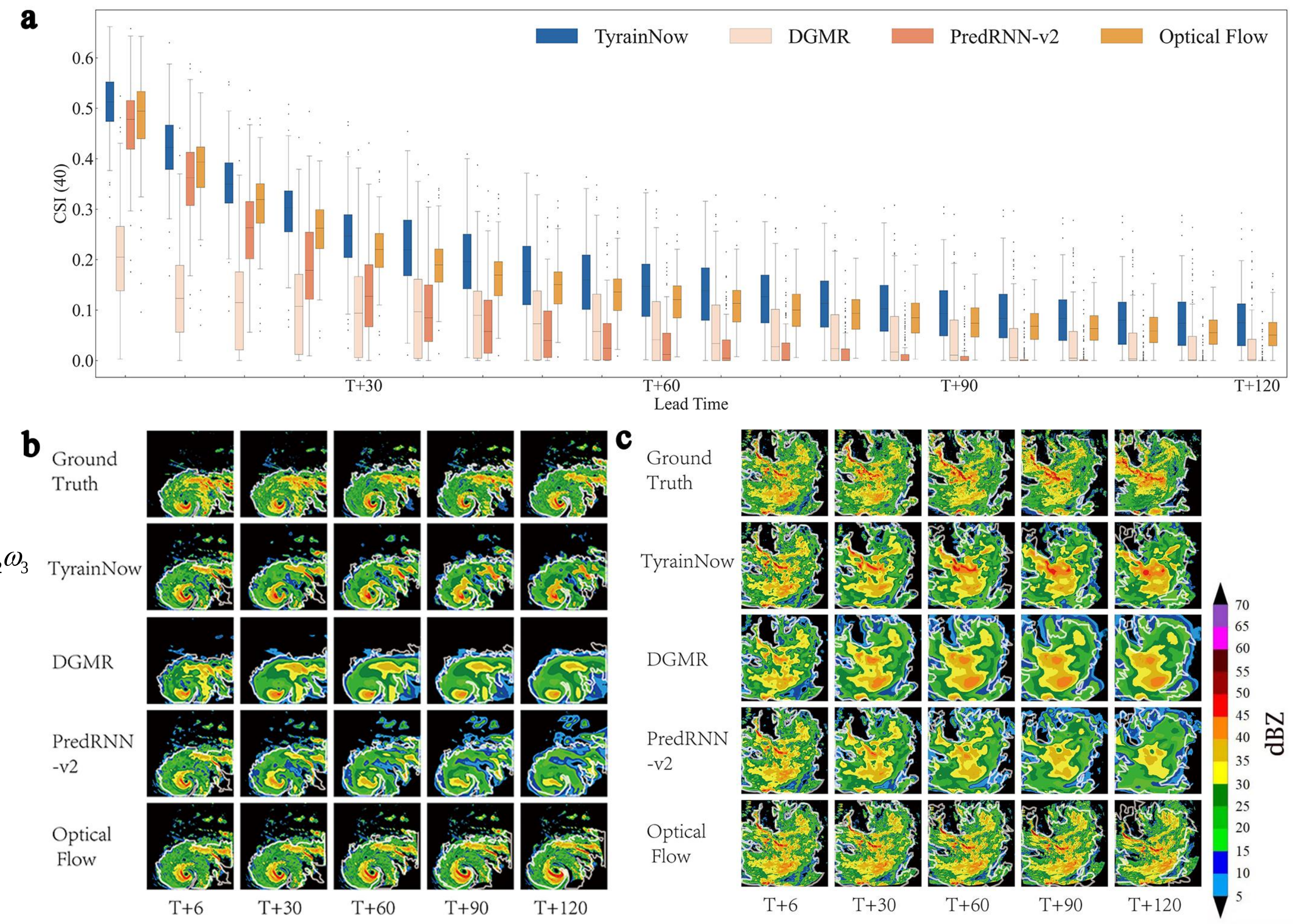



# AI in Tropical Cyclone Analysis and Nowcast (1)

## TC rainfall nowcast (TyrainNow)

Enhanced UNET model framework with spatial difference (SD) and temporal difference (TD) in the loss function together with structural similarity (SSIM) and learnable weight coefficients

