Aurora: A Foundation Model for the Earth System

Wessel Bruinsma

Microsoft Research AI for Science

Building a Dutch AI-Earth System Modeling Community
KNMI, De Bilt, 29 Jan 2025

The Aurora Team

Paris Perdikaris
University of Pennsylvania, formerly MSR

Wessel Bruinsma
MSR

Megan Stanley
MSR

Ana Lučić
University of Amsterdam, formerly MSR

Cristian Bodnar
Silurian, formerly MSR

Richard Turner
University of Cambridge, formerly MSR

Anna Vaughan
University of Cambridge

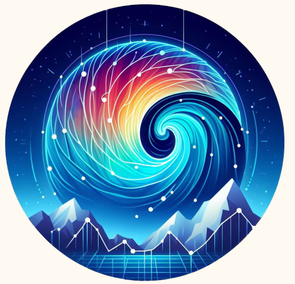
Johannes Brandstetter
JKU Linz, NXAI, formerly MSR

Patrick Garvan
Formerly MSR

Maik Riechert
MSR

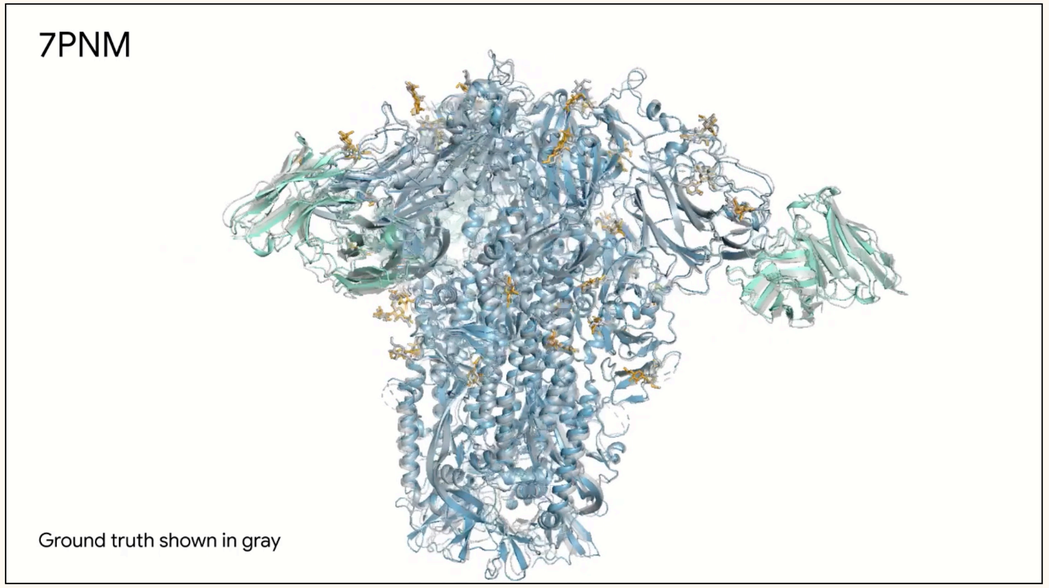
Max Welling
University of Amsterdam, CuspAI, formerly MSR

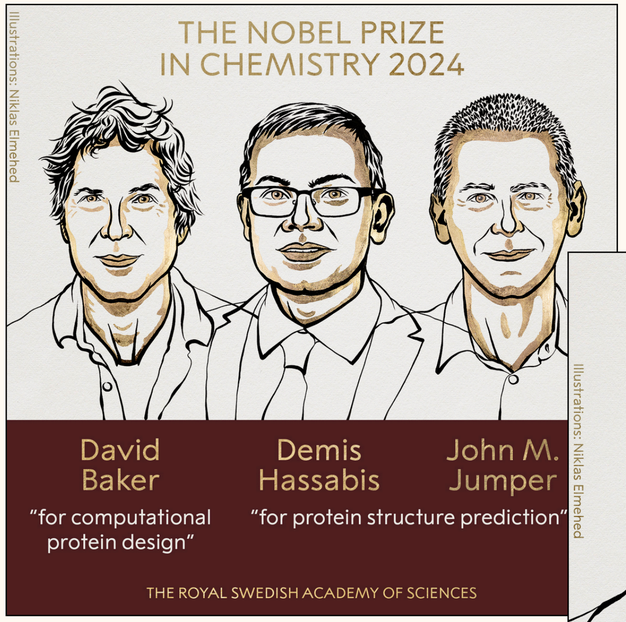
Elizabeth Heider
Book tour, formerly MSR

