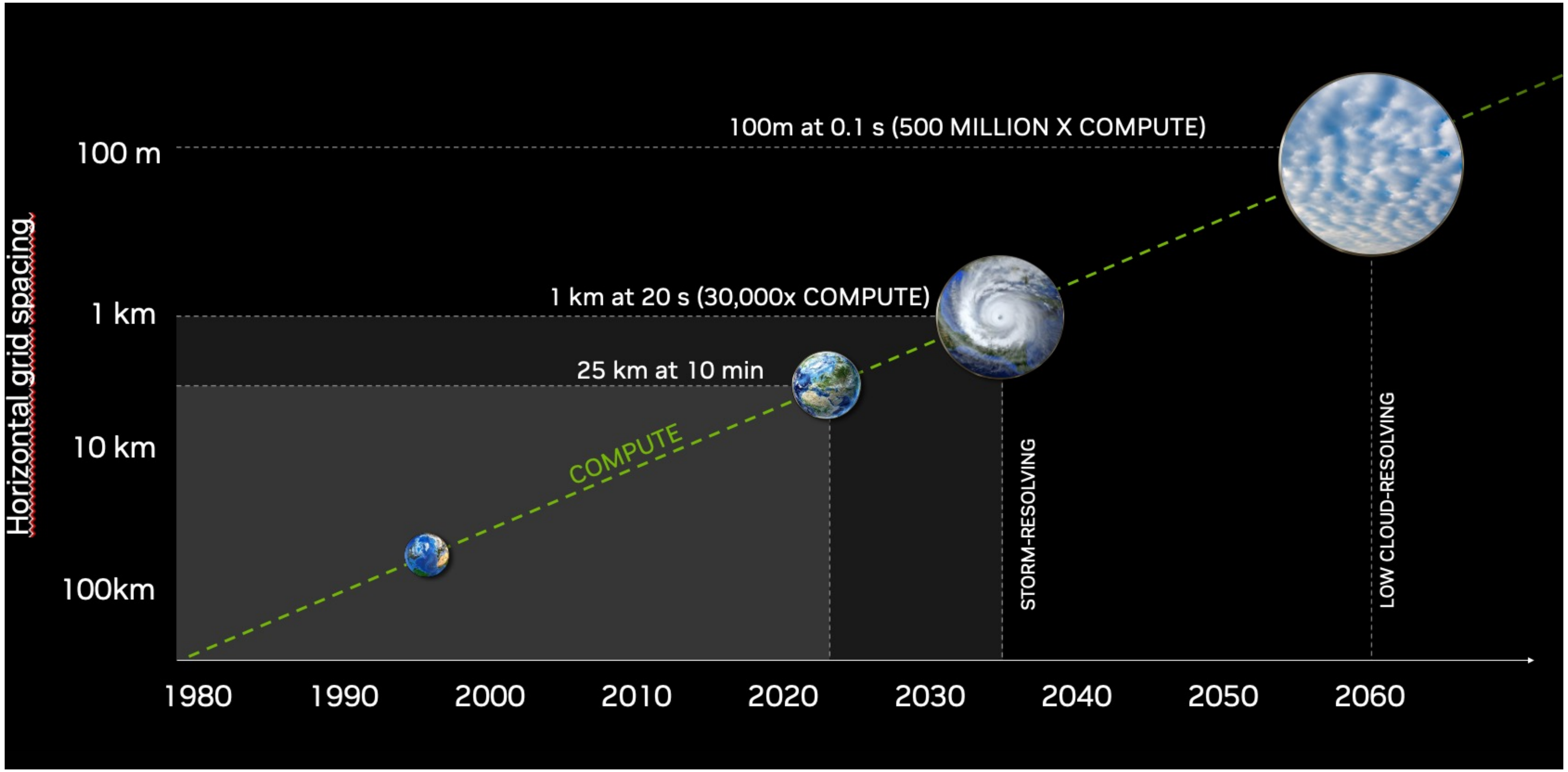
# Towards Kilometer-Scale Convection Allowing Model Emulation using Generative Diffusion Modeling

Jaideep Pathak, Senior Research Scientist | FPAW Fall Meeting / October 31 2024

# AI Could Side-Step Moore's Law With Implications for Weather Forecasts

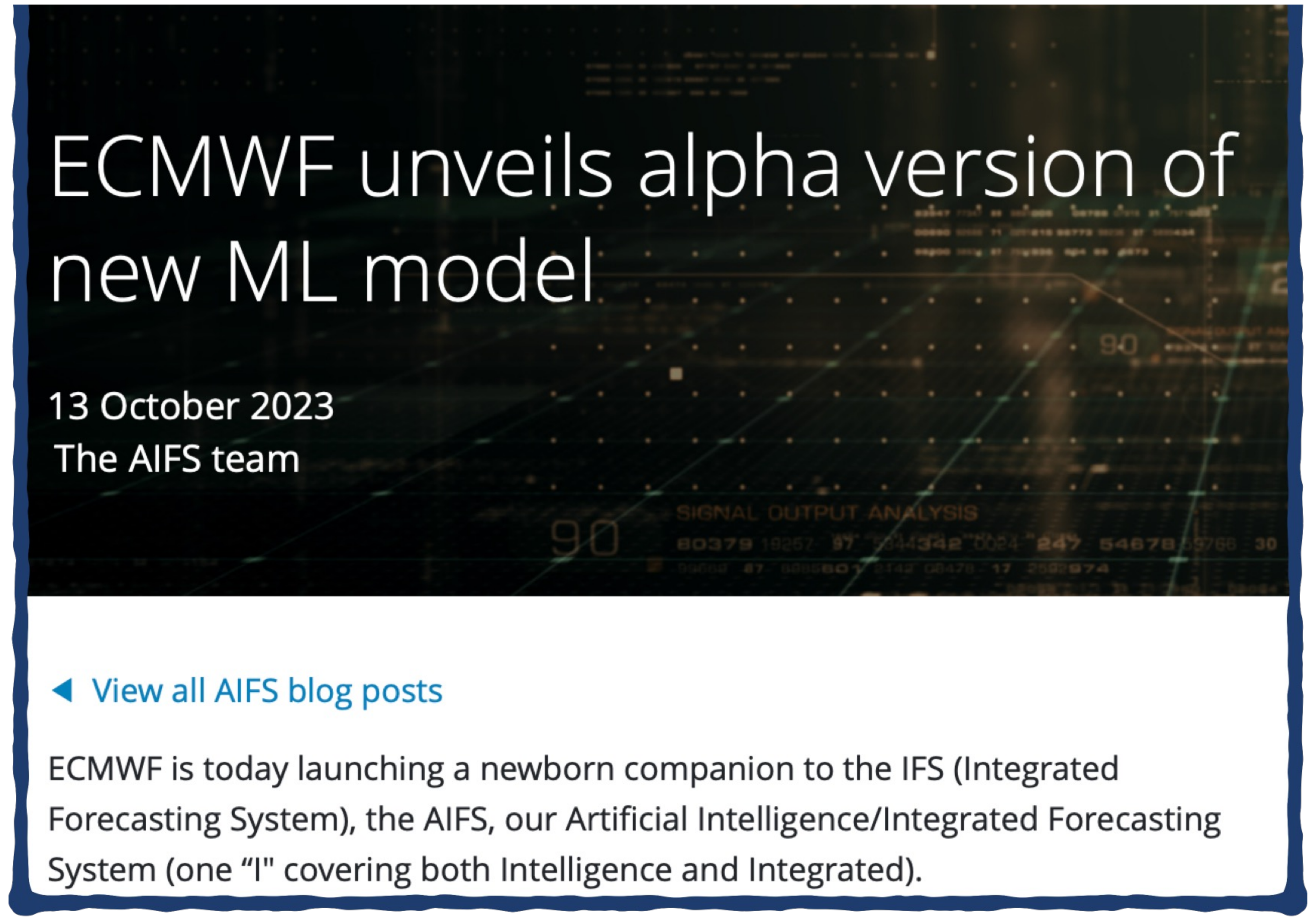
Can Breakthroughs in AI for Atmospheric Simulation Unlock Bigger Ensembles & Higher Resolution?



