Earth-2: Digital Twins for Weather and Climate

Karthik Kashinath, Principal Engineer and Scientist, AI-HPC

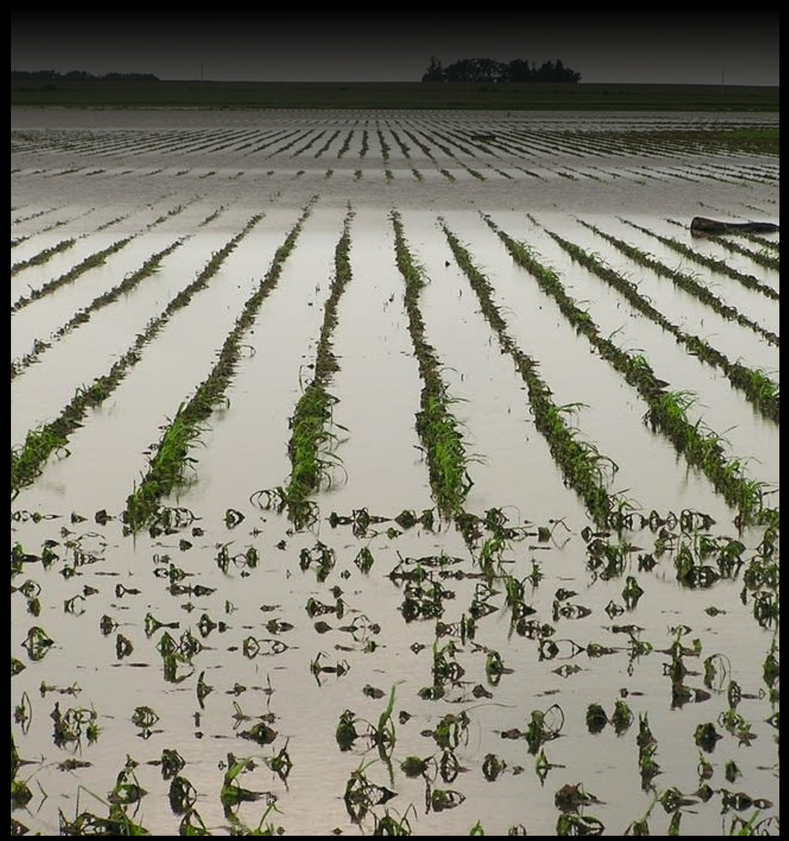
Engineering Lead, Earth-2 Initiative

The Future Under Climate Change will be Harsh

We urgently need better tools to prepare for it

WILDFIRE PREVENTION

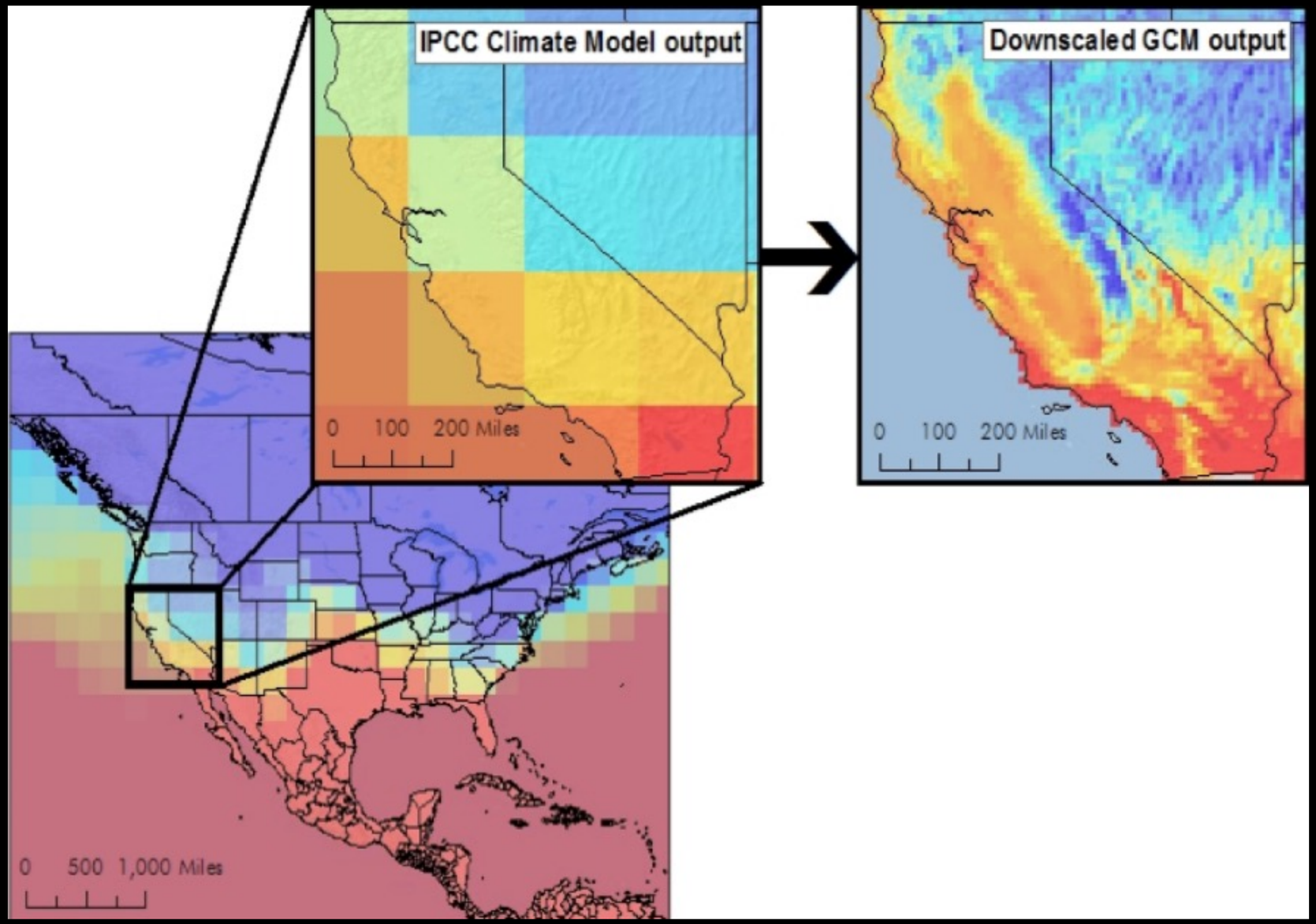
WATER MANAGEMENT



CROP FORECASTING

Today's Climate Predictions are too Low Resolution

Detailed Plans require Detailed Information and Predictions with Credible Cloud Feedbacks



High Resolution Climate Prediction is a Computational Challenge

Today's climate models are too low resolution. Brute force PDE solvers are decades away from what is needed.

