Aurora: A Foundation Model for the Earth System

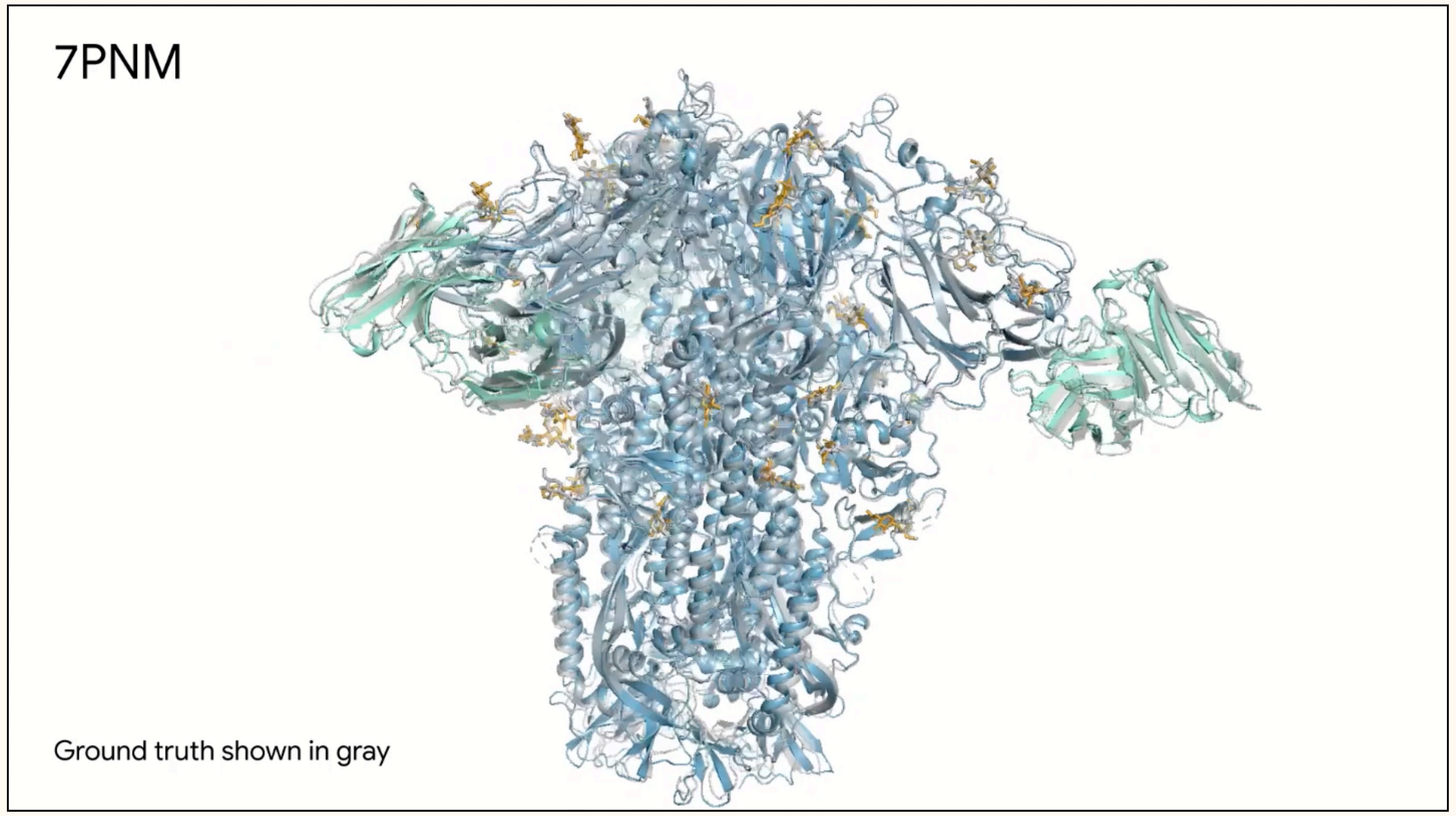
Wessel Bruinsma

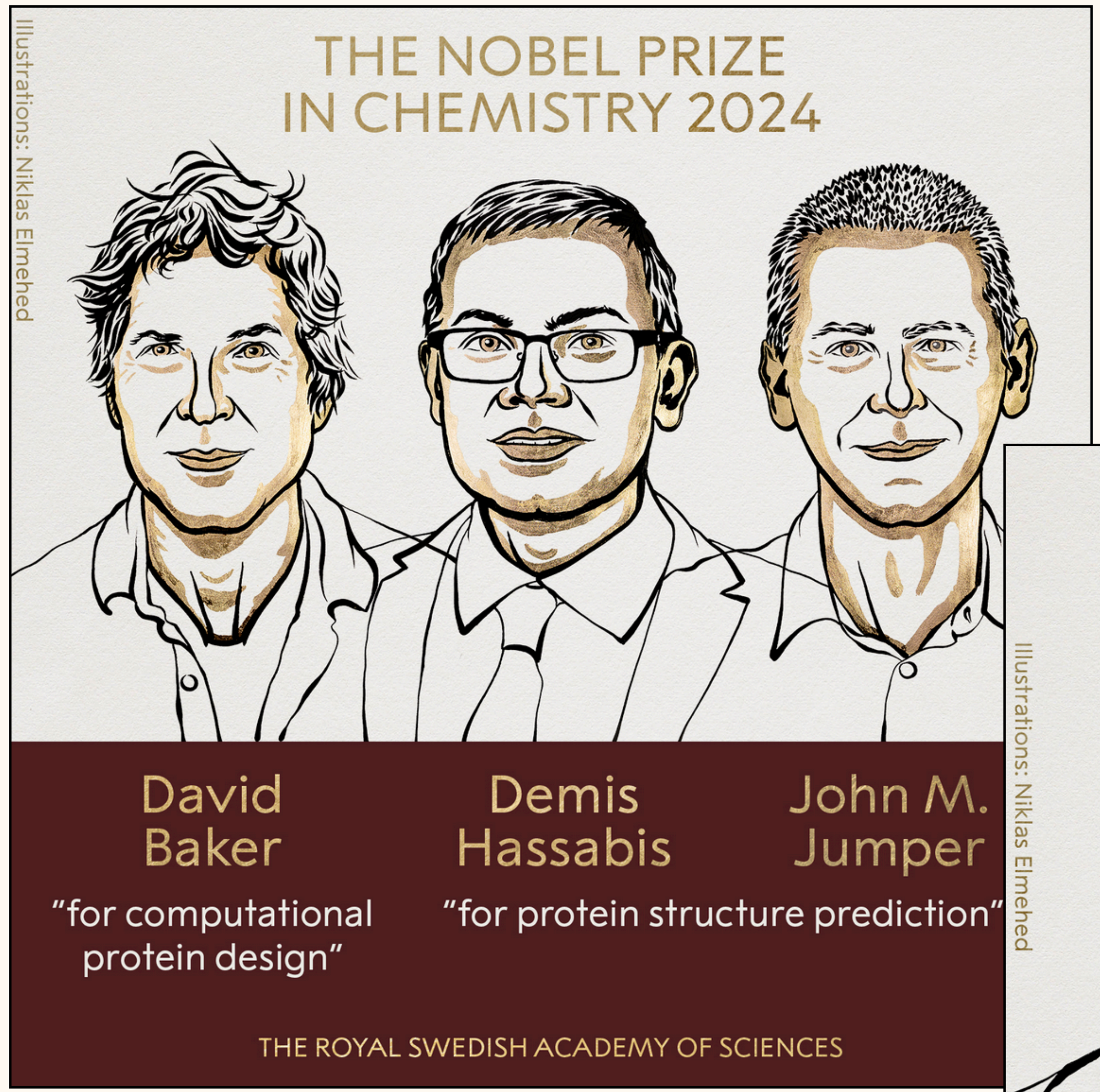
The Alan Turing Institute
Work was done at Microsoft Research