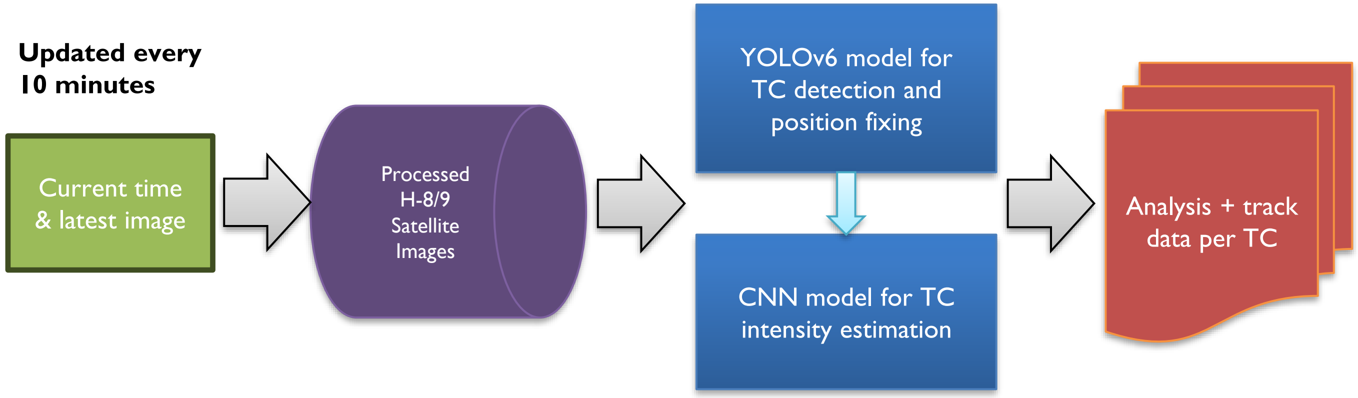
$$L(\mathbf{W},\omega_1,\omega_2,\omega_3)=\frac{1}{2\omega_1^2}\mathrm{SSIM}(\mathbf{W})+\frac{1}{2\omega_2^2}\mathrm{SD}(\mathbf{W})+\frac{1}{2\omega_3^2}\mathrm{TD}(\mathbf{W})+\log\omega_1\omega_2\omega_3$$