# 2023 was a milestone year for AI weather prediction

Global 25-km AI weather forecasting has exited its infancy

- Several AI/ML weather models are now as accurate or better than state-of-the-art numerical weather prediction at global 25km resolution.
- AI weather models offer massive speedups of over 10,000x and huge ensemble sizes



ECMWF unveils alpha version of new ML model

13 October 2023
The AIFS team

◀ View all AIFS blog posts

ECMWF is today launching a newborn companion to the IFS (Integrated Forecasting System), the AIFS, our Artificial Intelligence/Integrated Forecasting System (one “I" covering both Intelligence and Integrated).

# Diverse AI Architectures With Skill Exceeding Physics Models

## Scores for 2020 evaluated against ECMWF IFS analysis or ERA5 reanalysis

**Pressure** — 500hPa geopotential RMSE [kg$^2$/m$^2$]

| | | 1 | 3 | 5 | 7 | 10 |
|---|---|---|---|---|---|---|
| sical models | IFS HRES | 42 | 135 | 304 | 521 | 801 |
| sical models | IFS ENS Mean | 42 | 132 | 277 | 439 | 621 |
| ML / hybrid models | Pangu-Weather | 44 | 133 | 294 | 501 | 778 |
| ML / hybrid models | GraphCast | 39 | 124 | 274 | 468 | 731 |
| ML / hybrid models | FuXi | 40 | 125 | 276 | 433 | 631 |
| ML / hybrid models | SphericalCNN | 54 | 161 | 338 | 546 | 815 |
| ML / hybrid models | NeuralGCM 0.7° | 37 | 115 | 267 | 469 | 751 |
| ML / hybrid models | NeuralGCM ENS Mean | 43 | 126 | 266 | 424 | 606 |
| | Climatology | 820 | 820 | 820 | 820 | 820 |

**Temperature** — 850hPa temperature RMSE [K]

| | | 1 | 3 | 5 | 7 | 10 |
|---|---|---|---|---|---|---|
| sical models | IFS HRES | 0.62 | 1.16 | 1.82 | 2.63 | 3.63 |
| sical models | IFS ENS Mean | 0.65 | 1.11 | 1.62 | 2.17 | 2.80 |
| ML / hybrid models | Pangu-Weather | 0.62 | 1.05 | 1.71 | 2.51 | 3.54 |
| ML / hybrid models | GraphCast | 0.51 | 0.94 | 1.56 | 2.33 | 3.36 |
| ML / hybrid models | FuXi | 0.54 | 0.97 | 1.59 | 2.14 | 2.91 |
| ML / hybrid models | SphericalCNN | 0.73 | 1.18 | 1.86 | 2.64 | 3.62 |
| ML / hybrid models | NeuralGCM 0.7° | 0.54 | 0.97 | 1.58 | 2.38 | 3.42 |
| ML / hybrid models | NeuralGCM ENS Mean | 0.65 | 1.02 | 1.53 | 2.10 | 2.75 |
| | Climatology | 3.44 | 3.44 | 3.44 | 3.44 | 3.44 |

**Humidity** — 700hPa specific humidity RMSE [g/kg]

| | | 1 | 3 | 5 | 7 | 10 |
|---|---|---|---|---|---|---|
| sical models | IFS HRES | 0.55 | 0.96 | 1.27 | 1.53 | 1.81 |
| sical models | IFS ENS Mean | 0.51 | 0.84 | 1.06 | 1.22 | 1.38 |
| ML / hybrid models | Pangu-Weather | 0.53 | 0.88 | 1.19 | 1.47 | 1.79 |
| ML / hybrid models | GraphCast | 0.47 | 0.79 | 1.06 | 1.30 | 1.59 |
| ML / hybrid models | FuXi | | | | | |
| ML / hybrid models | SphericalCNN | 0.59 | 0.89 | 1.17 | 1.43 | 1.72 |
| ML / hybrid models | NeuralGCM 0.7° | 0.48 | 0.83 | 1.12 | 1.40 | 1.71 |
| ML / hybrid models | NeuralGCM ENS Mean | 0.54 | 0.81 | 1.02 | 1.19 | 1.37 |
| | Climatology | 1.59 | 1.59 | 1.59 | 1.59 | 1.59 |

**Wind Vector** — 850hPa wind vector RMSE [m/s]

| | | 1 | 3 | 5 | 7 | 10 |
|---|---|---|---|---|---|---|
| sical models | IFS HRES | 1.69 | 3.29 | 5.20 | 7.11 | 9.14 |
| sical models | IFS ENS Mean | 1.63 | 2.98 | 4.44 | 5.74 | 6.94 |
| ML / hybrid models | Pangu-Weather | 1.66 | 3.00 | 4.82 | 6.71 | 8.79 |
| ML / hybrid models | GraphCast | 1.42 | 2.76 | 4.44 | 6.22 | 8.17 |
| ML / hybrid models | FuXi | 1.47 | 2.80 | 4.49 | 5.64 | 7.02 |
| ML / hybrid models | SphericalCNN | 2.05 | 3.38 | 5.17 | 7.01 | 8.98 |
| ML / hybrid models | NeuralGCM 0.7° | 1.49 | 2.81 | 4.57 | 6.49 | 8.64 |
| ML / hybrid models | NeuralGCM ENS Mean | 1.76 | 2.88 | 4.28 | 5.59 | 6.83 |
| | Climatology | 7.89 | 7.89 | 7.89 | 7.89 | 7.89 |