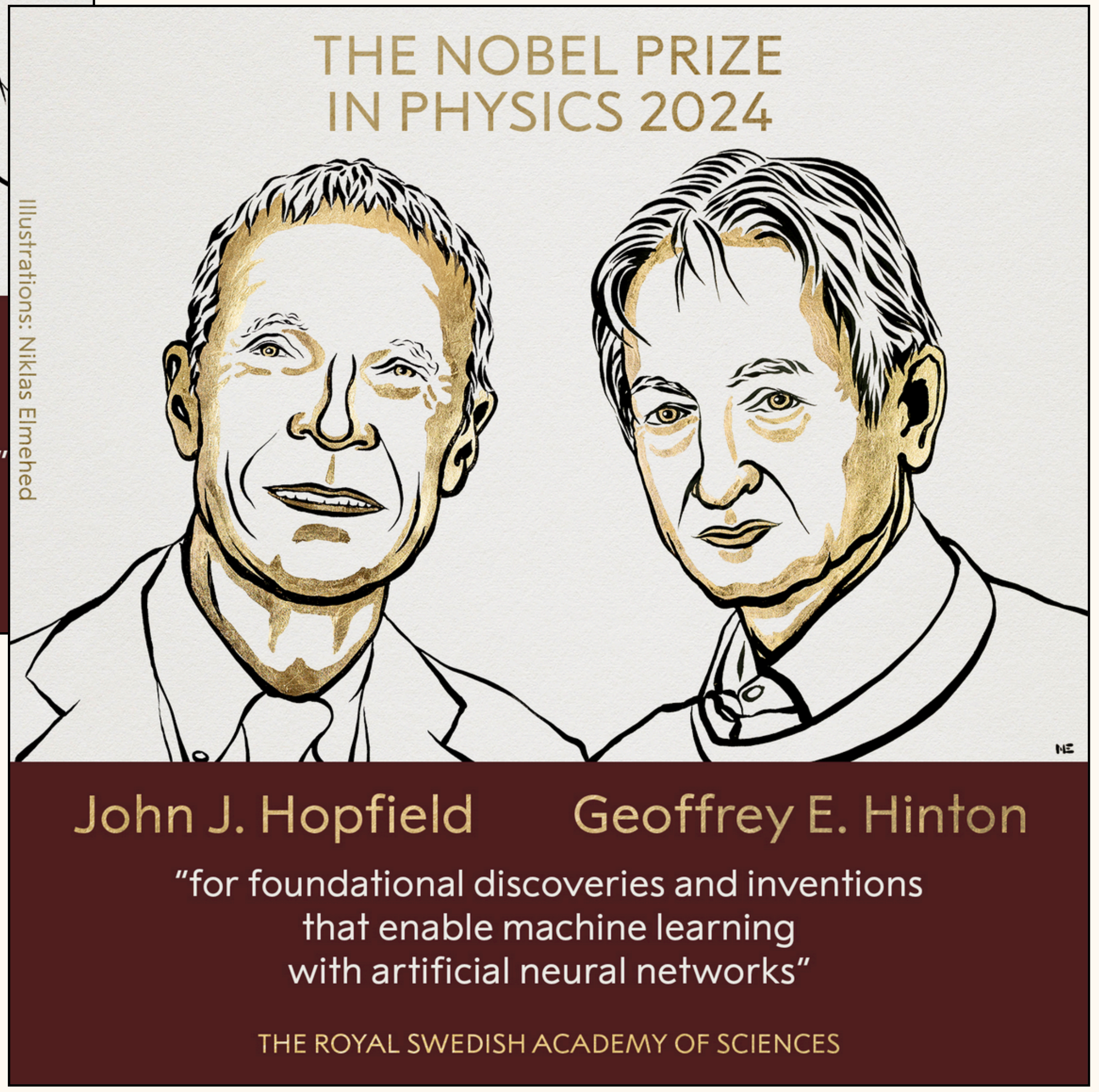
RMetS Early Careers and Student Conference 2025, 1 July 2025

The AI Revolution in Science

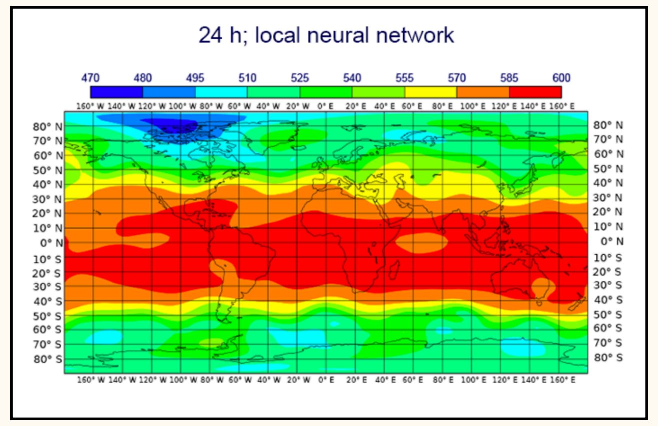
AlphaFold
Protein folding







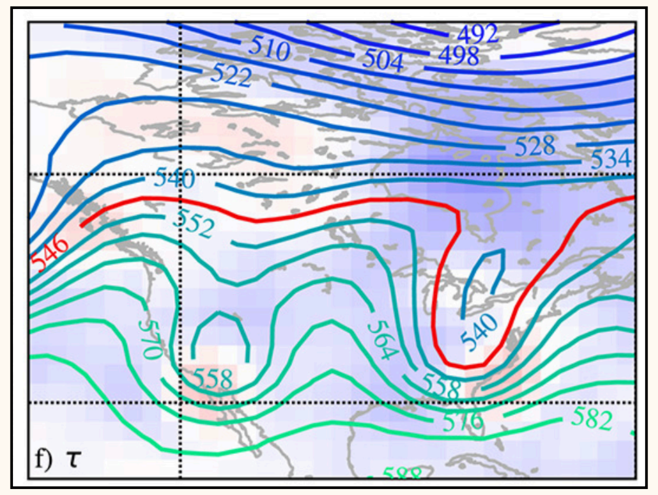
The AI Revolution in Weather Forecasting



2018

First serious efforts to compare AI models to physics baselines

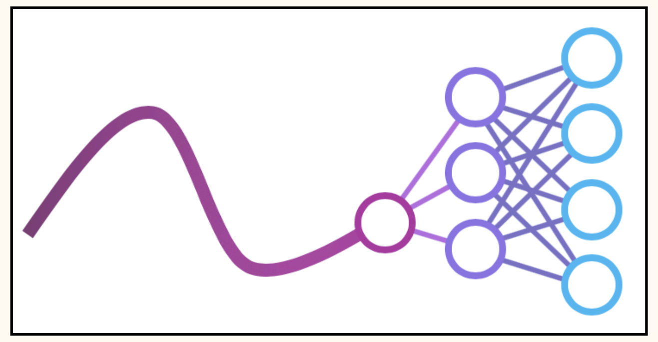
Dueben and Bauer (2018)



2019

AI models skillful to multiple days

Weyn et al. (2019)