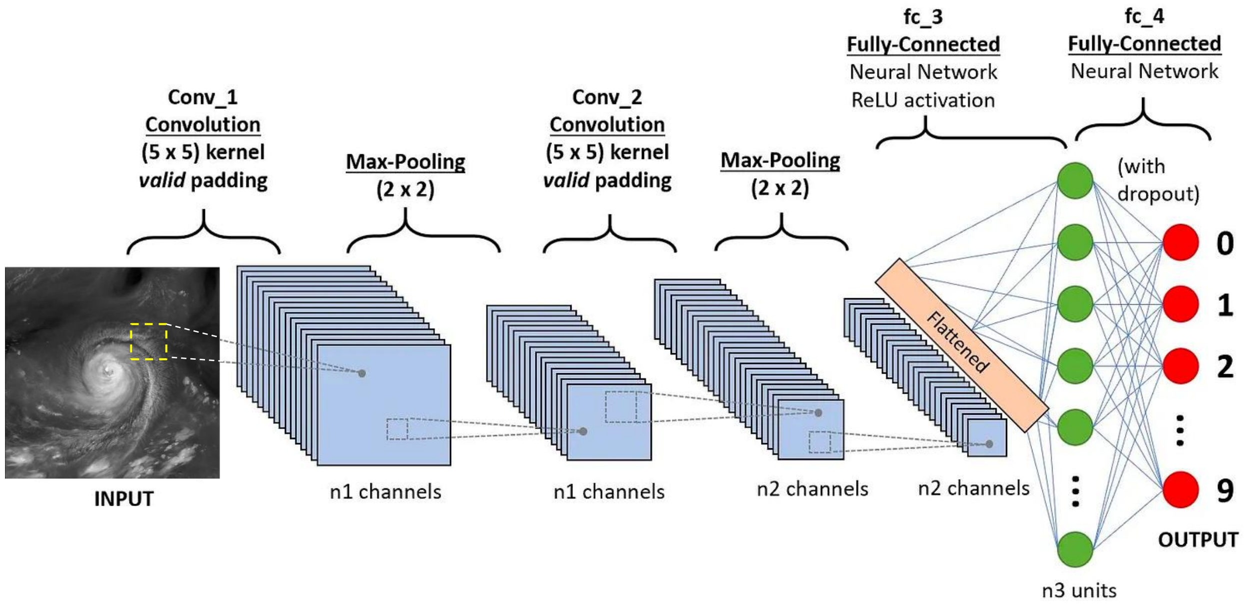


HONG KONG OBSERVATORY 香港天文台

# AI in Tropical Cyclone Analysis and Nowcast (2)

- **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
    - You Only Look Once (YOLO) v6 small object detection model
    - Ensemble approach
  - TC intensity estimation
    - Convolutional neural network (CNN) with 13 layers
  - Visualisation web-based platform



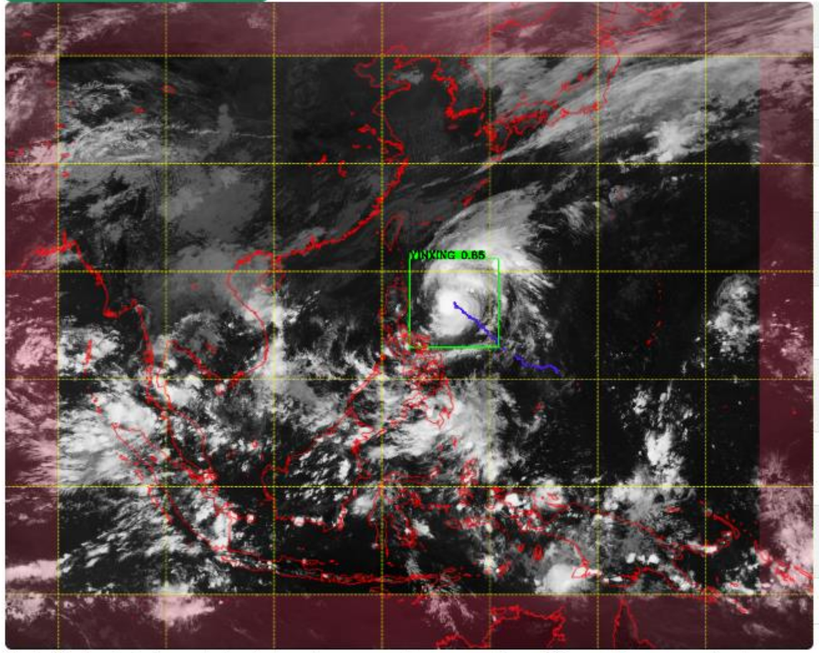



# AI-STORMVIS for H9 and FY4B

**AI-STORMVIS**
— One-Stop SatFix, track and intensity estimation.

Real-Time | Satellite + Objective Intensity | Past TC Verification | Optical Flow Motion | User Guide | Docs

Cur image: 2024.11.05/06:10Z | ImageTab | Toggle BBoxes | Lat: 40.14, Lon: 159.17



