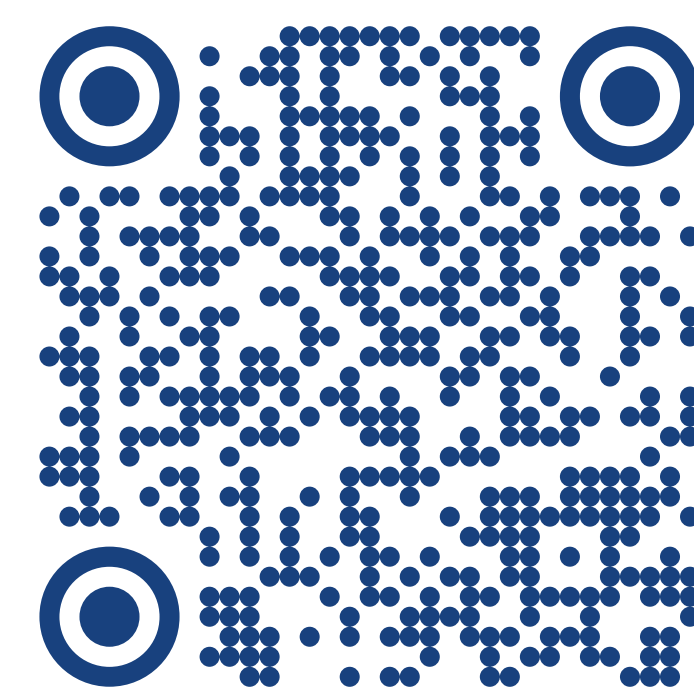


(Sparse) Attention to the Details: Preserving Spectral Fidelity in ML-based Weather Forecasting Models