Lead time [days]

−50 −20 −10 −5 −2 −1 1 2 5 10 20 50

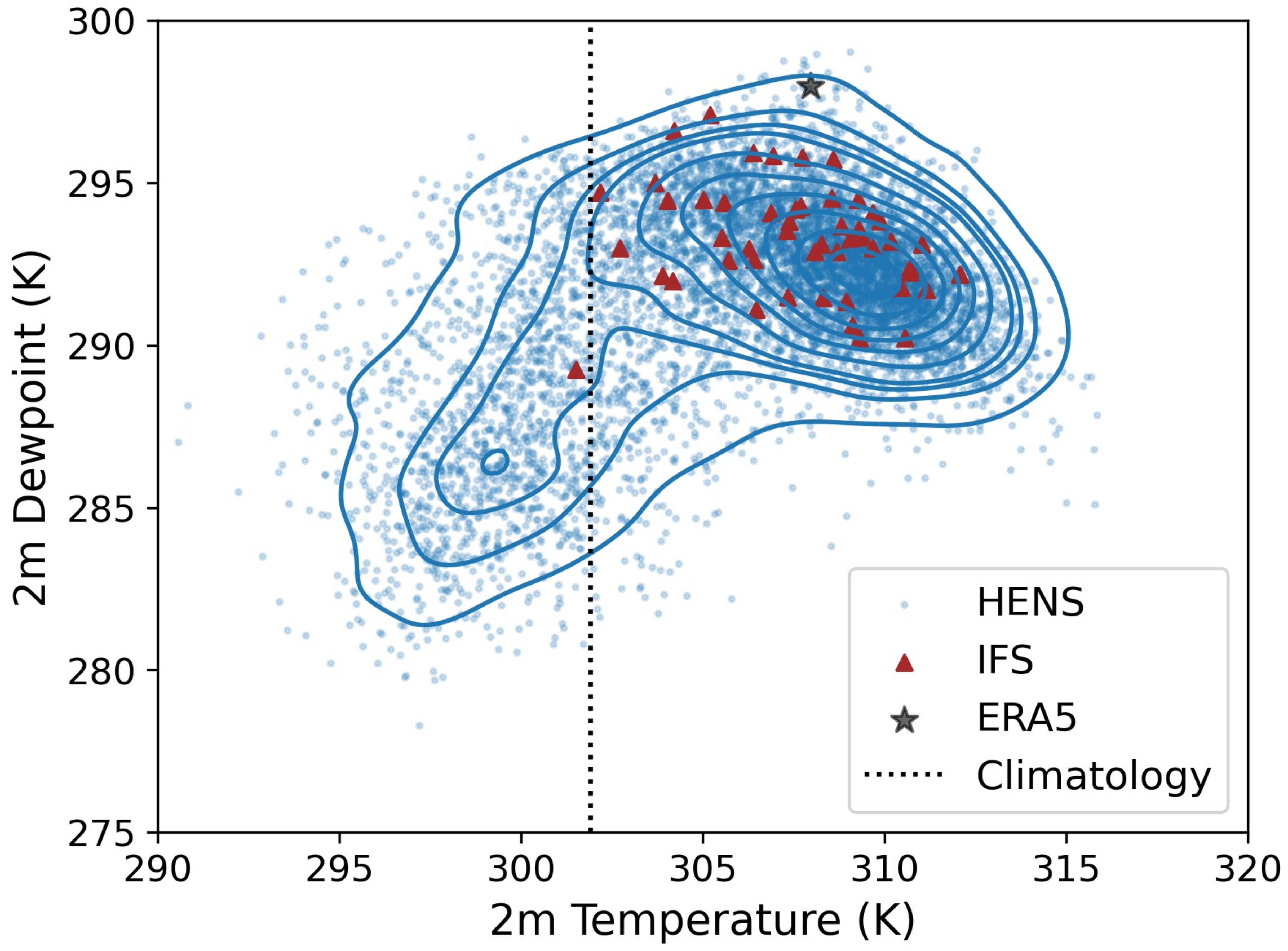
Better ⟵ % difference in RMSE vs IFS HRES ⟶ Worse

sites.research.google/weatherbench

# HENS: Huge ensemble sizes are possible at synoptic scale with AI

## Case Study with the NVIDIA FourCastNet-v2 model

- On August 23, 2023, Kansas City had an extreme heatwave, with 35°C air temperature, 56% relative humidity, and a heat index of 43°C.
- The 10-day IFS ensemble forecasts predicted warmer than average temperatures, but no members captured the combined magnitude of surface heat and humidity.
- HENS samples the tails of the forecast distribution and is able to capture the magnitude of the event.