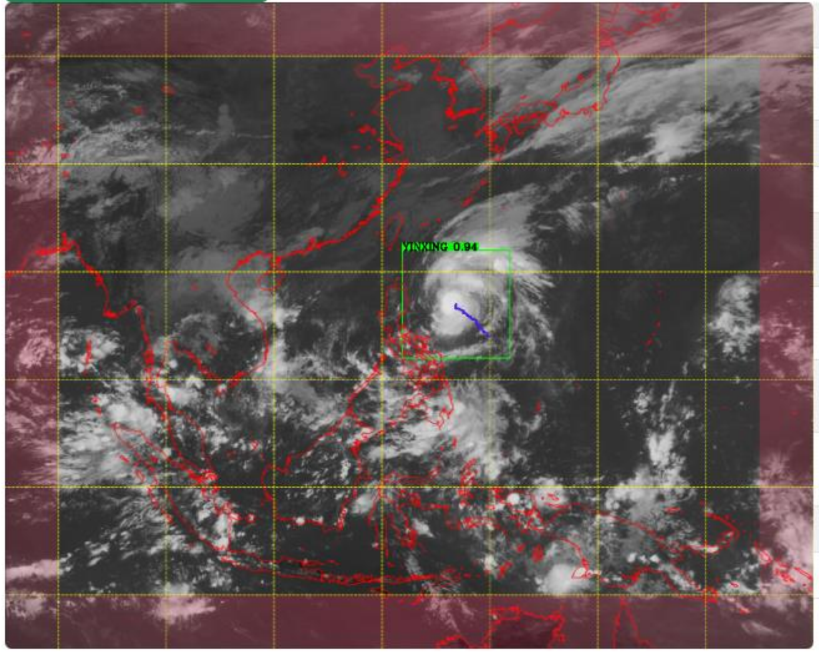

| # | 2435 YINXING |
|---|---|
| Pos | 17.12°N 126.64°E |
| CI | 3.2 (45 knots) |
| Rad | 4.1° |
| T-1 mvt | 21.0 km/h 11 kt NW (319) |
| T-3 mvt | 17.3 km/h 9 kt NW (312) |
| T-6 mvt | 24.9 km/h 13 kt NW (311) |
| T-12 mvt | 24.6 km/h 13 kt NW (308) |
| T-24 mvt | 29.3 km/h 15 kt NW (313) |
| Since | 11.03/09:20Z |

Note: Outputs in the red-shaded region may not be accurate as TCs could be recognised even without full image.

**AI-STORMVIS (FY4B Version)**
— One-Stop SatFix, track and intensity estimation.

Real-Time | Satellite + Objective Intensity | Past TC Verification | Optical Flow Motion | User Guide | Docs | Log

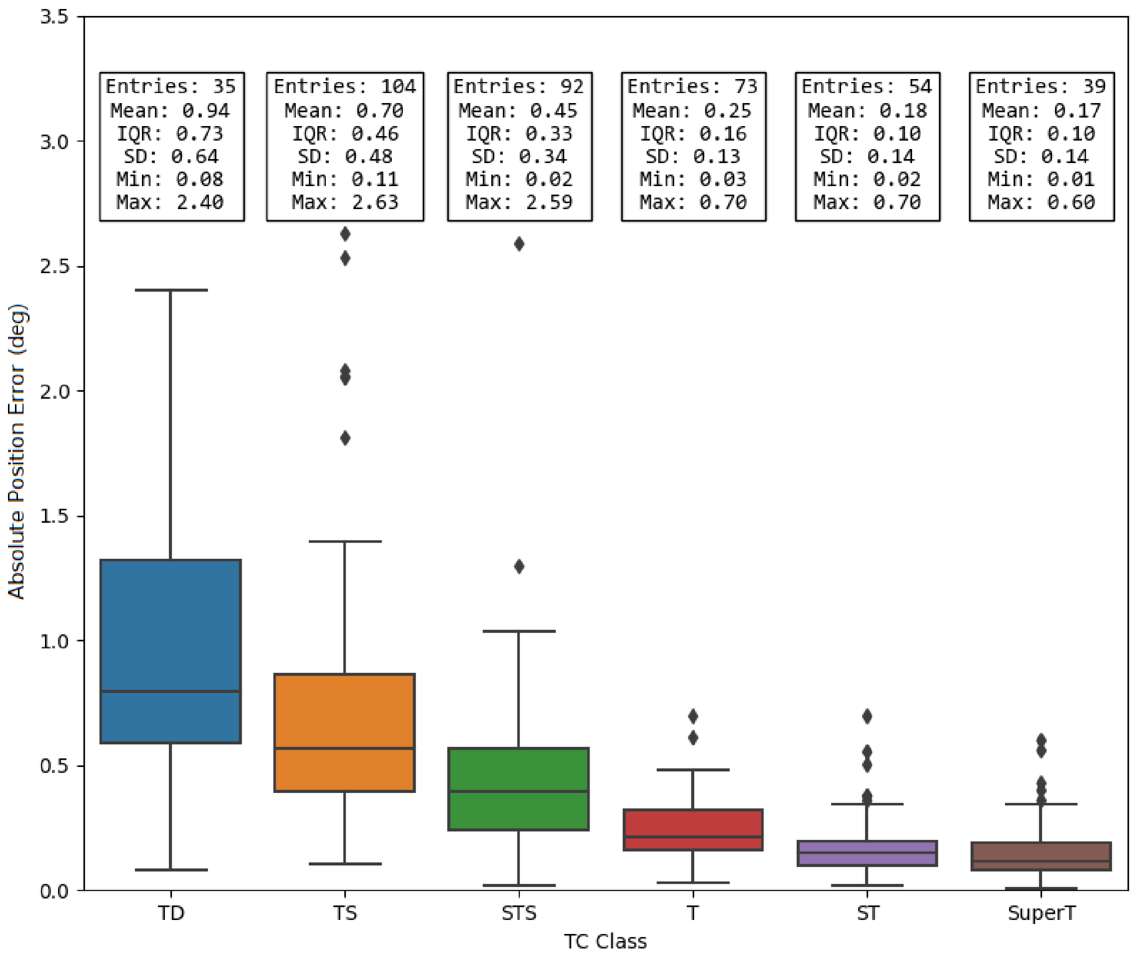
Cur image: 2024.11.05/06:00Z | ImageTab | Toggle BBoxes | Lat: 23.96, Lon: 91.46