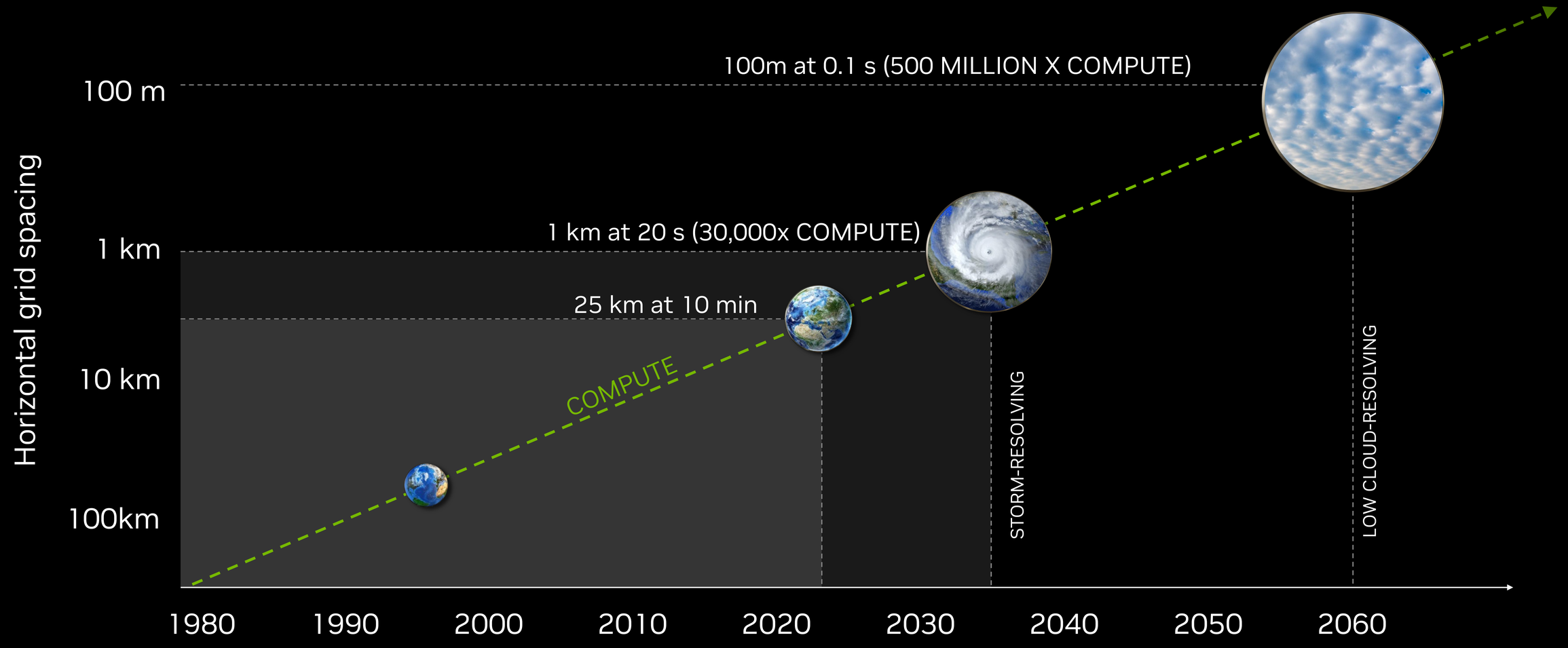
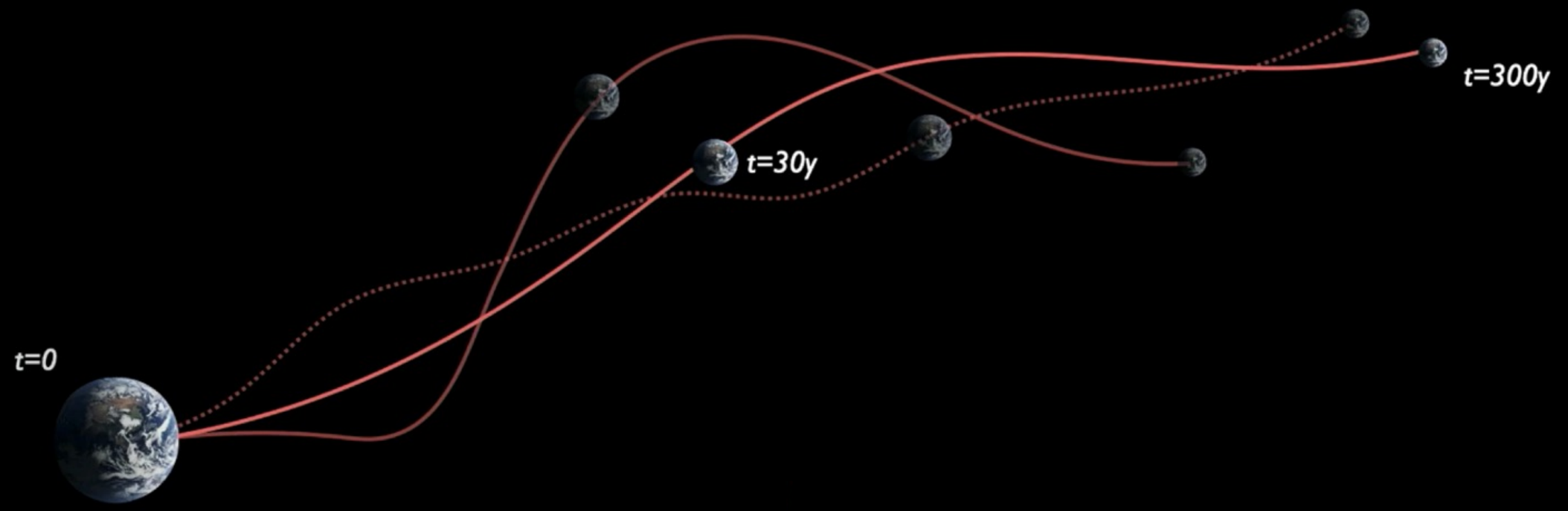