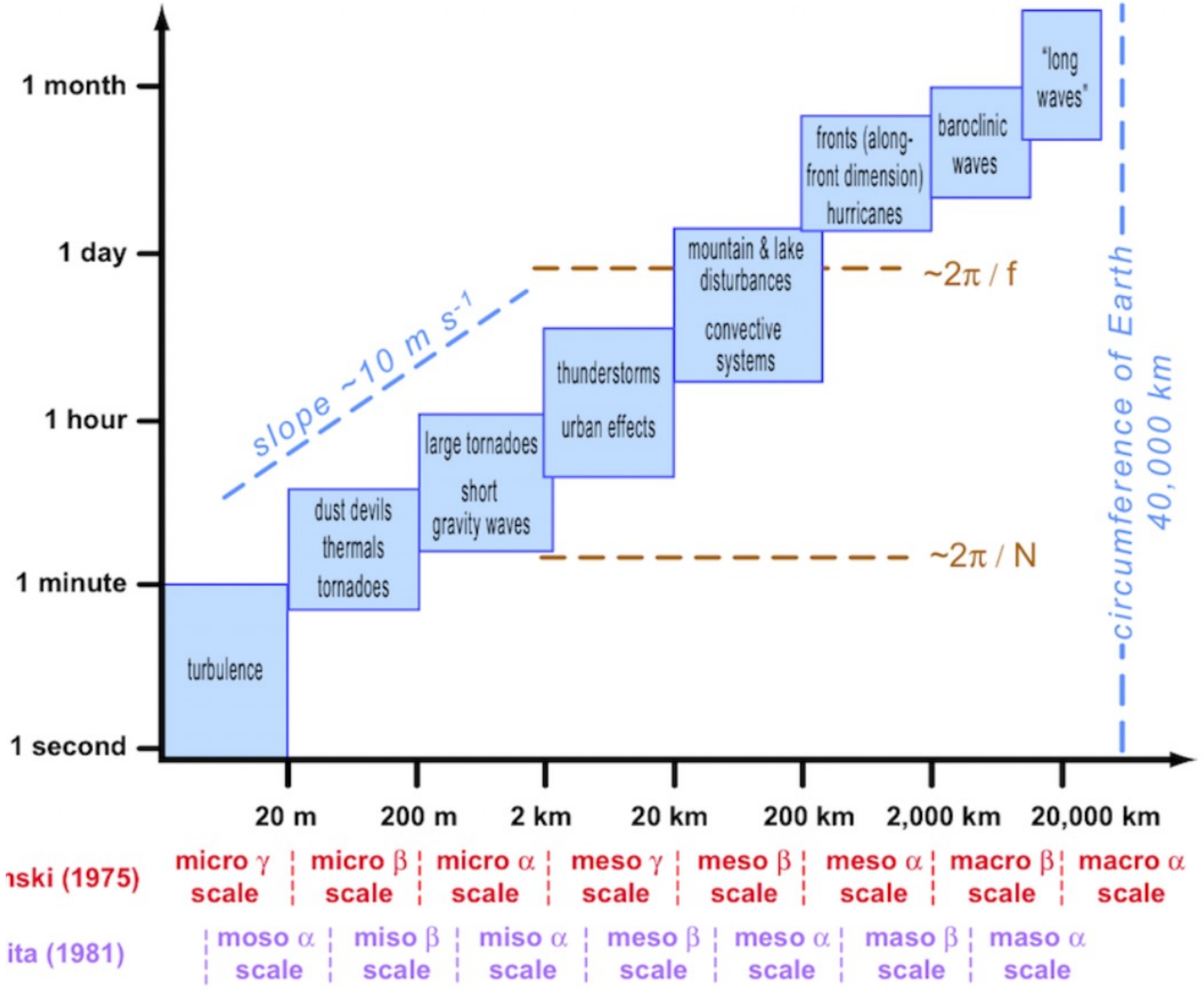

**Mahesh et al., Huge Ensembles Part I: Design of Ensemble Weather Forecasts using Spherical Fourier Neural Operators**

# Km-scale Weather Forecasting with ML

Atmospheric physics spans a large range of spatial and temporal scales

- Global weather models:
  - 10-30km resolution.
  - Negligible vertical acceleration of air, hydrostatic balance assumed.
  - Parametrized precipitation forecasts.
  - Global domain.

- Regional weather models
  - 1-5km resolution.
  - Hydrostatic balance is not assumed resulting in buoyancy and convection.
  - Explicitly modeled convective dynamics.
  - Capable of simulating thunderstorms.
  - Regional domain due to computational expense.