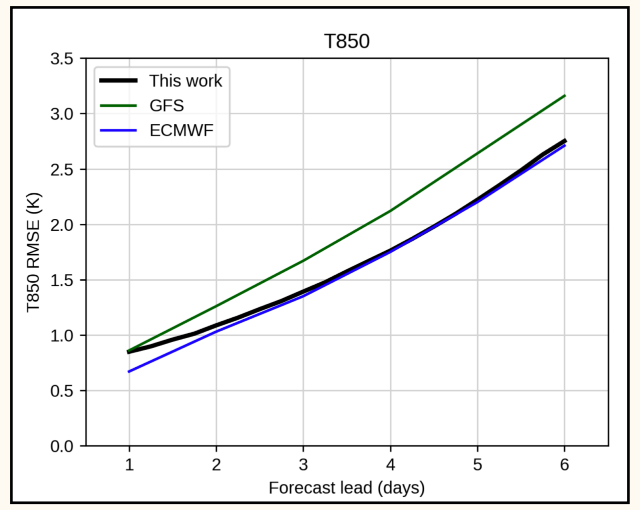


2020

WeatherBench starts to drive ML development

Rasp et al. (2020)

The AI Revolution in Weather Forecasting



2022

GNN outperforms GFS at 1°

Keisler (2022)

HUAWEI

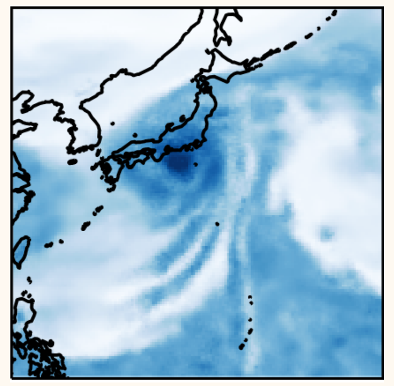
2022

Pangu-Weather outperforms HRES at 0.25°

Bi et al. (2023)

The AI Revolution in Weather Forecasting

2022–2023 Tech companies start to work in this space

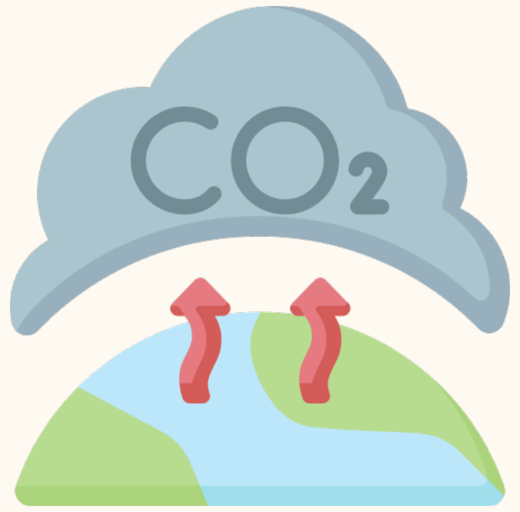


2023 GenCast outperforms IFS ensemble
Price et al. (2024)

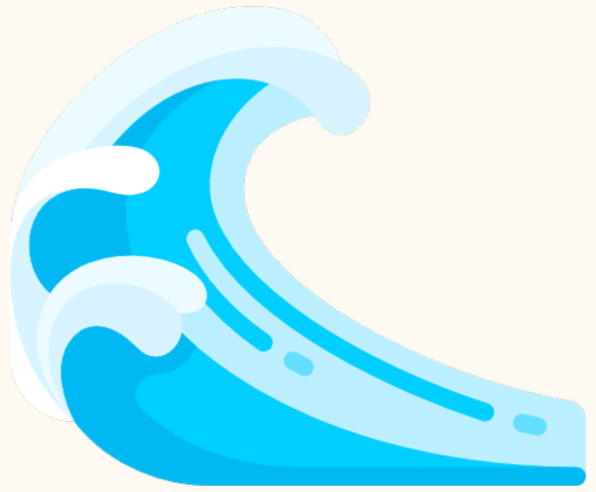
2024 ECMWF launches AIFS

What About Other Forecasting Tasks?

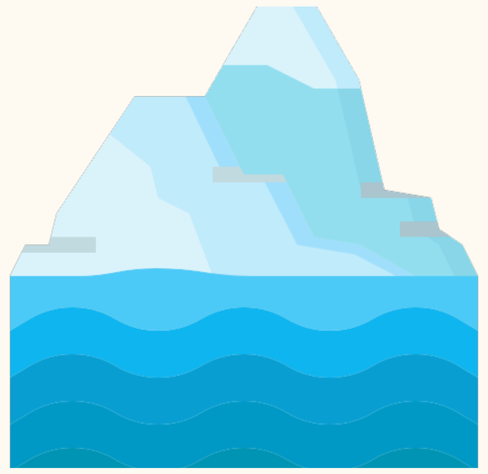
Air pollution



Atmospheric composition



Waves



Sea ice

Ocean

- Current models are impressive, but **limited to one setting**.
- Unified approach?

Aurora

pretraining

- Train a single neural network a *large* body of Earth system data
- Learn general-purpose representation of dynamics that govern atmospheric and oceanic flow
- Slow and data hungry

fine-tuning

- Leverage learned representation to **efficiently adapt to new domains**!
- Fast and data efficient

Aurora: a **foundation model** for the Earth system

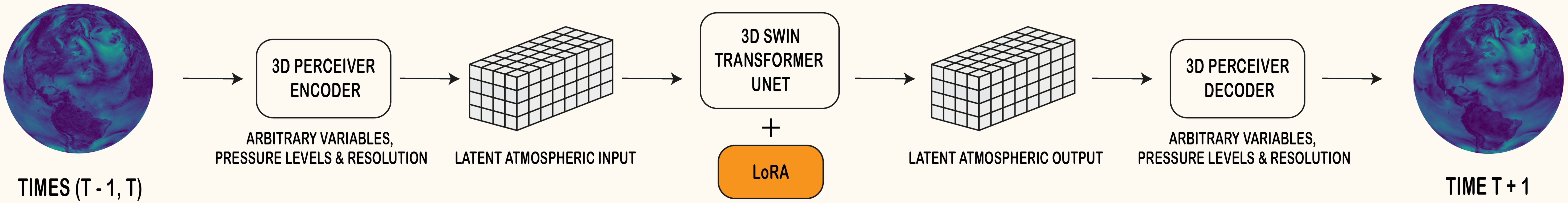
The Model

- Predict global state of **any variables** at **any resolution** 6 h ahead:

$$\hat{X}^{t+6\,\mathrm{h}} = \Phi(X^{t}, X^{t-6\,\mathrm{h}}),$$

$$\begin{aligned} \hat{X}^{t+12\,\mathrm{h}} &= \Phi(\hat{X}^{t+6\,\mathrm{h}}, X^{t}), \\ \hat{X}^{t+18\,\mathrm{h}} &= \Phi(\hat{X}^{t+12\,\mathrm{h}}, \hat{X}^{t+6\,\mathrm{h}}), \\ &\vdots \end{aligned}$$

- Transformer-based encoder–decoder architecture:

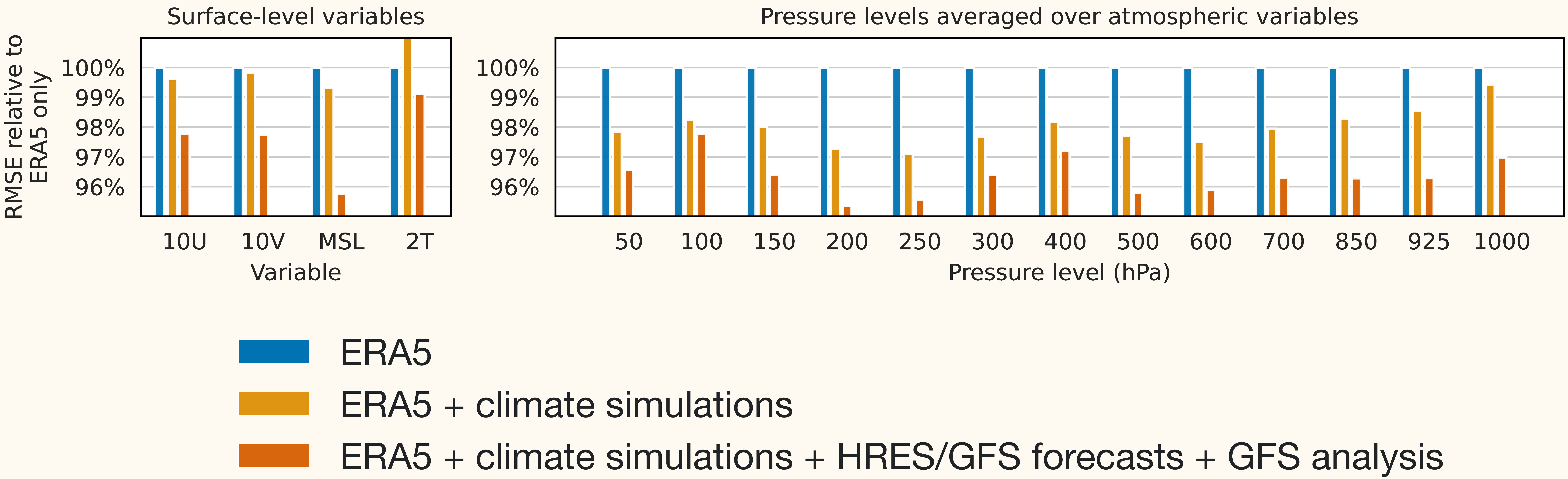


Pretraining

Name	Resolution	Timeframe	Surf. variables	Atmos. variables	Levels	Steps	Size
ERA5	$0.25° \times 0.25°$	1979–2020	2T, 10U, 10V, MSL	U, V, T, Q, Z	13	368.18 k	105.50 TB
HRES-0.25 forecasts	$0.25° \times 0.25°$	2016–2020	2T, 10U, 10V, MSL	U, V, T, Q, Z	13	149.81 k	42.93 TB
IFS-ENS-0.25	$0.25° \times 0.25°$	2018–2020	2T, 10U, 10V, MSL	U, V, T, Q, Z	3	6.69 M	527.54 TB
IFS-ENS-0.25 mean	$0.25° \times 0.25°$	2018–2020	2T, 10U, 10V, MSL	U, V, T, Q, Z	3	133.71 k	10.55 TB
GFS forecasts	$0.25° \times 0.25°$	Feb 2015–2020	2T, 10U, 10V, MSL	U, V, T, Q, Z	13	354.40 k	101.56 TB
GFS T0	$0.25° \times 0.25°$	Feb 2015–2020	2T, 10U, 10V, MSL	U, V, T, Q, Z	13	8.64 k	2.48 TB
GEFS reforecasts	$0.25° \times 0.25°$	2000–2019	2T, MSL	U, V, T, Q, Z	7	2.96 M	454.61 TB
CMCC-CM2-VHR4	$0.25° \times 0.25°$	1950–2014	2T, 10U, 10V, MSL	U, V, T, Q	7	94.96 k	12.62 TB
ECMWF-IFS-HR	$0.45° \times 0.45°$	1950–2014	2T, 10U, 10V, MSL	U, V, T, Q, Z	7	94.96 k	4.75 TB
MERRA-2	$0.625° \times 0.50°$	1980–2020	2T, 10U, 10V, MSL	U, V, T, Q	13	119.81 k	5.58 TB
					Total	10.97 M	1,268.12 TB

- 150 000 steps on 32 GPUs (A100)
- The magic: **data scaling** and **model scaling**!

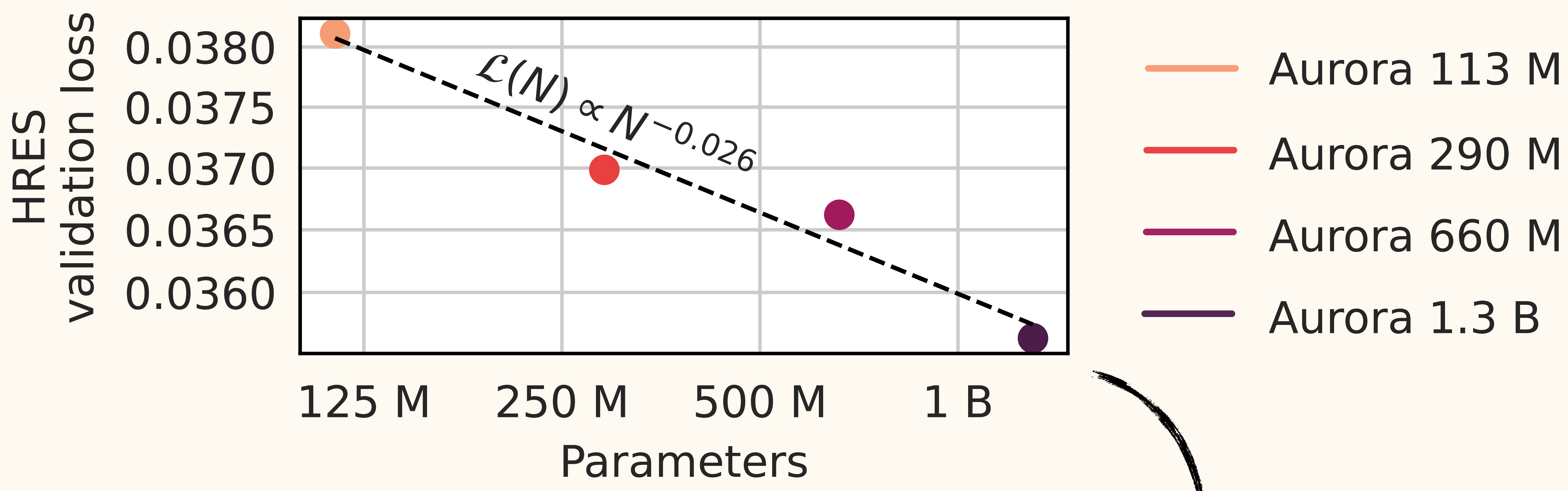
Data Scaling



Model Scaling

Economical to train big models!