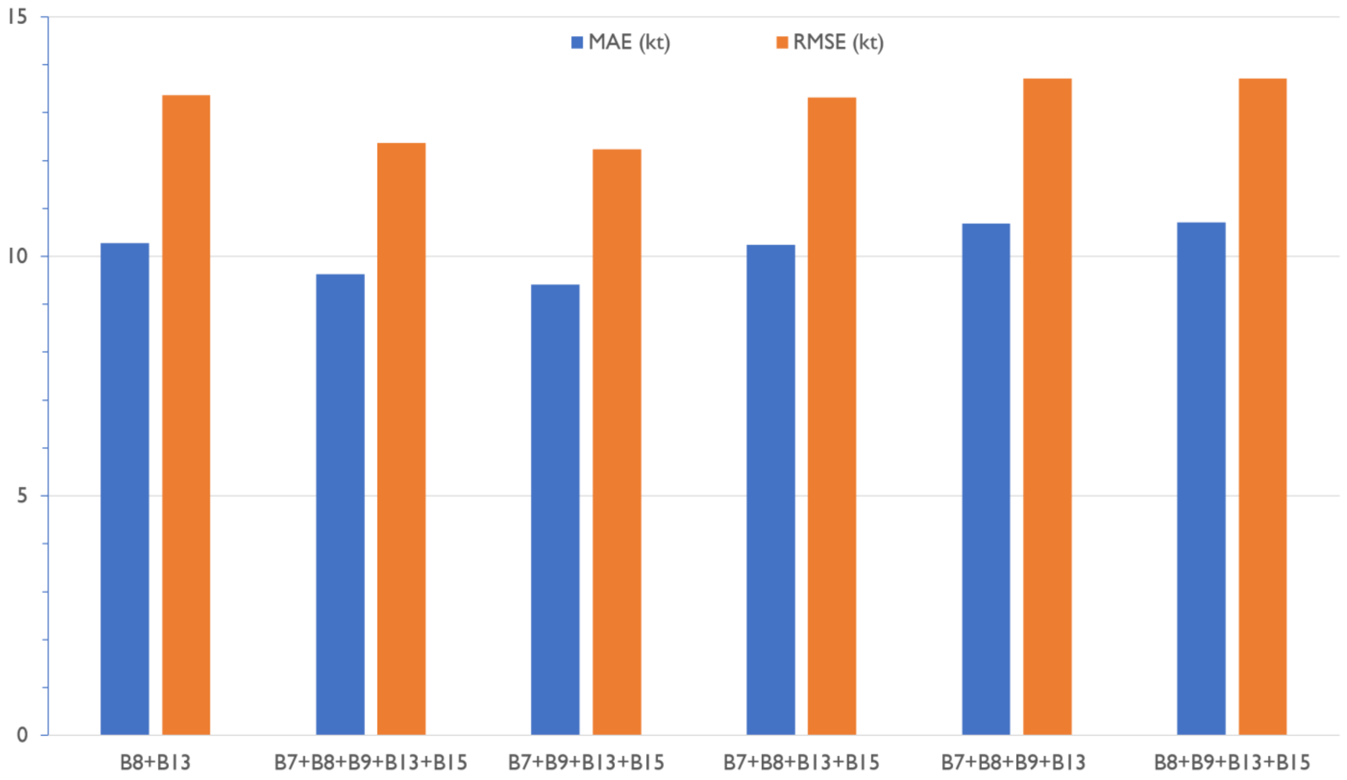

| # | 2435 YINXING |
|---|---|
| Pos | 17.02°N 126.84°E |
| CI | -999.0 |
| Rad | 5.0° |
| T-1 mvt | 42.6 km/h 23 kt N (1) |
| T-3 mvt | 35.2 km/h 19 kt NW (311) |
| T-6 mvt | 34.5 km/h 18 kt NW (311) |
| T-12 mvt | 0.0 km/h 0 kt E (90) |
| T-24 mvt | - |
| Since | 11.04/11:45Z |