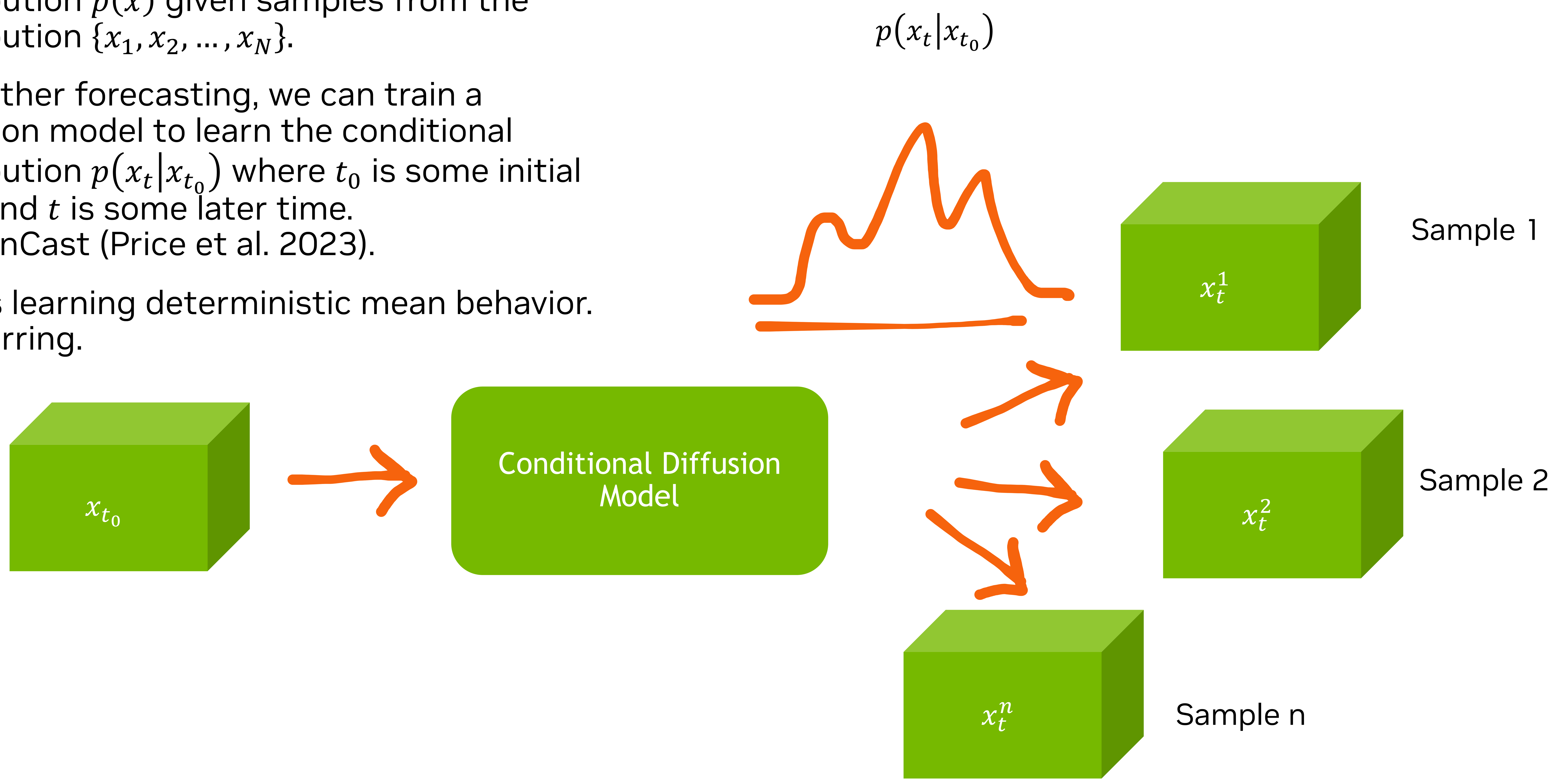



Markowski & Richardson

# Generative Diffusion Models

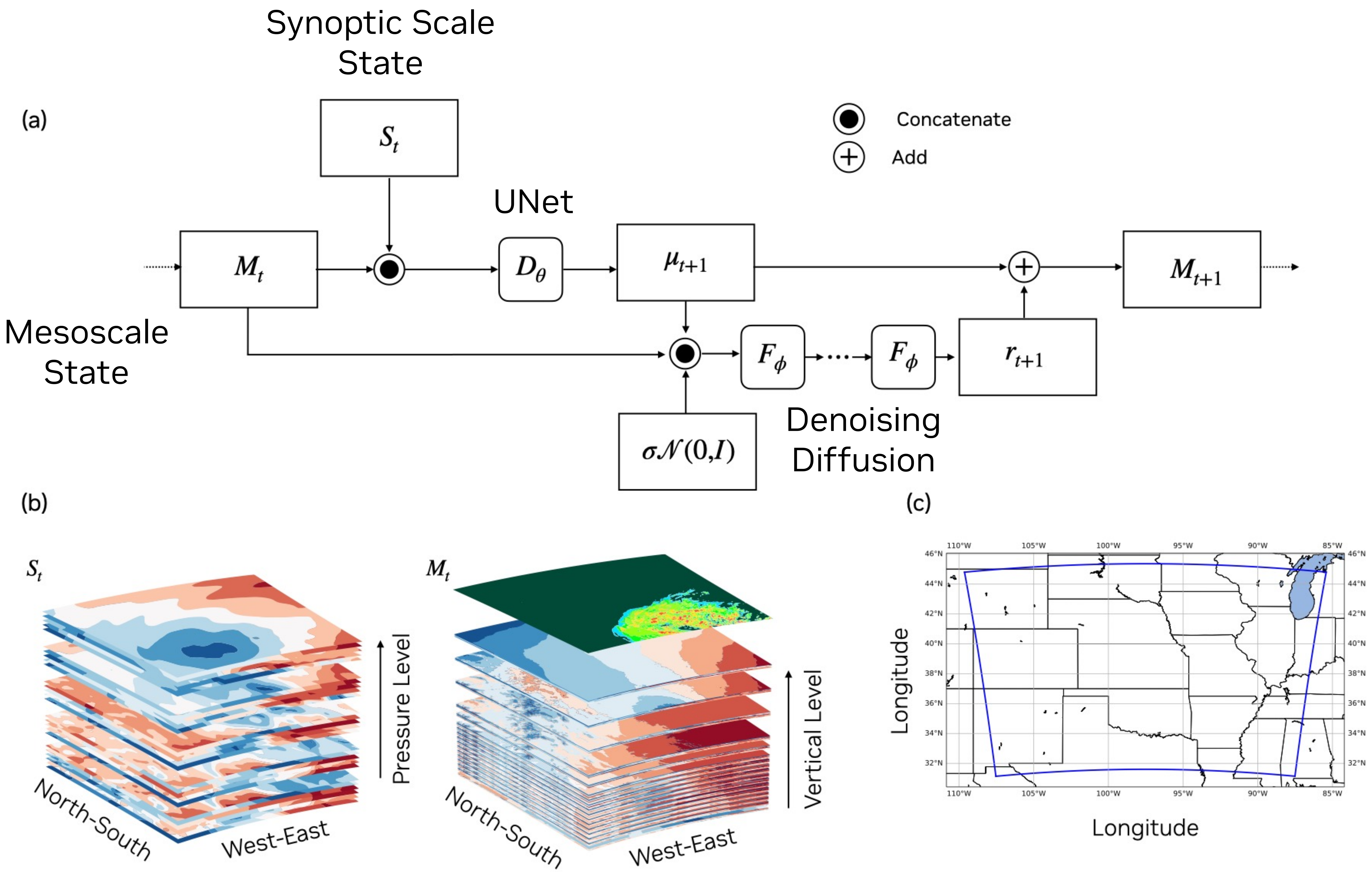
Learn a distribution p(x)

- Diffusion models allow you to learn a distribution $p(x)$ given samples from the distribution $\{x_1, x_2, \dots, x_N\}$.
- In weather forecasting, we can train a diffusion model to learn the conditional distribution $p(x_t|x_{t_0})$ where $t_0$ is some initial time and $t$ is some later time. C.f. GenCast (Price et al. 2023).
- Avoids learning deterministic mean behavior. No blurring.



# StormCast – Km-Scale Generative Convection Allowing Model

Beyond downscaling: NVIDIA's first high-resolution AI km-scale weather prediction prototype