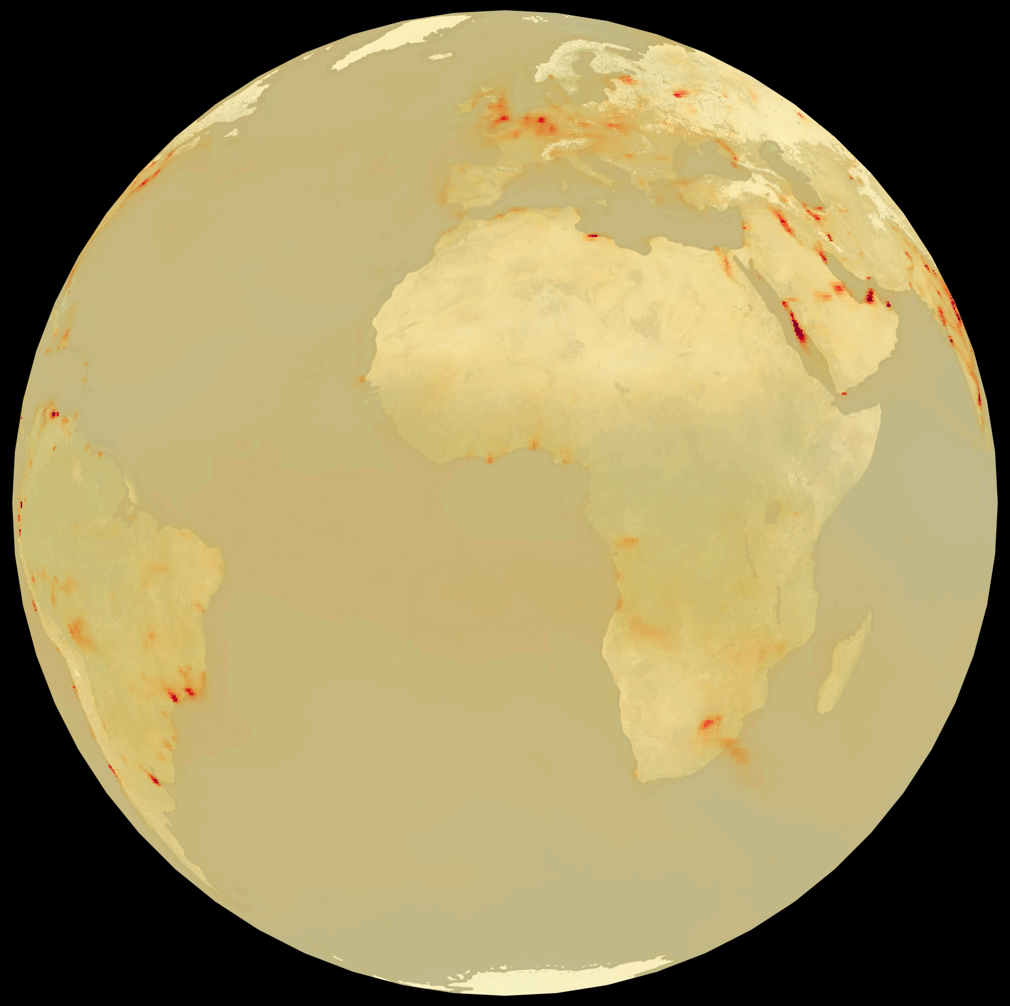
Validation loss at fixed computational budget



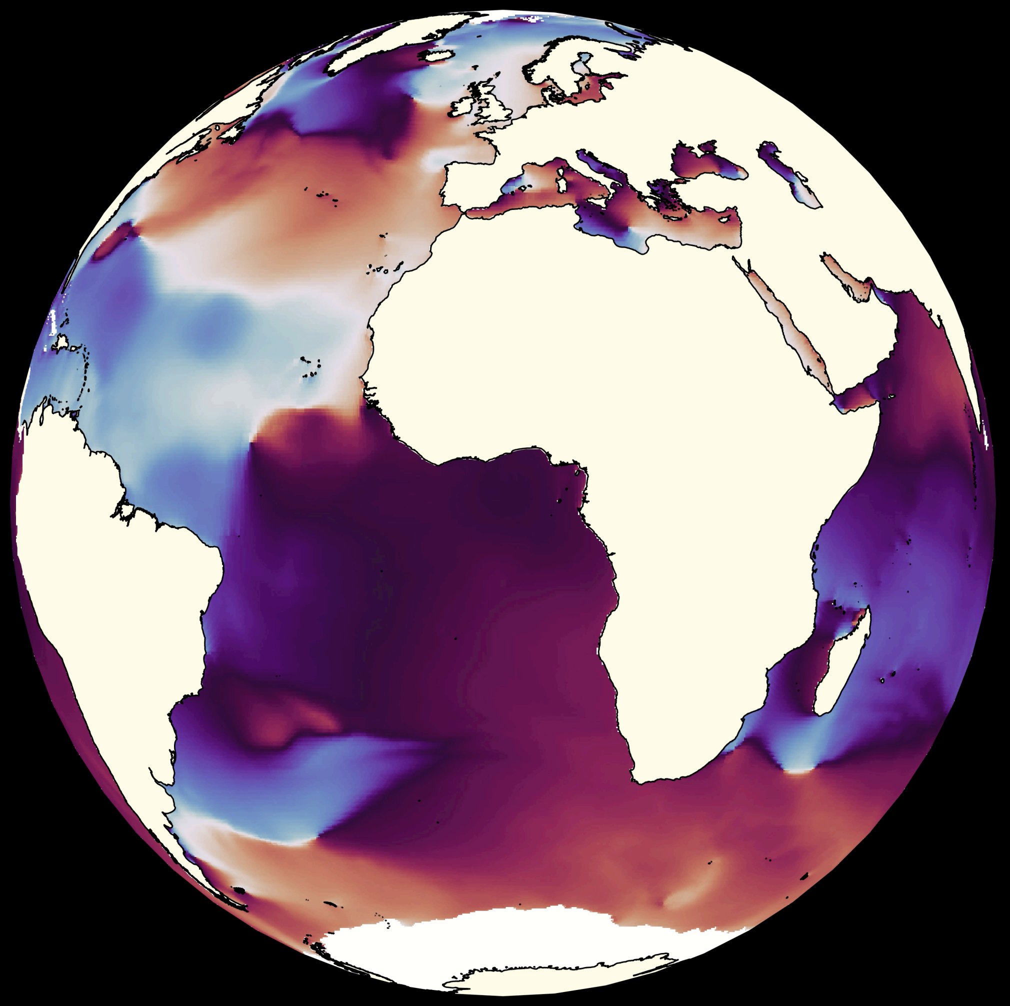
6% reduction per 10× model size!

Fine-Tuning Applications

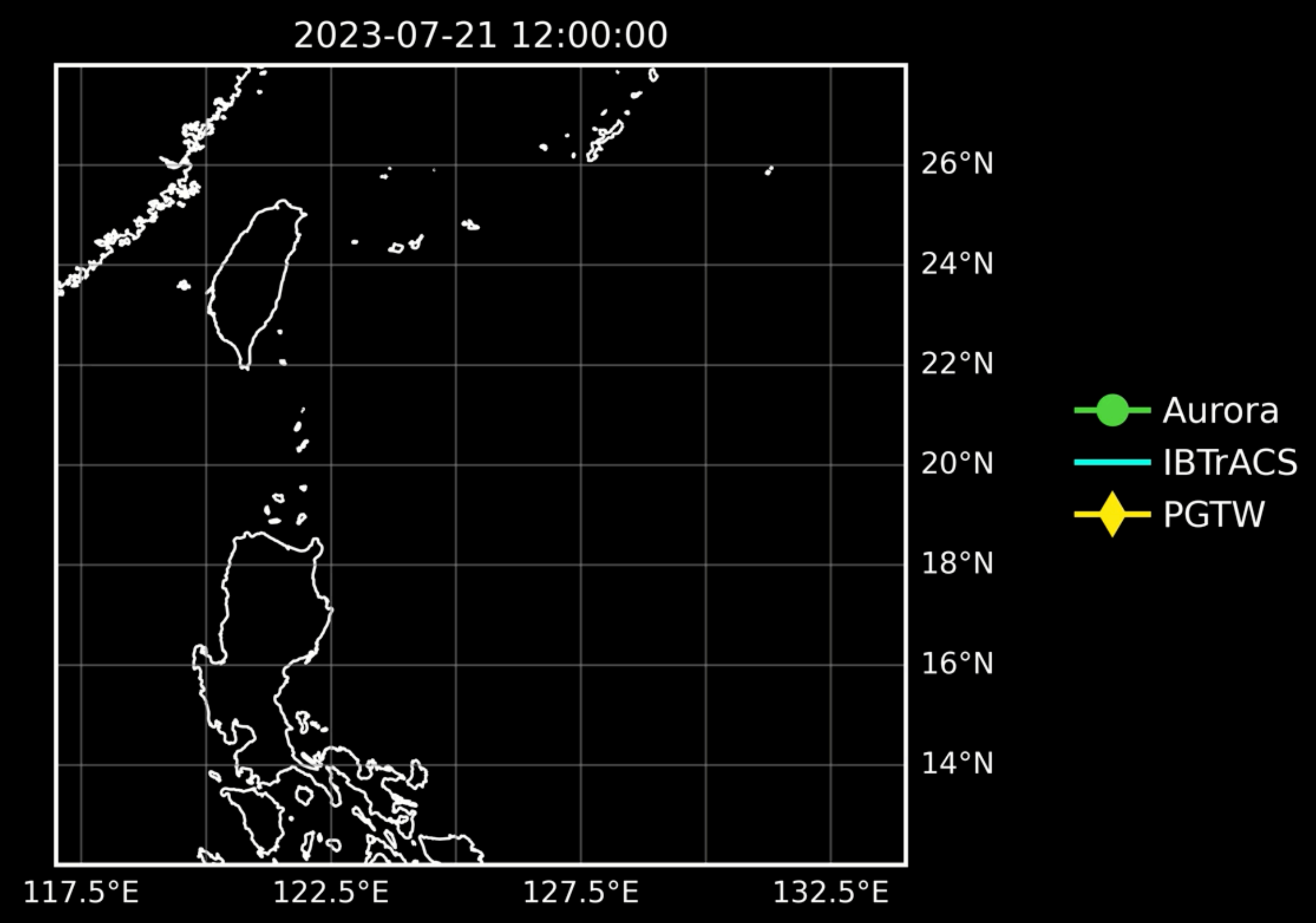
Operational in all settings!