paper

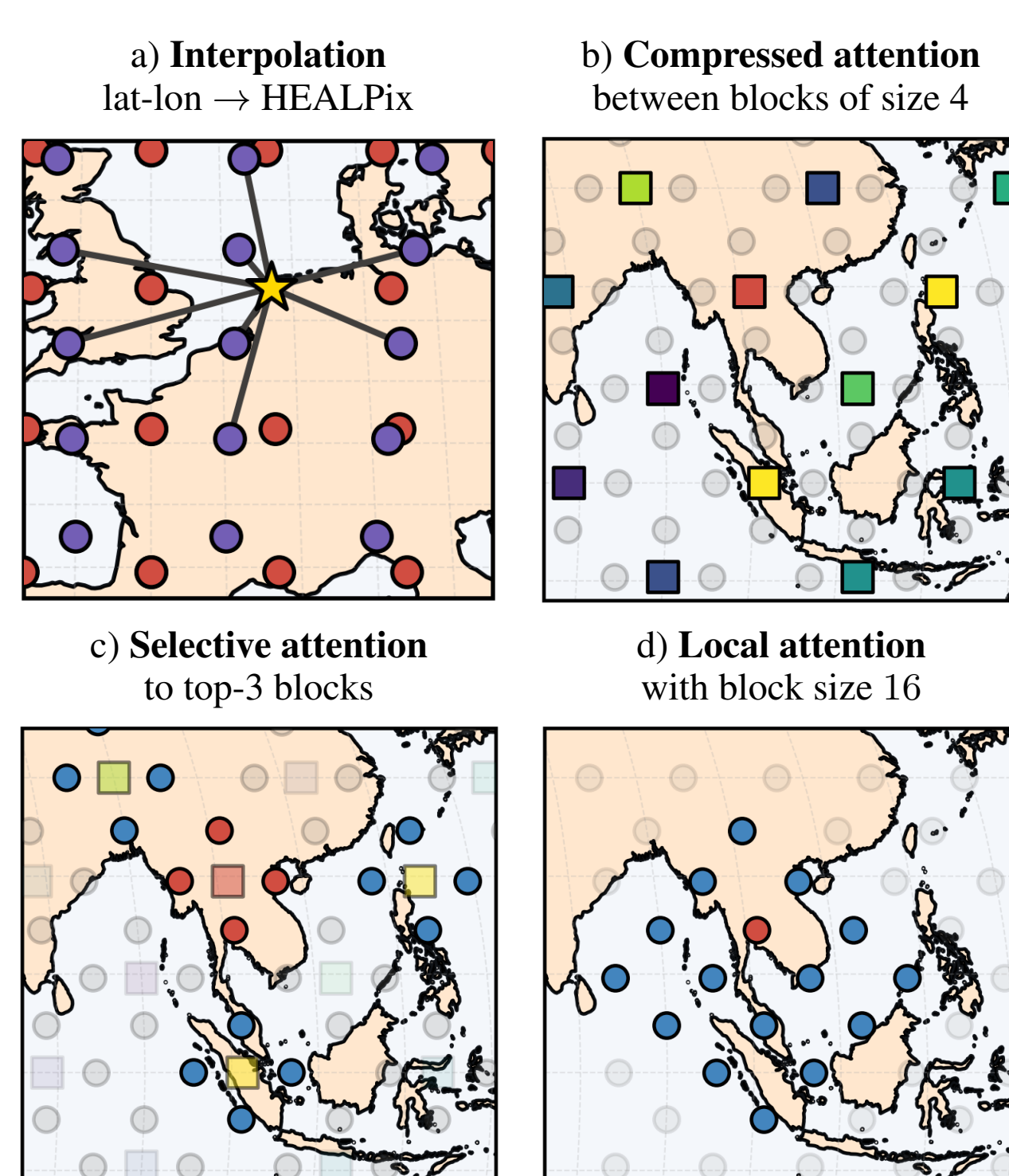
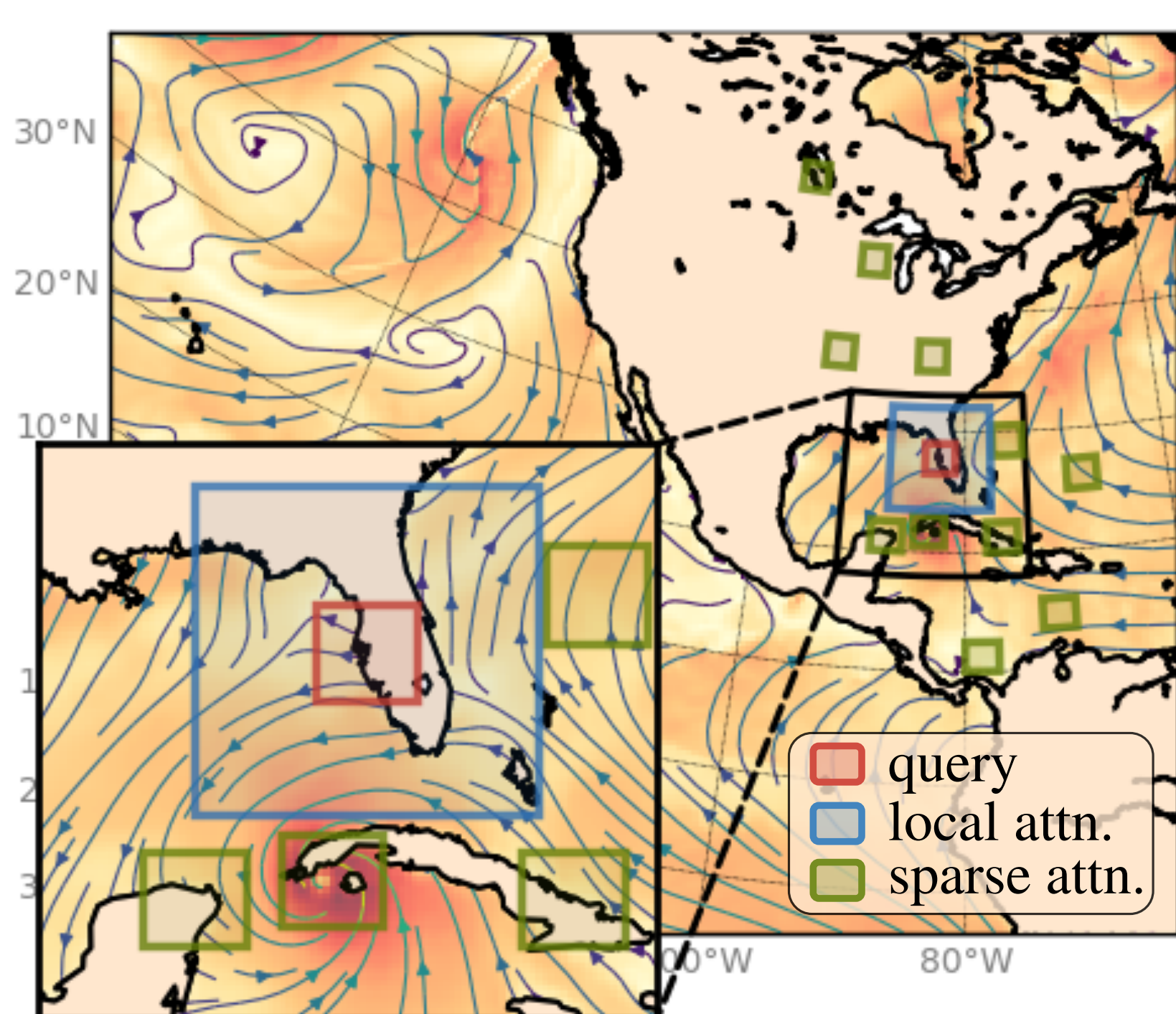