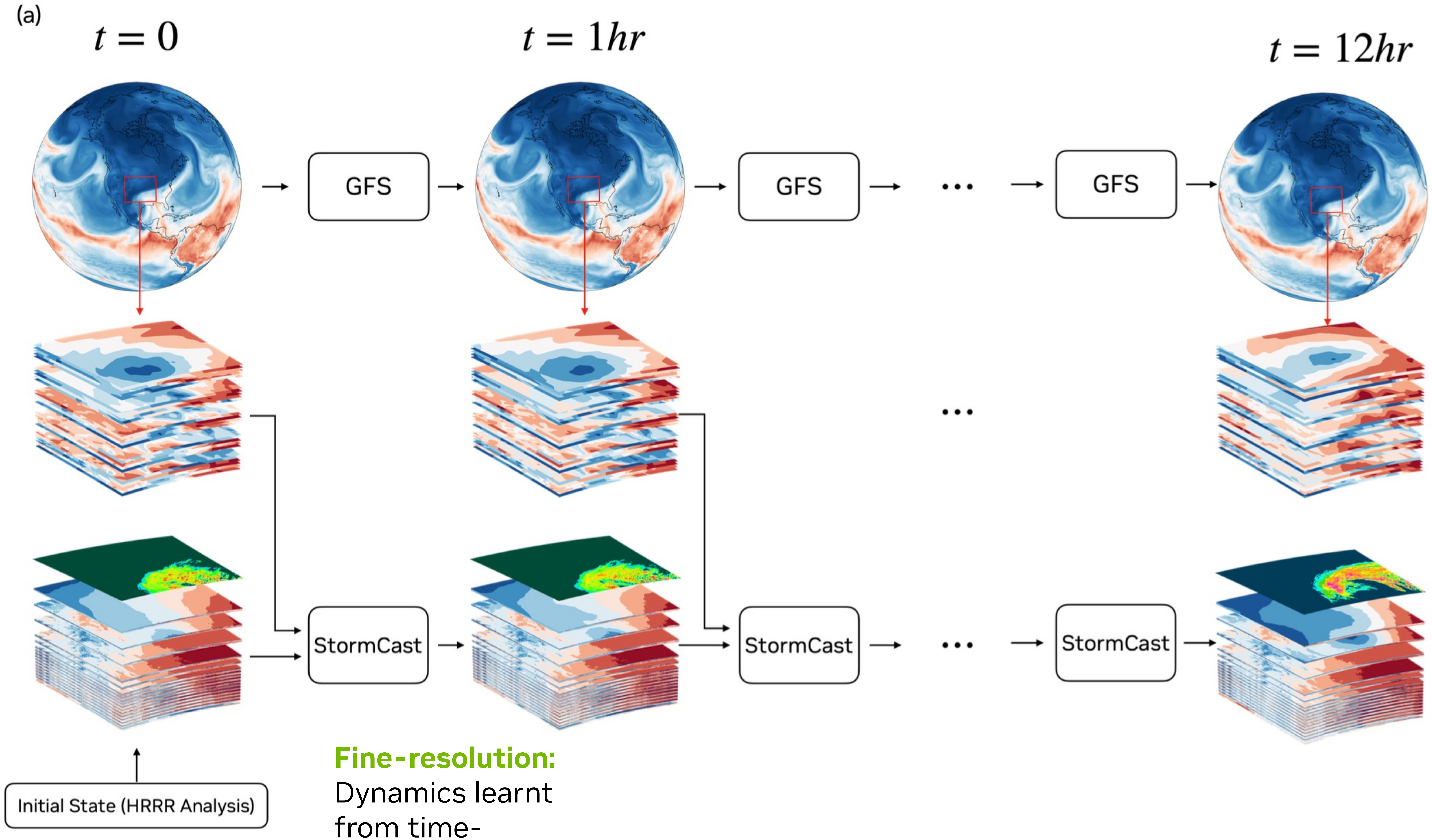

NVIDIA

# StormCast – Km Scale Generative Convection Allowing Model

## A Multi-Scale Inference Setup

**Coarse-resolution:** ML model is made aware of a prior, 25-km global forecast.

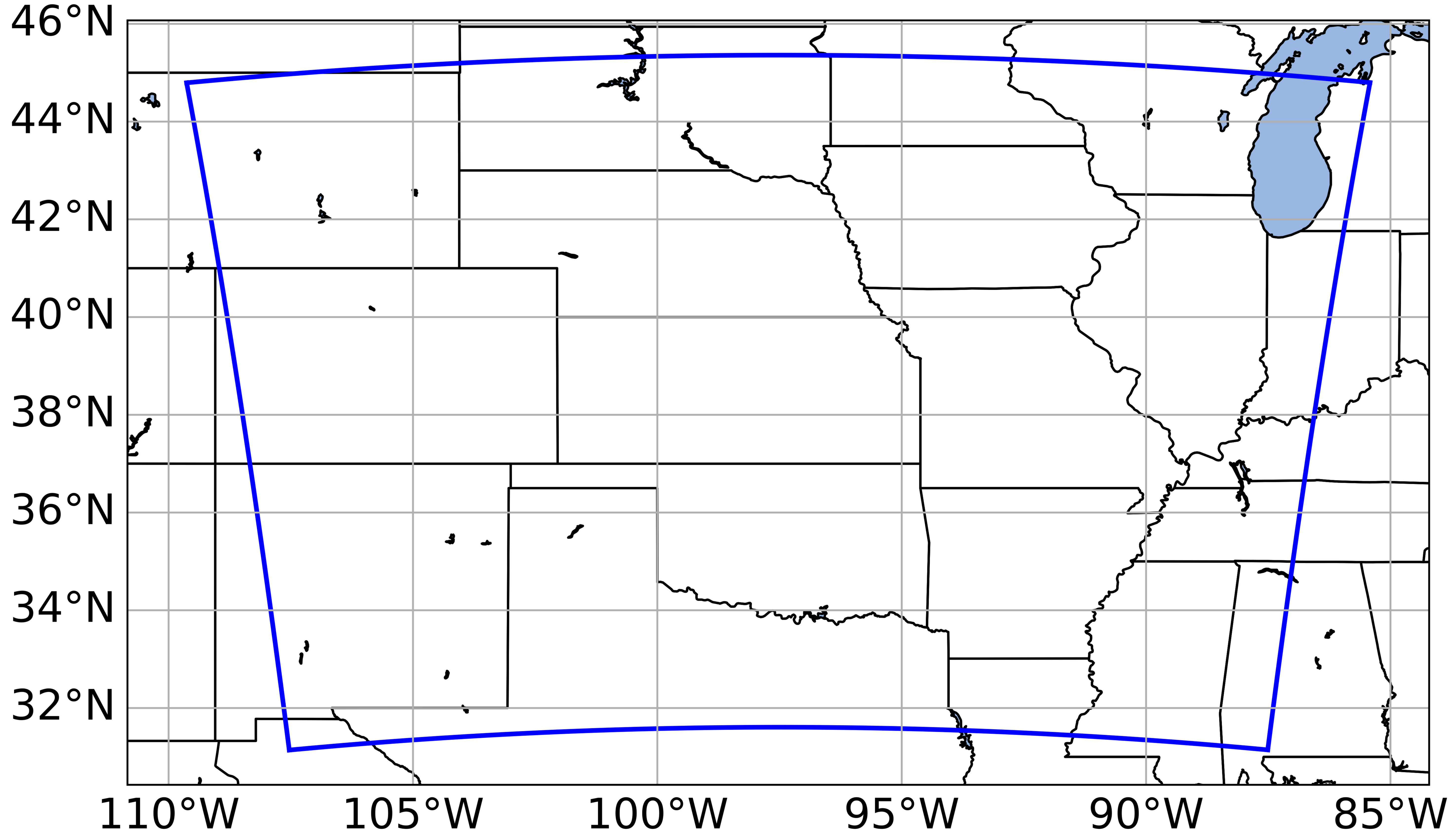
**Fine-resolution:** Is initialized with radar-assimilating HRRR...




**Fine-resolution:** Dynamics learnt from time-stepping km-scale state.

# Domain Extent: Experimental Central US Proving Ground

1536 x1920 km (512 x 640 pixels)



# StormCast – Km Scale Generative Convection Allowing Model

## State vector definition

**Km-scale state :**

- 6 dynamical variables across ~ 16 HRRR vertical levels + scalars.

| ERA5 | | |
|---|---|---|
| Parameter | Pressure levels (hPa) | Height levels (m) |
| Zonal Wind (u) | 1000, 850, 500, 250 | 10 |
| Meridional Wind (v) | 1000, 850, 500, 250 | 10 |
| Geopotential Height (z) | 1000, 850, 500, 250 | None |
| Temperature (t) | 1000, 850, 500, 250 | 2 |
| Humidity (q) | 1000, 850, 500, 250 | None |
| Total Column of Water Vapour (tcwv) | Integrated | - |
| Mean Sea Level Pressure (mslp) | surface | - |
| Surface Pressure (sp) | surface | - |
| HRRR | | |
| Parameter | Hybrid model levels (index) | Height levels (m) |
| Zonal Wind (u) | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 15, 20, 25, 30 | 10 |
| Meridional Wind (v) | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 15, 20, 25, 30 | 10 |
| Geopotential Height (z) | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 15, 20, 25, 30 | None |
| Temperature (t) | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 15, 20, 25, 30 | 2 |
| Humidity (q) | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 15, 20, 25, 30 | None |
| Pressure (p) | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 15, 20 | None |
| Max. Composite Radar Reflectivity (refc) | - | Integrated |
| Mean Sea Level Pressure (mslp) | - | Surface |
| Orography | - | Surface |
| Land/Water Mask | - | Surface |