# Verification of AI-STORMVIS

## Distribution of absolute error in TC position fix



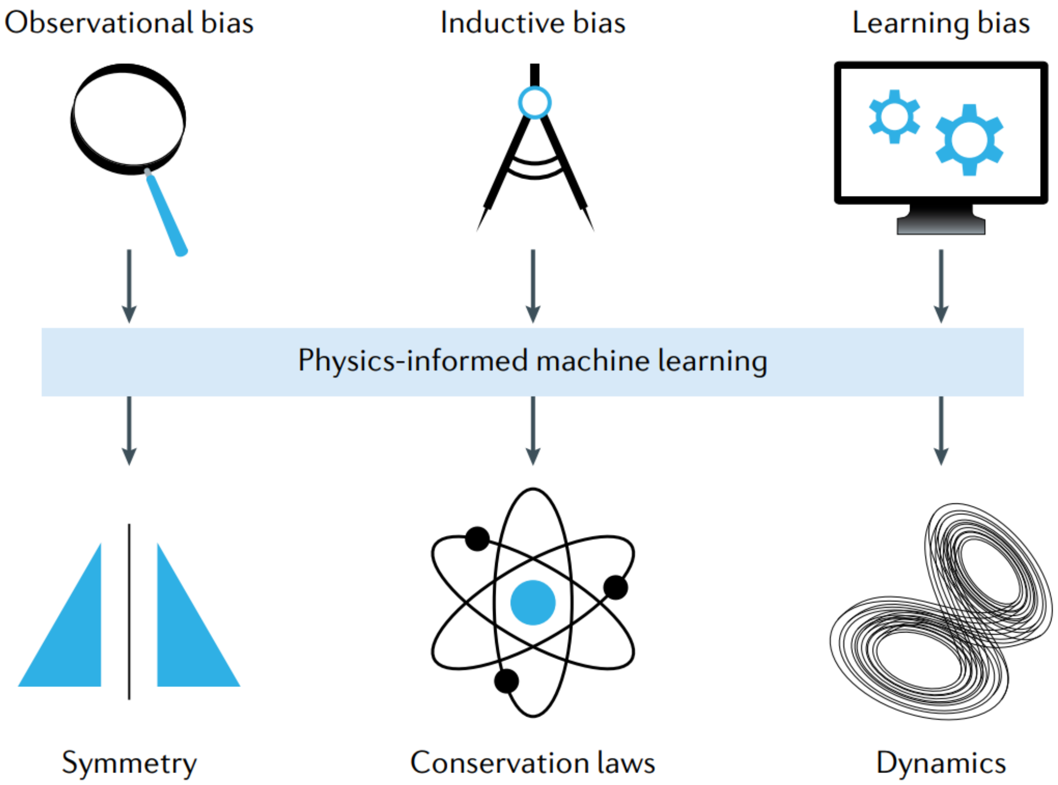

## MAE and RMSE of estimated maximum winds for AI-STORMVIS using different combination of IR channels

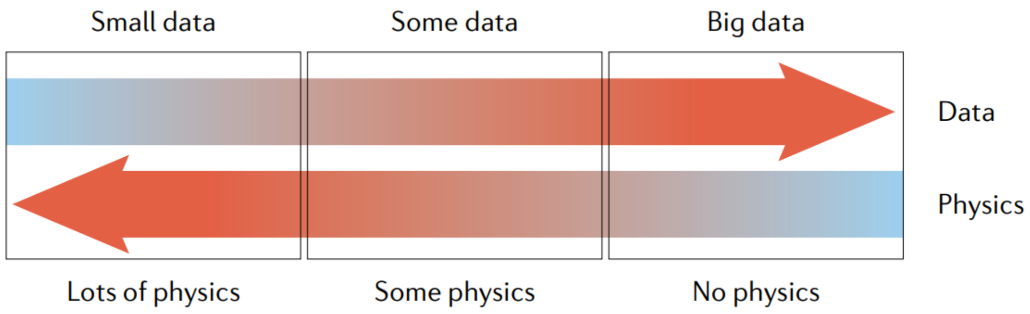


# Physics Informed Machine Learning

**Hybrid approach based on physical laws with numerical methods**

| **Sufficient data with augmentation** | **Specialized NN (e.g. Graph Neural Network)** | **Constraint in Loss Functions** |
|---|---|---|