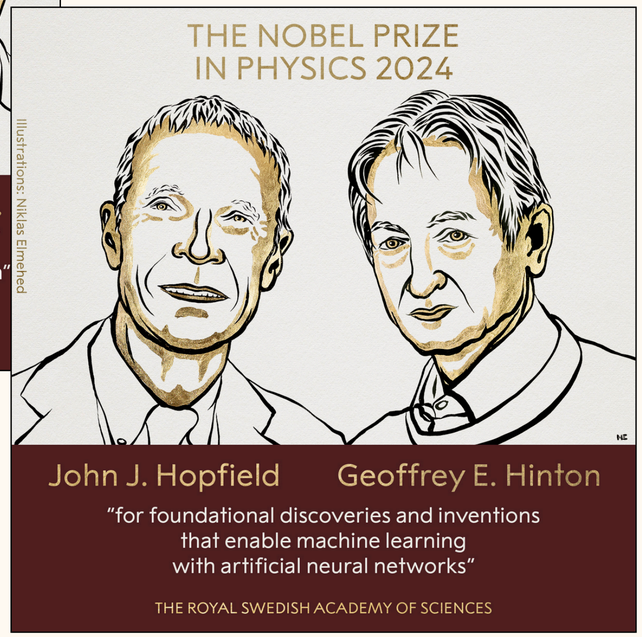


The AI Revolution in Science

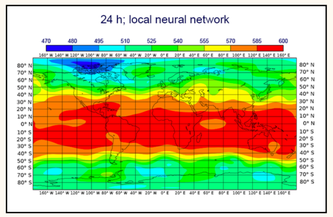
AlphaFold
Protein folding







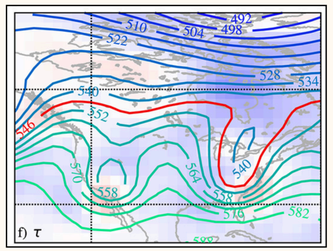
The AI Revolution in Weather Forecasting



2018

First serious efforts to compare AI models to physics baselines

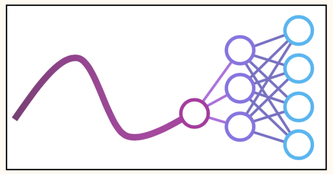
Dueben and Bauer (2018)



2019

AI models skillful to multiple days

Weyn et al. (2019)

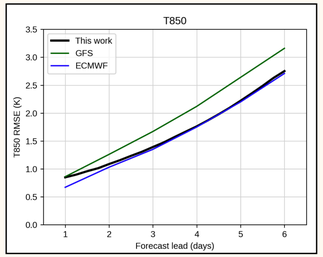


2020

WeatherBench starts to drive ML development

Rasp et al. (2018)

The AI Revolution in Weather Forecasting



HUAWEI

2022

GNN outperforms GFS at 1°

Keisler (2022)

2022

Pangu-Weather outperforms HRES at 0.25°

Bi et al. (2023)

The AI Revolution in Weather Forecasting

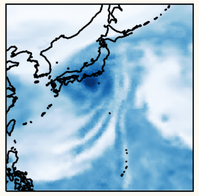
NVIDIA

Microsoft

Google DeepMind

2022–2023

Tech companies start to work in this space



2023

GenCast outperforms IFS ensemble

Price et al. (2024)

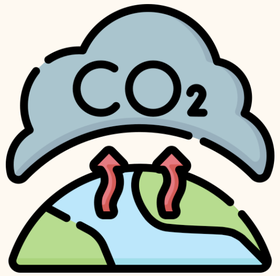
ECMWF

2024

ECMWF launches AIFS

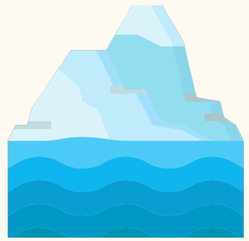
What About Other Forecasting Tasks?