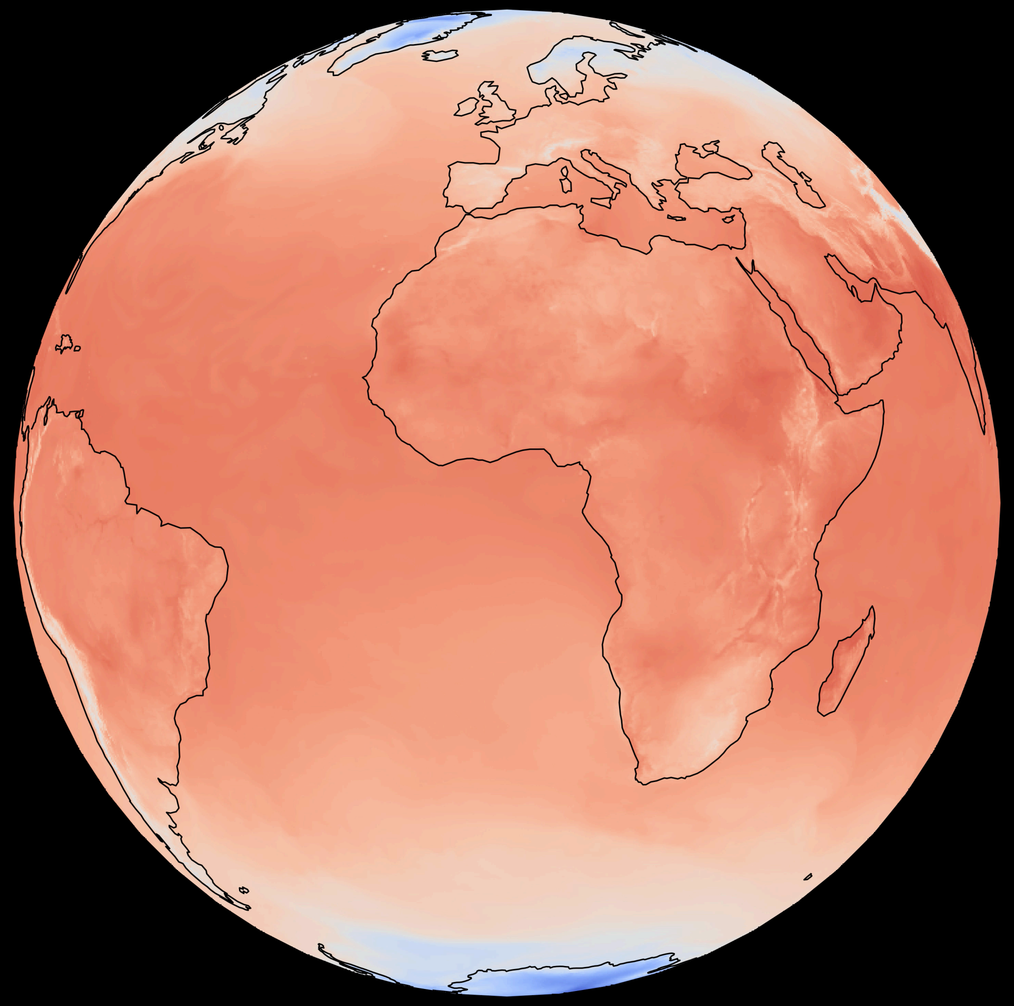
Atmospheric comp.
and air pollution



Ocean
waves



Tropical cyclone
tracks



High-resolution
weather

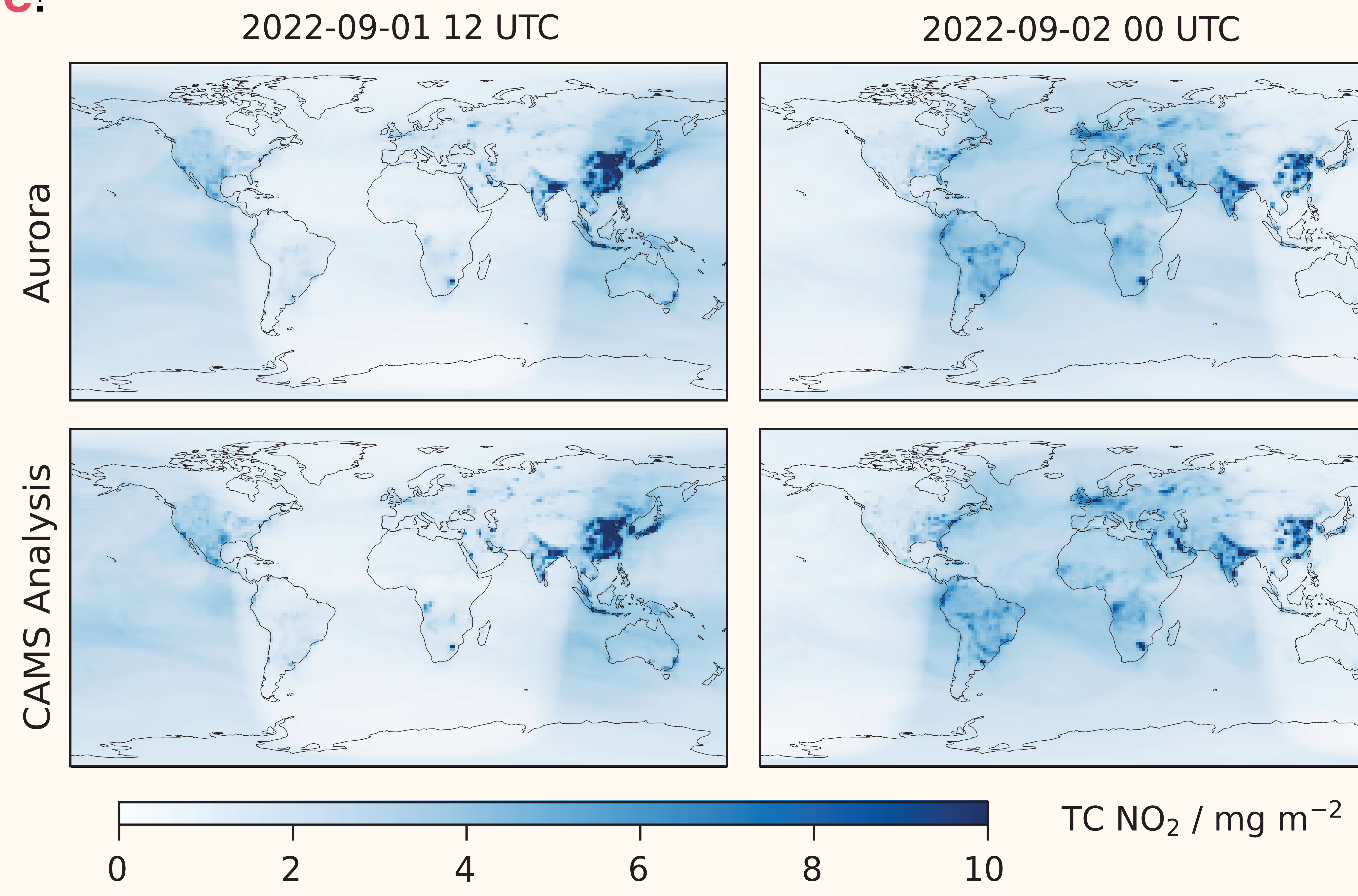
Animations created by Megan Stanley

Air Pollution Forecasting

Coupled to IFS, ~10x more expensive:
~16 node-hours per hour lead time!

- **Setup:** model PM_1, $PM_{2.5}$, PM_{10}, CO, NO, NO_2, SO_2, O_3