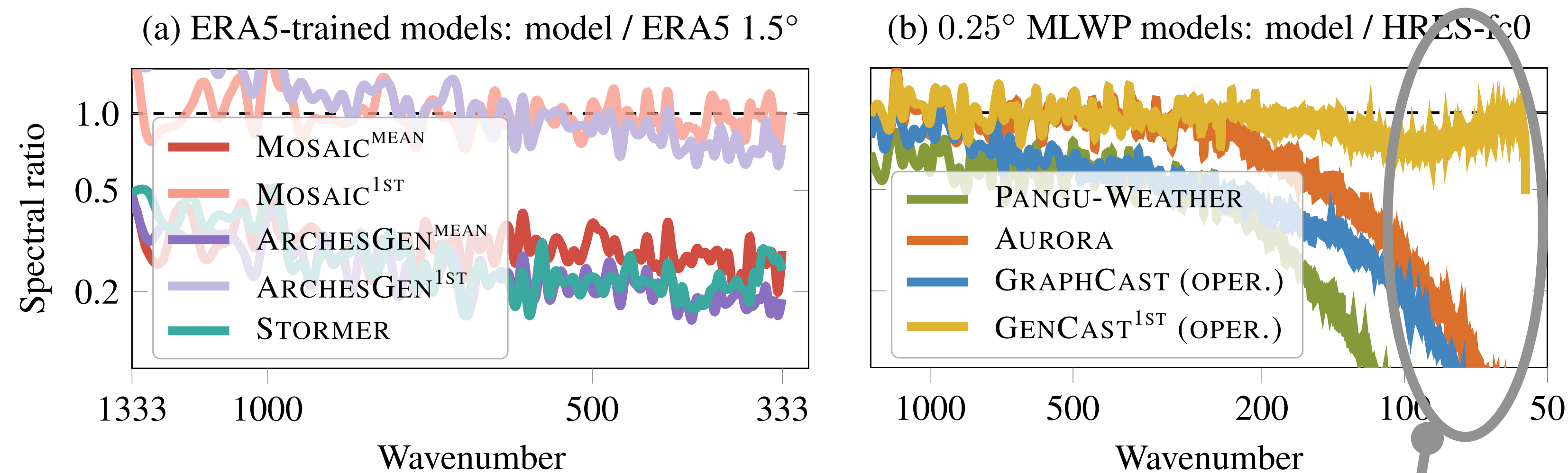
Maksim Zhdanov Ana Lucic Max Welling Jan-Willem van de Meent