Table 1: Parameters from the ERA5 and HRRR dataset that are used for training the StormCast model

| Hybrid Level Indices | 1 | 2 | 3 | 4 | | 6 | 7 | 8 | 9 | 10 | 11 | 13 | 15 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Altitude (m) | 125 | 150 | 200 | 280 | 400 | 560 | 750 | 970 | 1210 | 1500 | 1800 | 2500 | 3500 | 6300 |

**Lead time:**

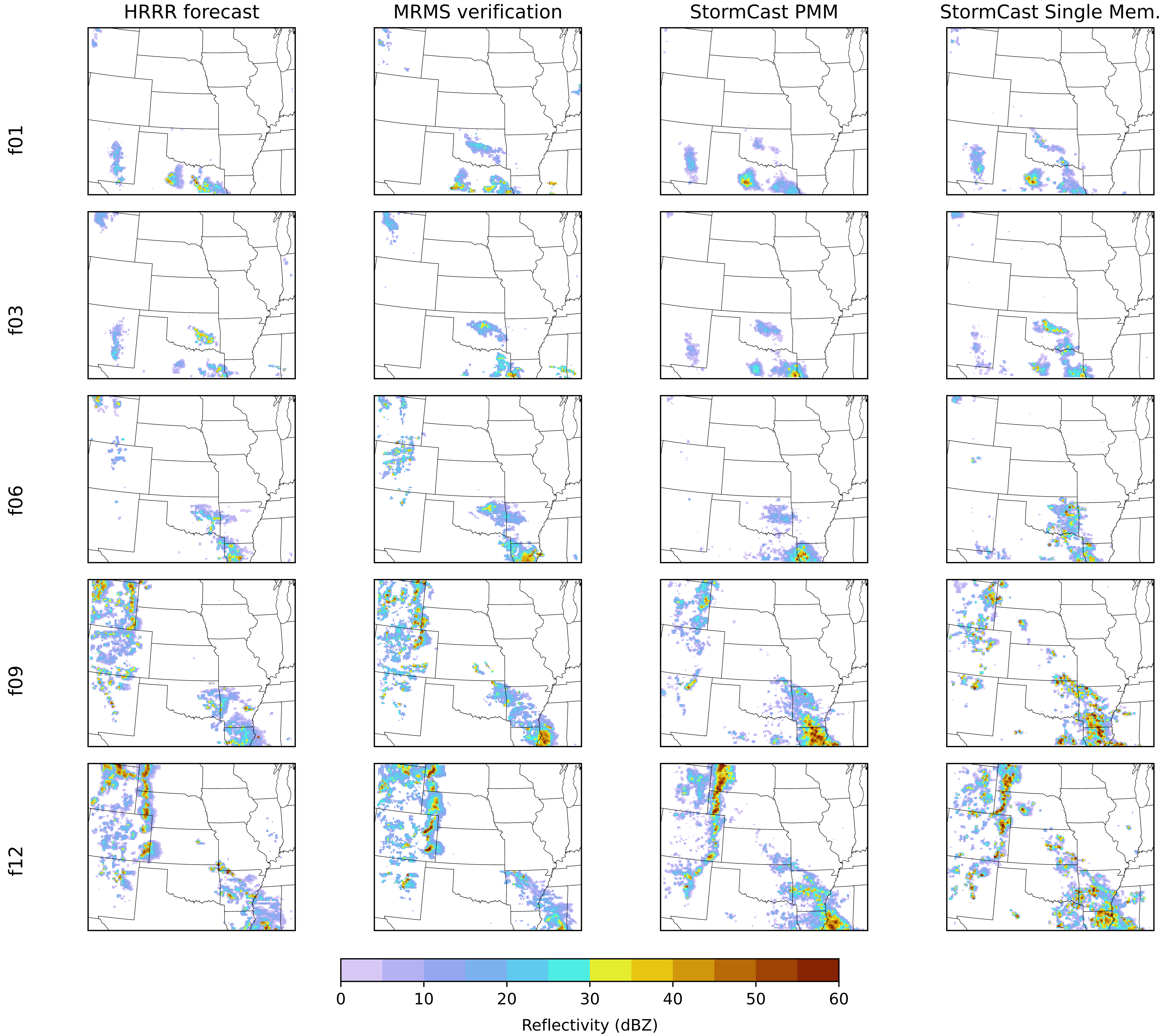
1 hour

3 hours

6 hours

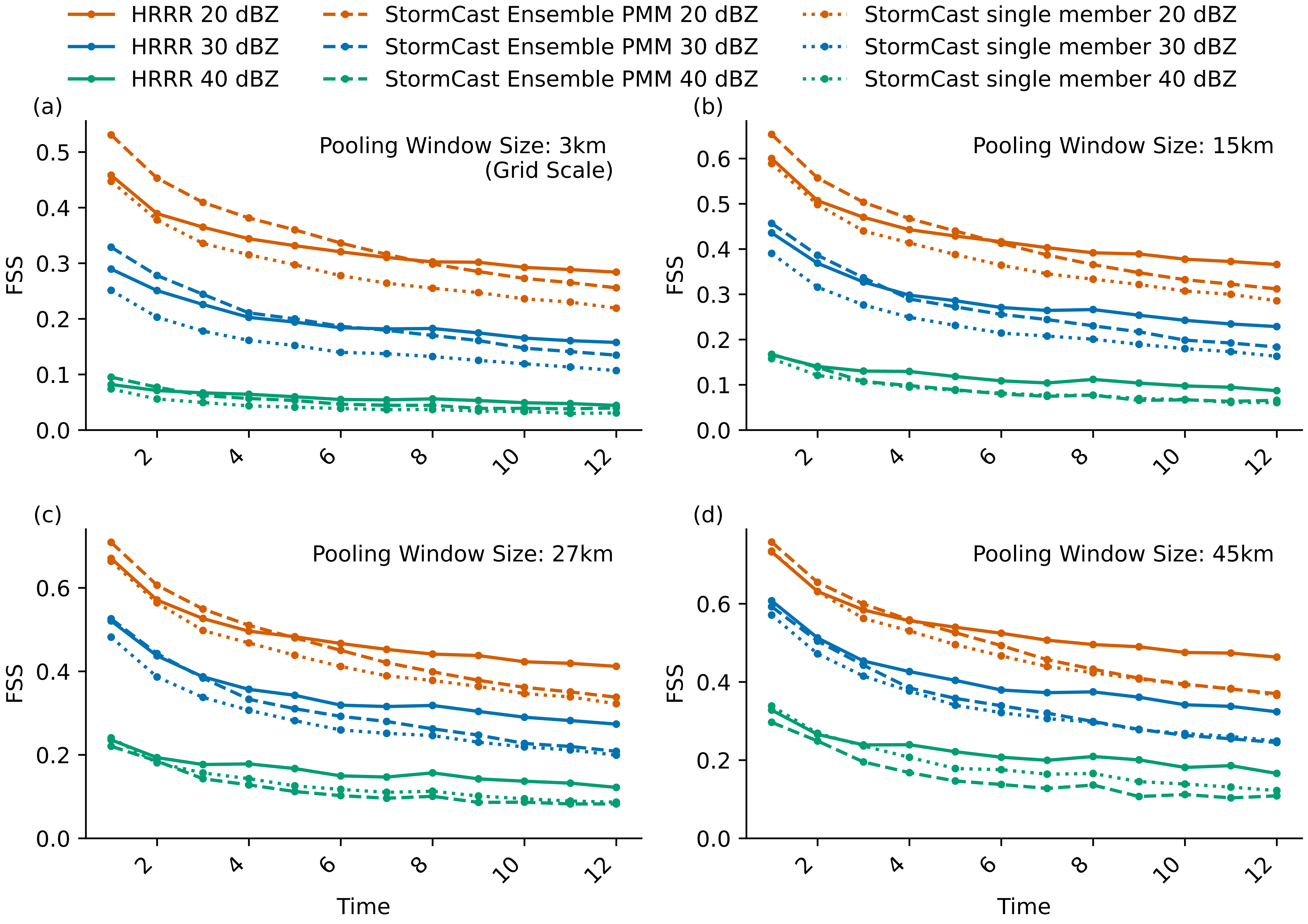
9 hours

12 hours

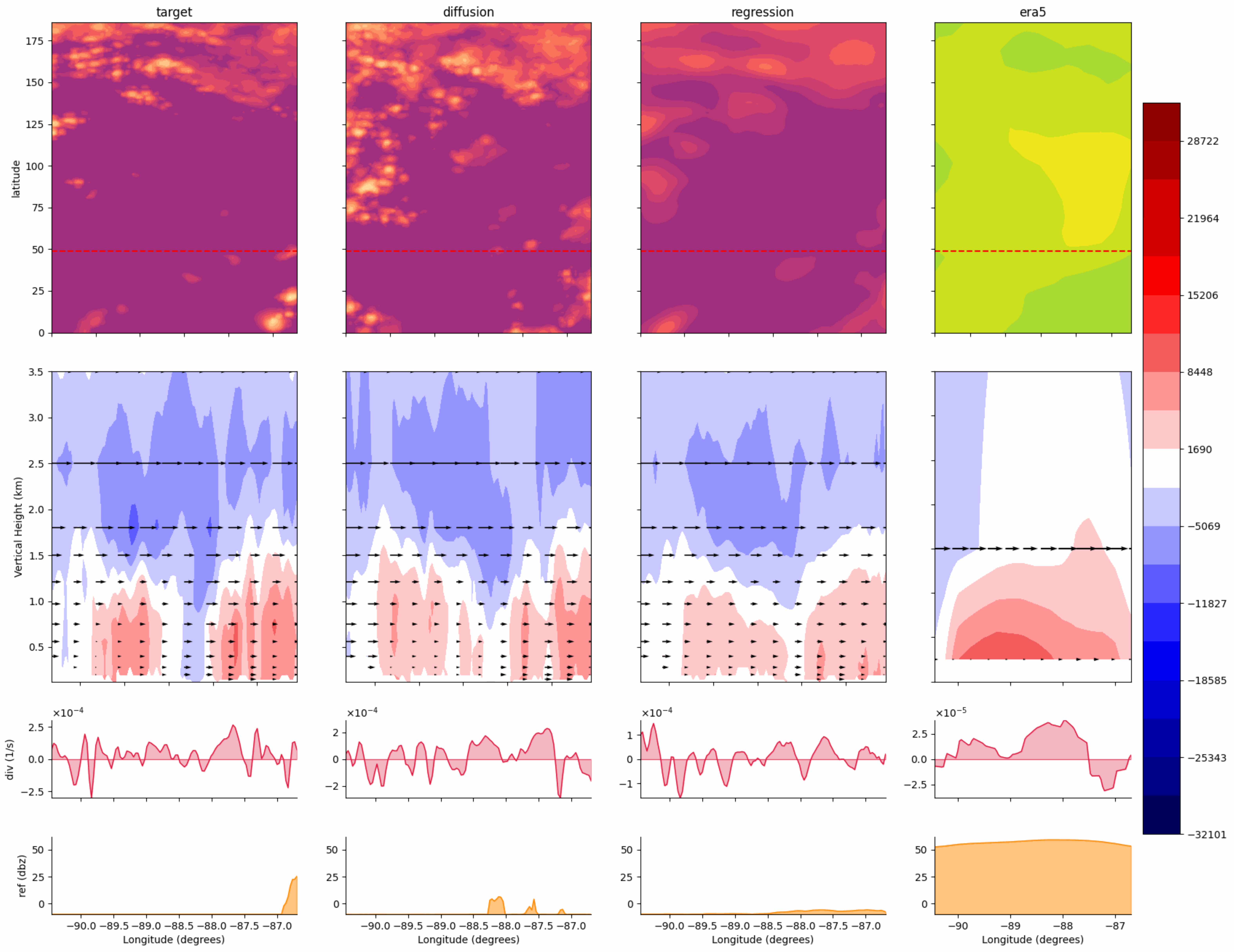


# Results: Forecast Skill Comparison

## Competitive skill with HRRR using 5-member ensemble PMM



# Convective scale motions



# Results: Spectra

High-fidelity km-scale spectra without blurring

- **StormCast**: An encouraging new ML model for km-scale prediction.
- **Results**: 5-member radar forecasts with PMM surpassing HRRR deterministic skill across 1-5h lead times.
- **State**: 96 prognostic km-scale vars – incl. hor. winds, temperature, pressure & humidity across16 HRRR model levels.
- **Design:** Multi-scale & Stochastic Architecture
  - Deterministic backbone with diffusion correction.
  - Forecasts conditioned by GFS during rollout & initialized by HRRR analysis.

- **Possibilities:** How much can we improve forecast accuracy and uncertainty with 100s to 1000s of ensemble members from AI mesoscale models?
- **Preprint:** https://arxiv.org/abs/2408.10958