Air pollution



Atmospheric composition

Waves

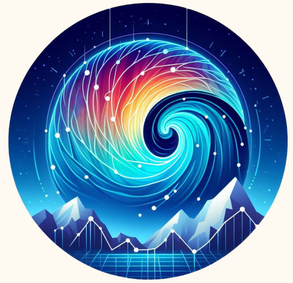


Sea ice

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

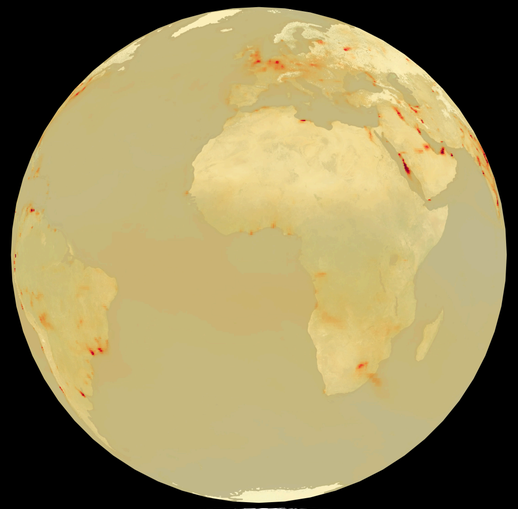
fine-tuning

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

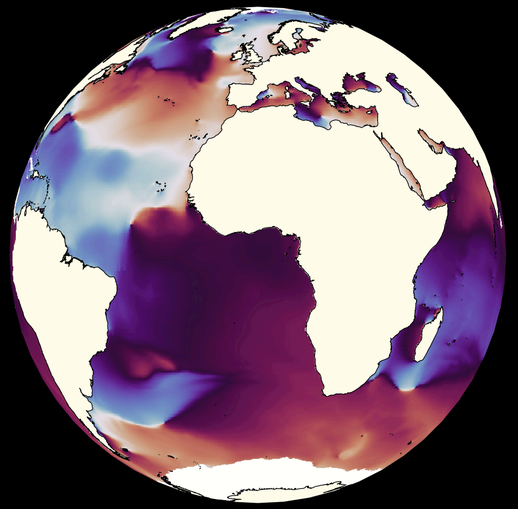
Aurora: a **foundation model** for the Earth system

Aurora

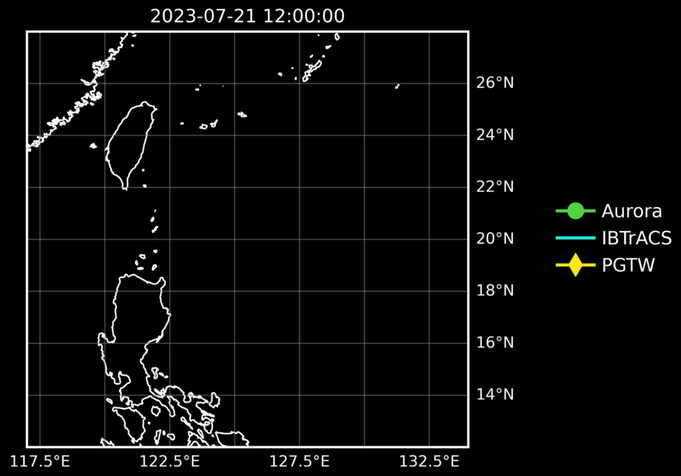
Fine-Tuning Applications



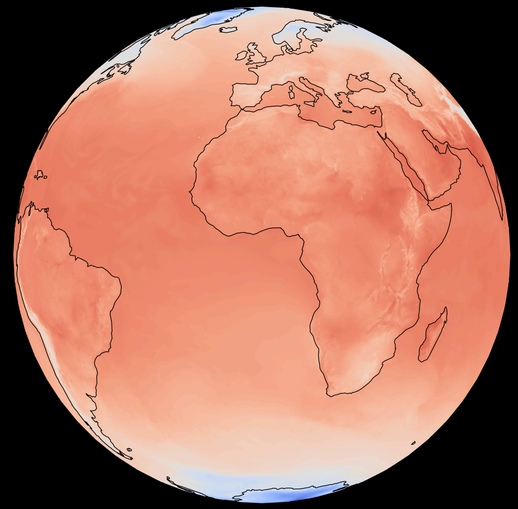
Atmospheric comp. and air pollution



Ocean waves



Tropical cyclone track

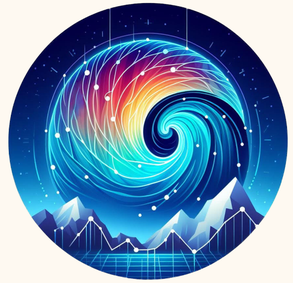


High-resolution weather

Animations created by Megan Stanley

Aurora

Air Pollution Forecasting



- **Setup:** model levels of 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

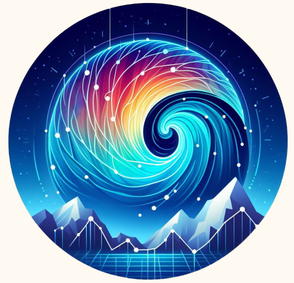
Coupled to IFS, ~10x more expensive:

~16 node-hours per hour lead time!