Figure adapted from: Schneider, T., Teixeira, J., Bretherton, C. *et al.* *"Climate goals and computing the future of clouds". Nature Climate Change* **7,** 3–5 (2017)

It is Hard to Interact with High Resolution Climate Prediction Data.

"We can compute km-scale predictions, but can't effectively extract information content, let alone interact with it"

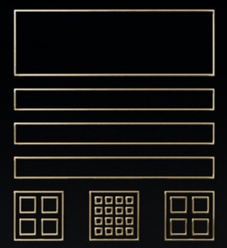
-- Bjorn Stevens.



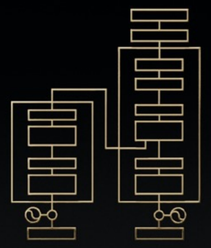
Advances in High-Performance Computing and Machine Learning

Promise million-X speedups

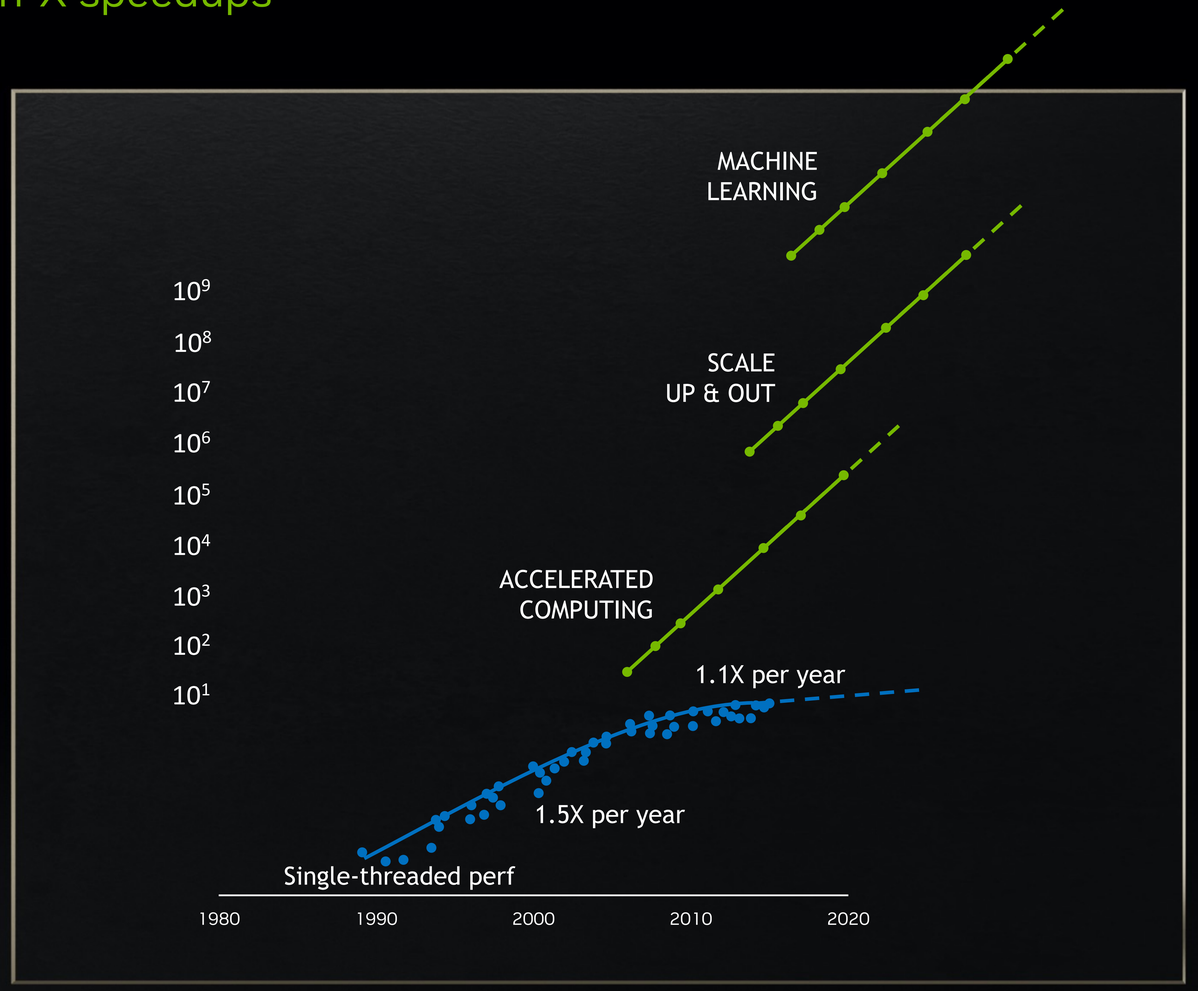