- **Data:** Copernicus Atmospheric Monitoring Service (CAMS) analysis
- **Baseline:** CAMS forecasts

Aurora: **~0.5 s per hour lead time**



Overall:

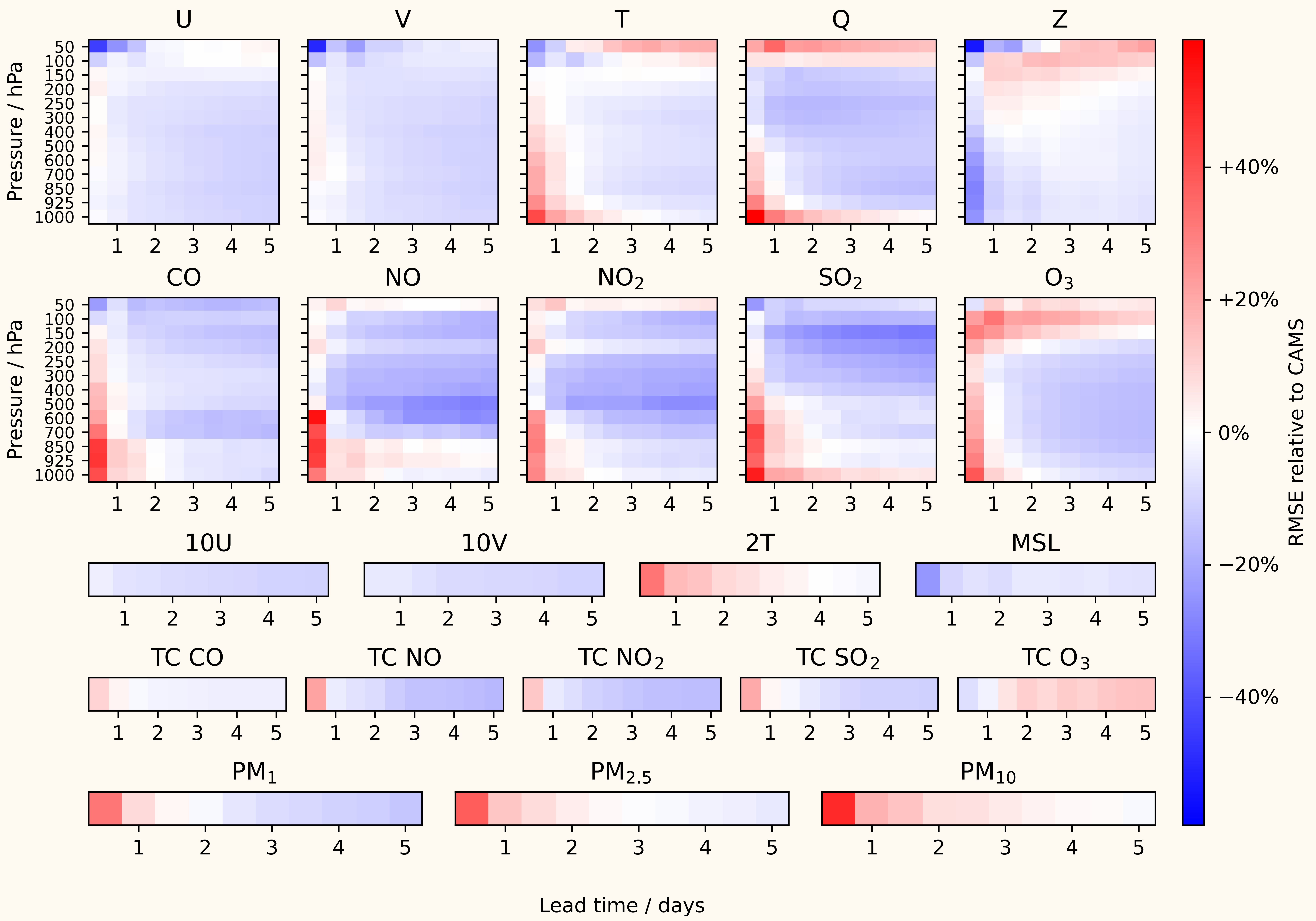
Competitive on 95%
(≤ 20% RMSE)

Better on 75%

Three days:

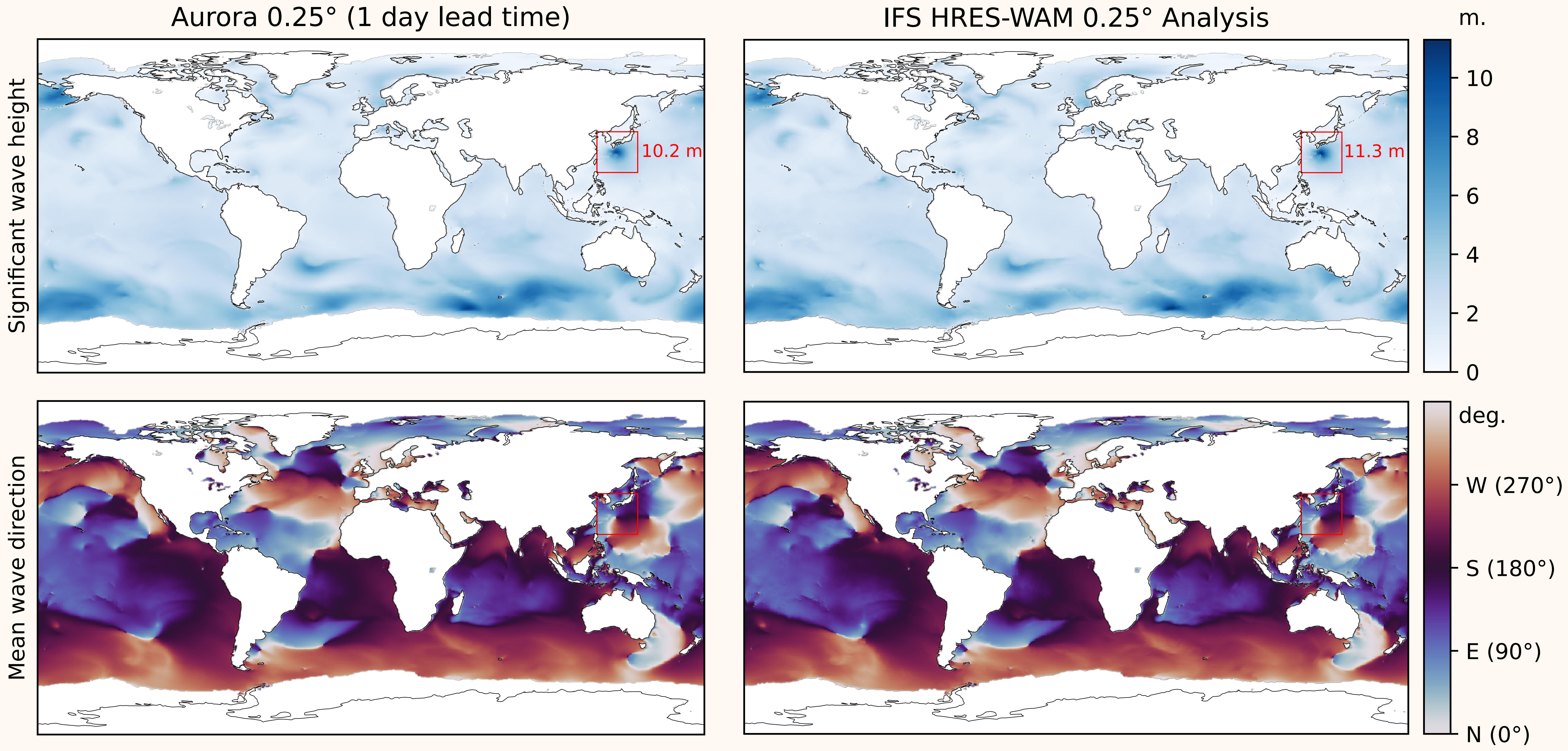
Competitive on 100%
(≤ 20% RMSE)