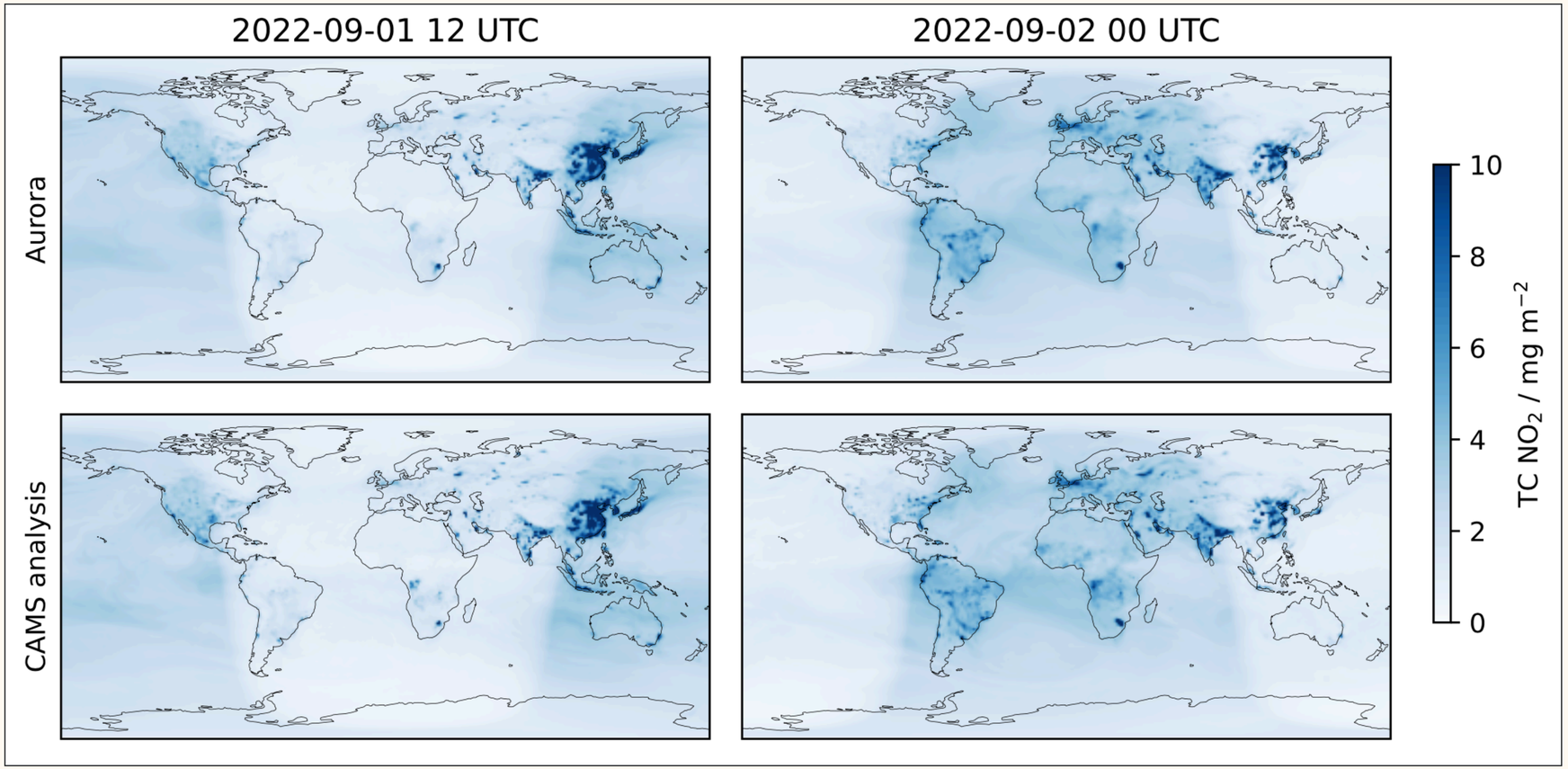
Aurora: **~0.5 s per hour lead time**

Aurora

Air Pollution Forecasting (2)