AMLab, University of Amsterdam, CuspAI

Motivation

Large Weather Models are fast and skilful, but they suppress fine-scale details. We argue that the **spectral degradation** has two sources:

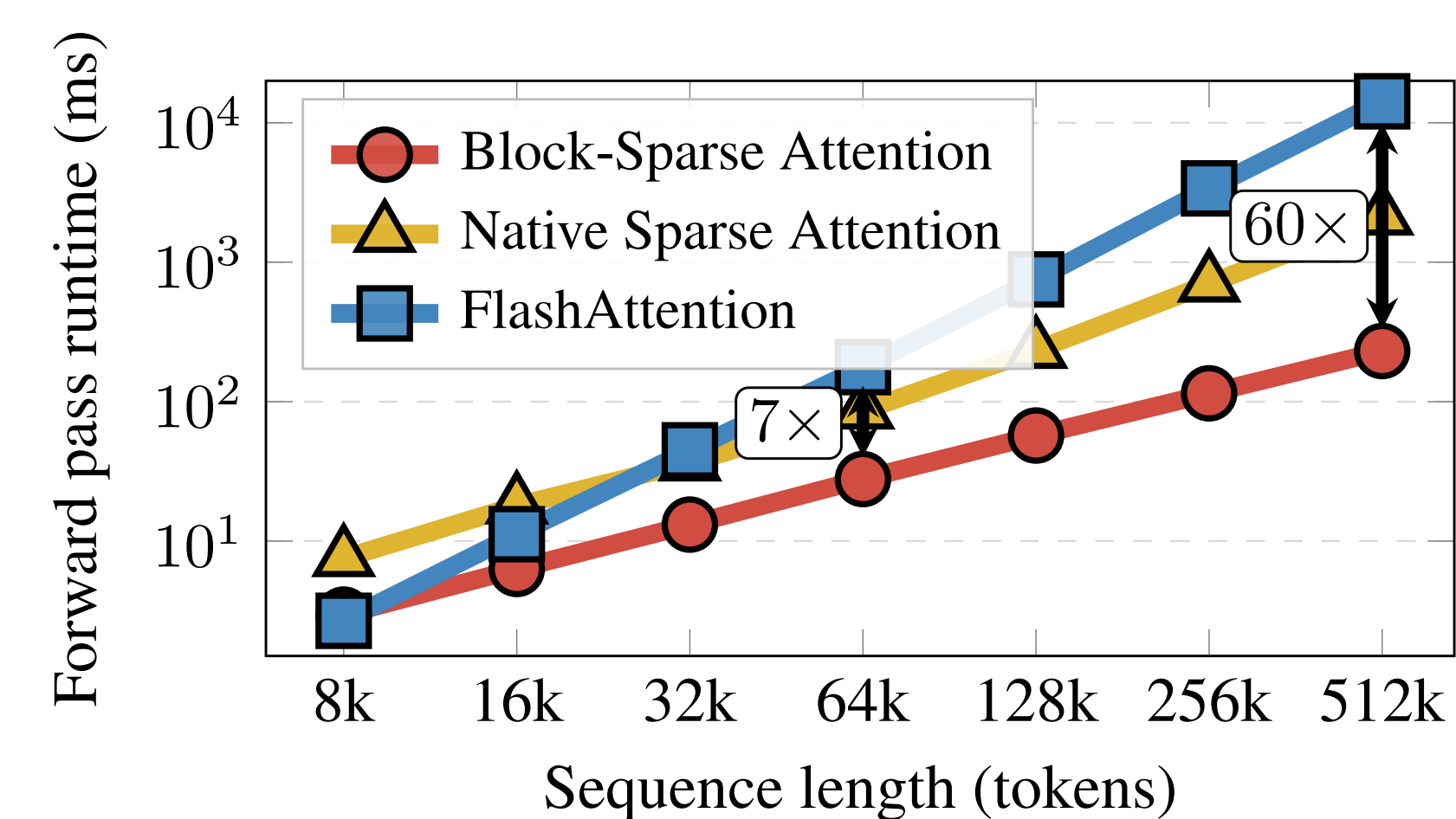
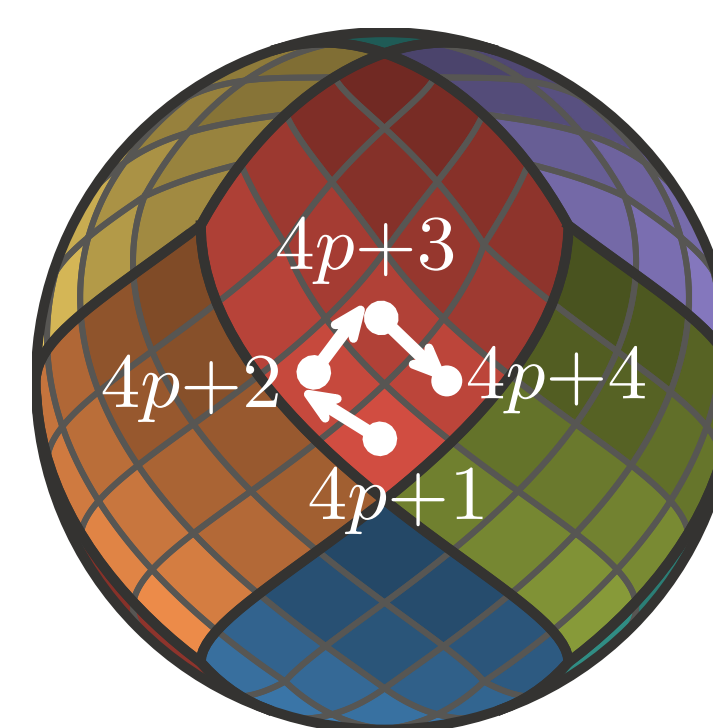
- 1) **Statistical**: deterministic model predicting the ensemble mean.
- 2) **Architectural**: projecting to coarse latent grids introduces an information bottleneck.