Better on 86%



Ocean Wave Forecasting

- **Setup:** model height, direction, and period of wave components
- **Data:** HRES-WAM analysis
- **Baseline:** HRES-WAM forecasts



Overall:

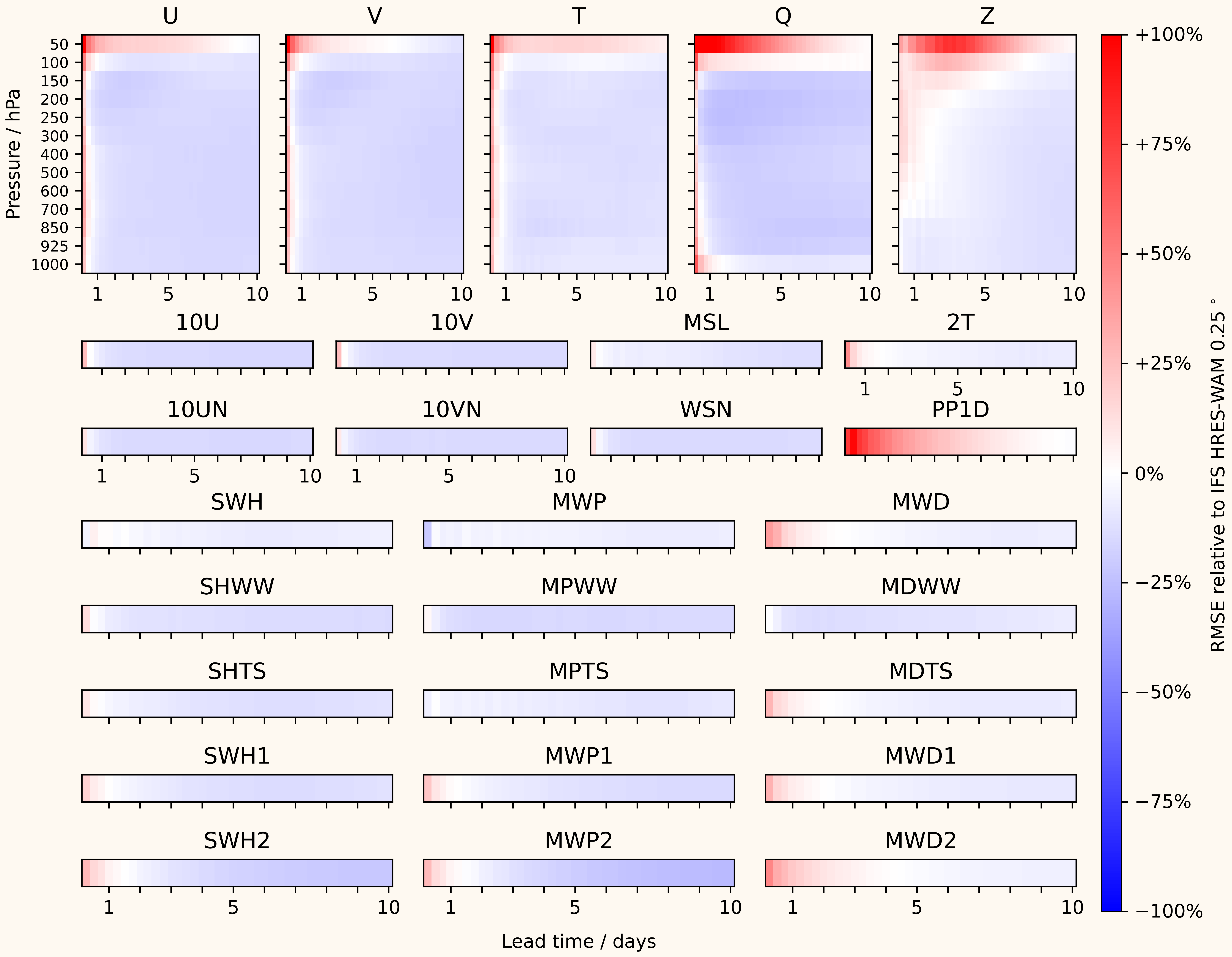
Competitive on 96%
(≤ 20% RMSE)

Better on 86%

Three days:

Competitive on all but PP1D
(≤ 20% RMSE)

Better on 91%



Open Source

- All models open source under MIT licence!
- Details docs with examples

```python
import torch

from aurora import Aurora, Batch, rollout

model = Aurora()
model.load_checkpoint()

model.eval()
model.to("cuda")

batch = Batch(...)

with torch.inference_mode():
    for prediction in rollout(model, batch, steps=10):
        ...  # Do something with `prediction`.

```

https://github.com/microsoft/aurora

pip install microsoft-aurora

The Weather Forecasting Pipeline

The Aurora Team

Paris Perdikaris

University of Pennsylvania, formerly MSR

Wessel Bruinsma

The Alan Turing Institute, formerly MSR

Cristian Bodnar

Silurian, formerly MSR

Ana Lučić

University of Amsterdam, formerly MSR

Megan Stanley

MSR

Richard Turner

U. of Cambridge, The Alan Turing Institute, formerly MSR

Anna Allen

University of Cambridge, The Alan Turing Institute

Johannes Brandstetter

JKU Linz, Emmi AI, formerly MSR

Patrick Garvan

IONQ, formerly MSR

Maik Riechert

MSR

Max Welling

University of Amsterdam, CuspAI, formerly MSR

Elizabeth Heider

Book tour, formerly MSR

Conclusion

- Medium-term weather forecasting has seen incredible progress
- **Pretraining–fine-tuning paradigm** to extend these advancements to other domains
- Aurora only scratches the surface!

wessel.ai/pdf/aurora
wessel.ai/pdf/aardvark
hi@wessel.ai

Bodnar, C., Bruinsma, W.P., Lučić, A., Stanley M., Allen, A. *et al*. A foundation model for the Earth system. *Nature* **641**, 1180–1187 (2025).

Allen, A., Markou, S. *et al*. End-to-end data-driven weather prediction. *Nature* **641**, 1172–1179 (2025).