- Heterogeneous and spikey
- Anthropogenic factors
- Scarce
- Non-stationary



Overall:

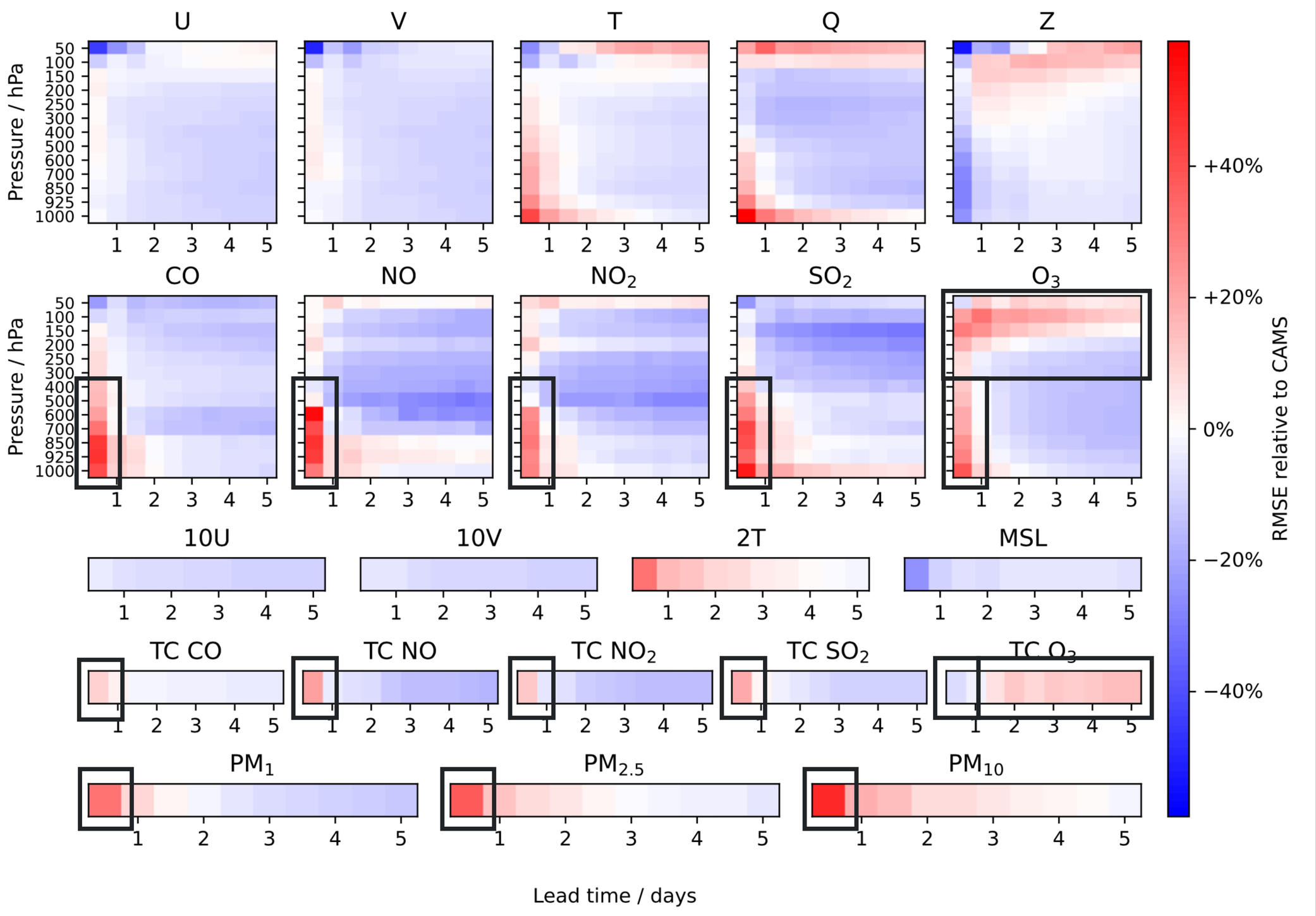
Competitive on 95%
(≤ 20% RMSE)

Better on 75%

Three days:

Competitive on 100%
(≤ 20% RMSE)

Better on 86%



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
hi@wessel.ai