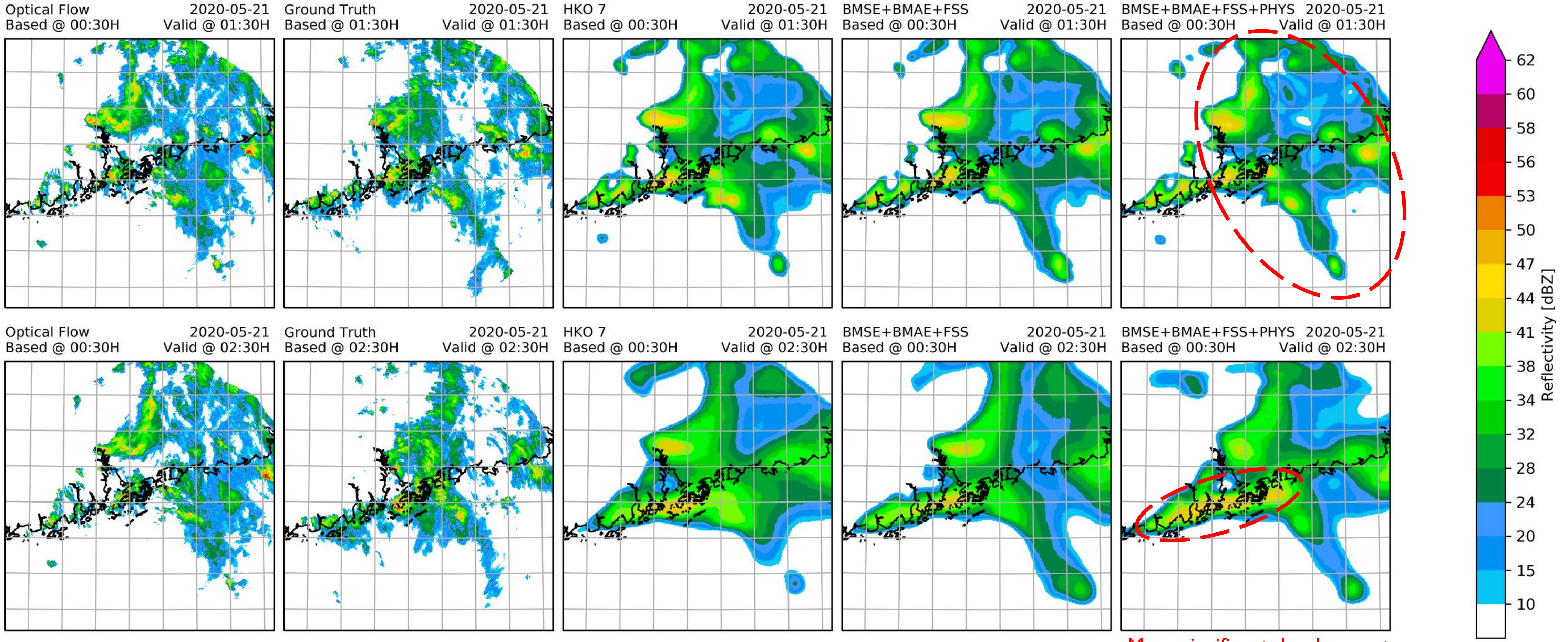




https://www.nature.com/articles/s42254-021-00314-5

| | Optical flow extrapolation | Actual | TrajGRU | TrajGRU with spatio-intensity error constraint | TrajGRU with physics and spatio-intensity error constraint |
|---|---|---|---|---|---|

1-hour nowcast

2-hour nowcast




More significant development due to physics constraint

# Thank you very much

WORLD
METEOROLOGICAL
ORGANIZATION