MOSAIC

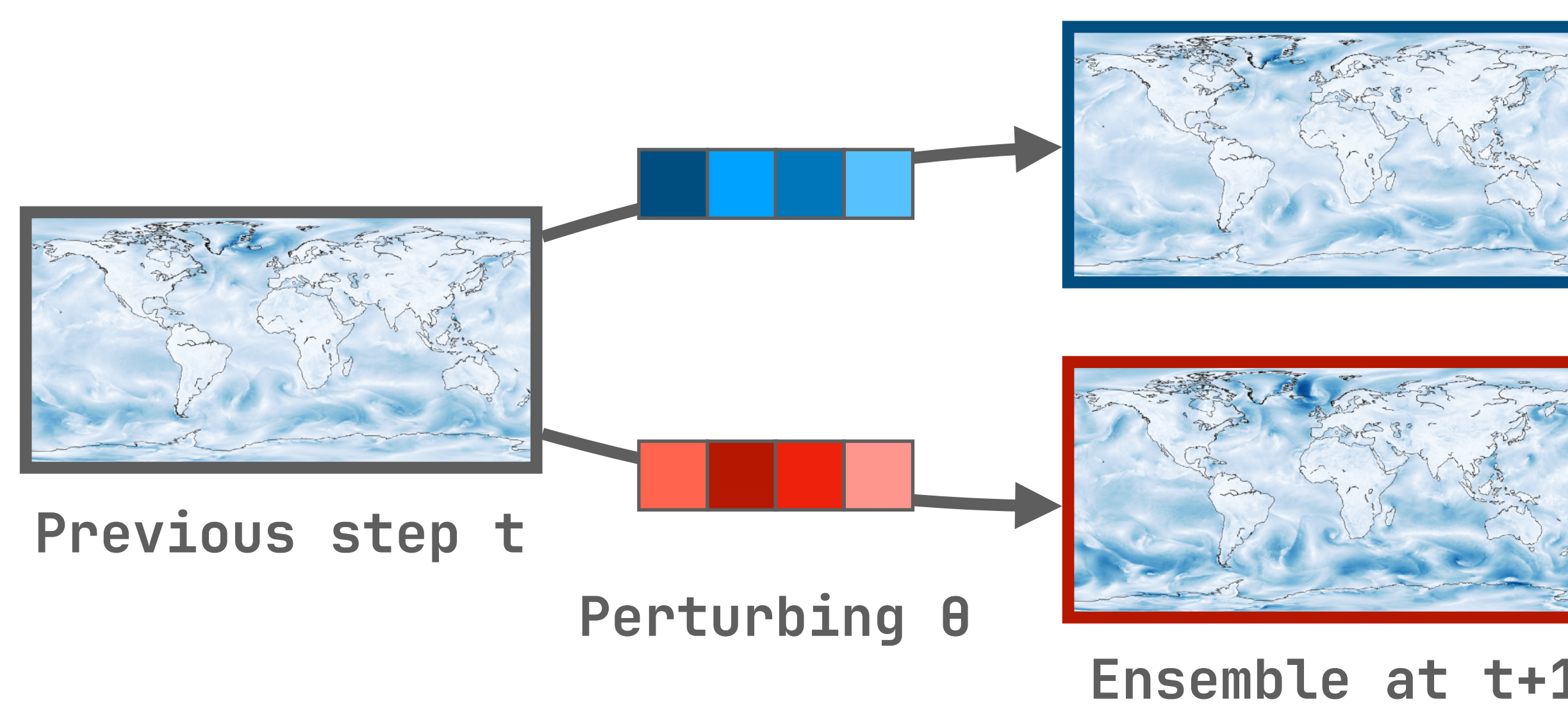
We introduce **MOSAIC**: a 1.5° probabilistic weather forecasting model which addresses both sources of spectral degradation:

Block-sparse attention: to avoid the information bottleneck of compressive encoding, MOSAIC processes weather data at native resolution using block-sparse attention on the HEALPix mesh.



Hierarchical grid

Learned functional perturbations: a noise vector is injected into the model weights, producing structured, input-dependent uncertainty, trained with CRPS.

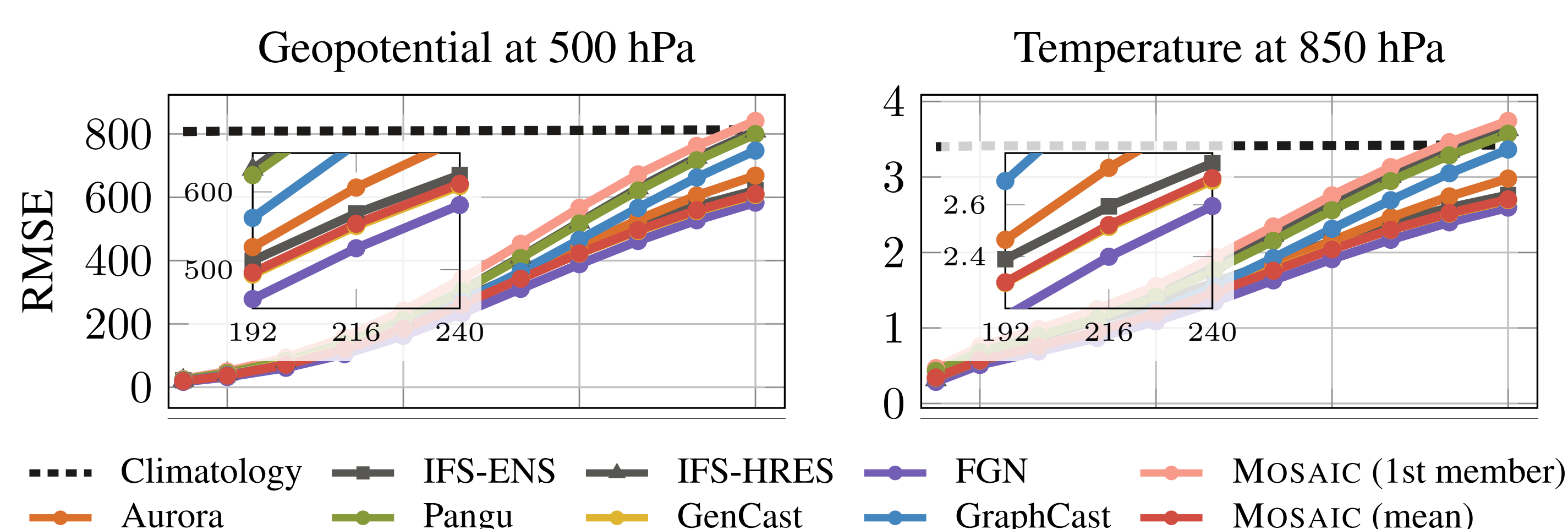
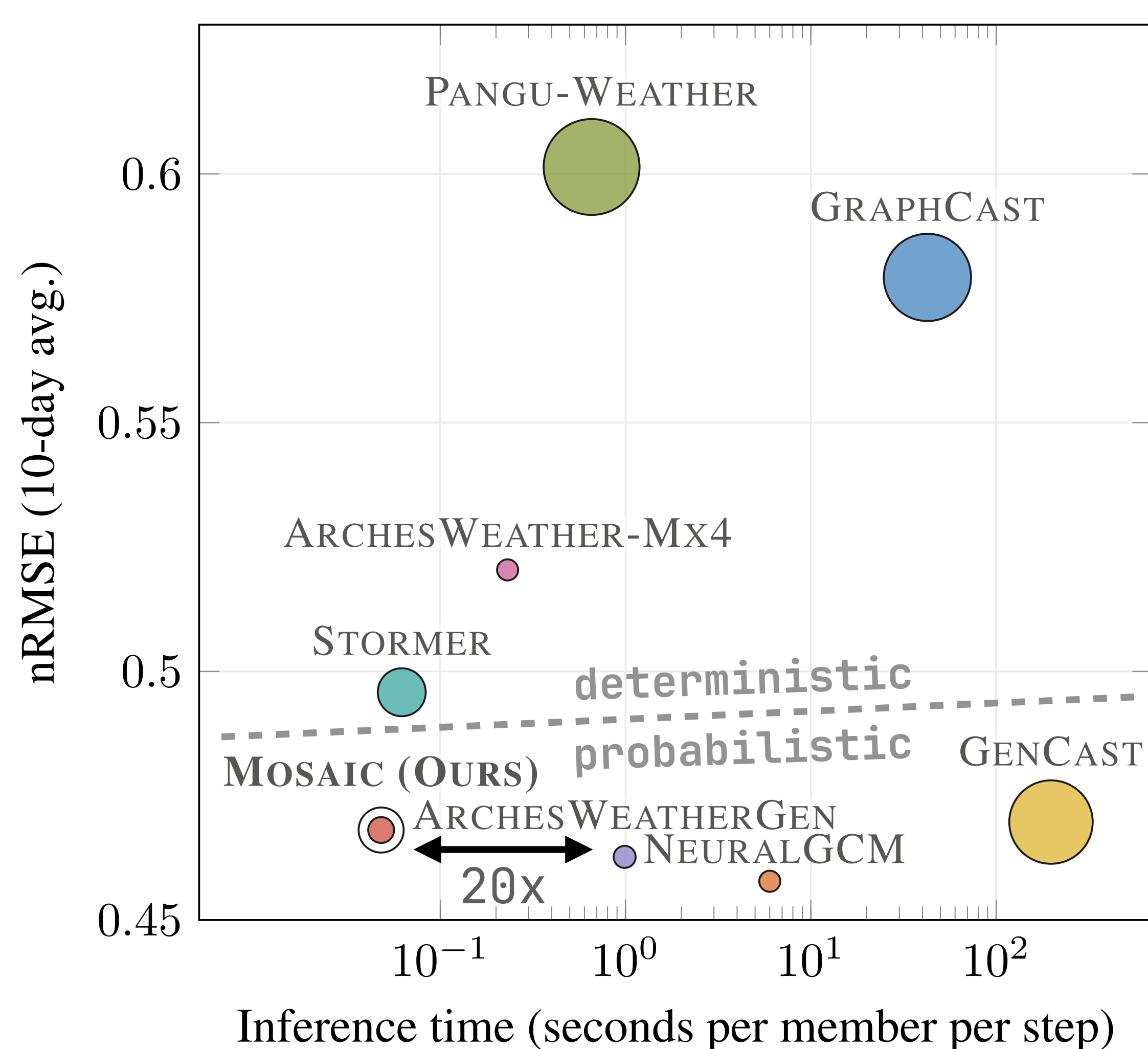


Results

MOSAIC is pretrained on ERA5 1979–2018 at 1.5° and finetuned on HRES-fc0 2016–2021, tested on 2022.

We demonstrate that MOSAIC at 1.5° resolution:

- Competitive with large models in terms of skill.
- Ensemble members match weather data spectra.
- Generates a 24-member, 10-day forecast in <12s.



MOSAIC is able to capture Hurricane Ian's landfall 7 days prior.

Hurricane Ian's exceptional death toll, explained

Florida has seen plenty of hurricanes. Why was Ian so deadly?