Accelerated Computing



AI

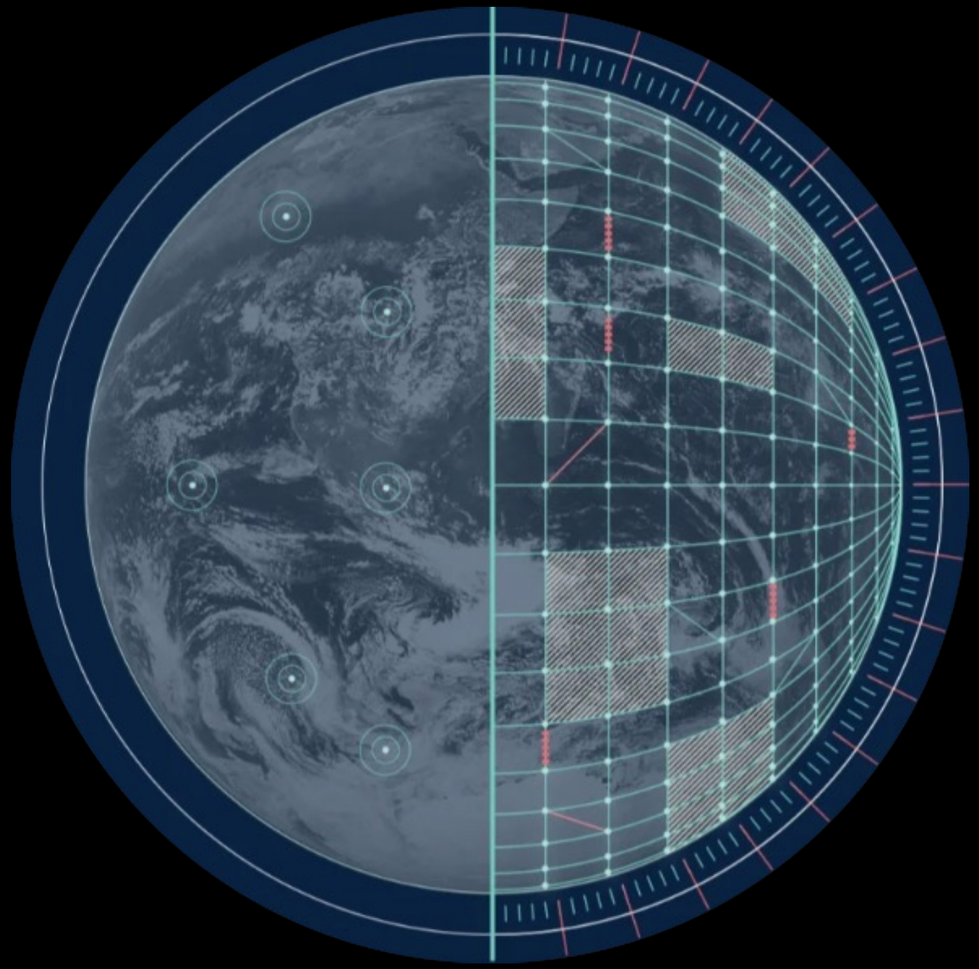


Data Center Scale

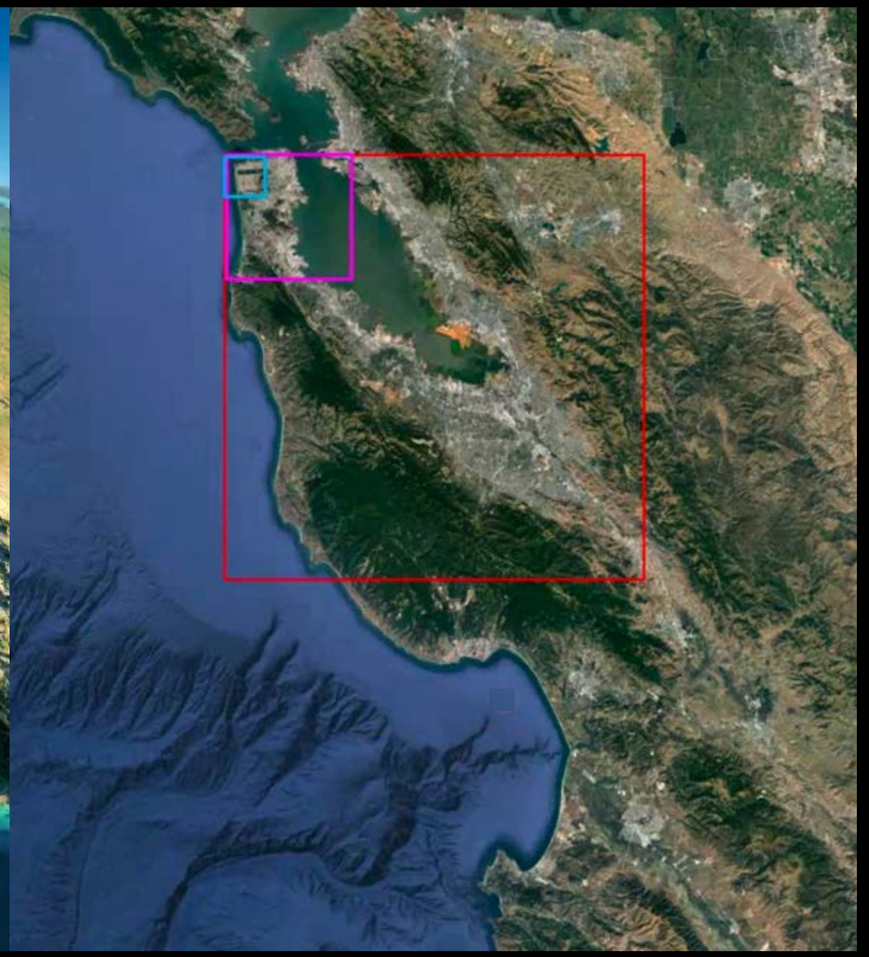


EARTH-2: Begins as a vision

Of a highly interactive climate information system for serving society with next-gen climate predictions.

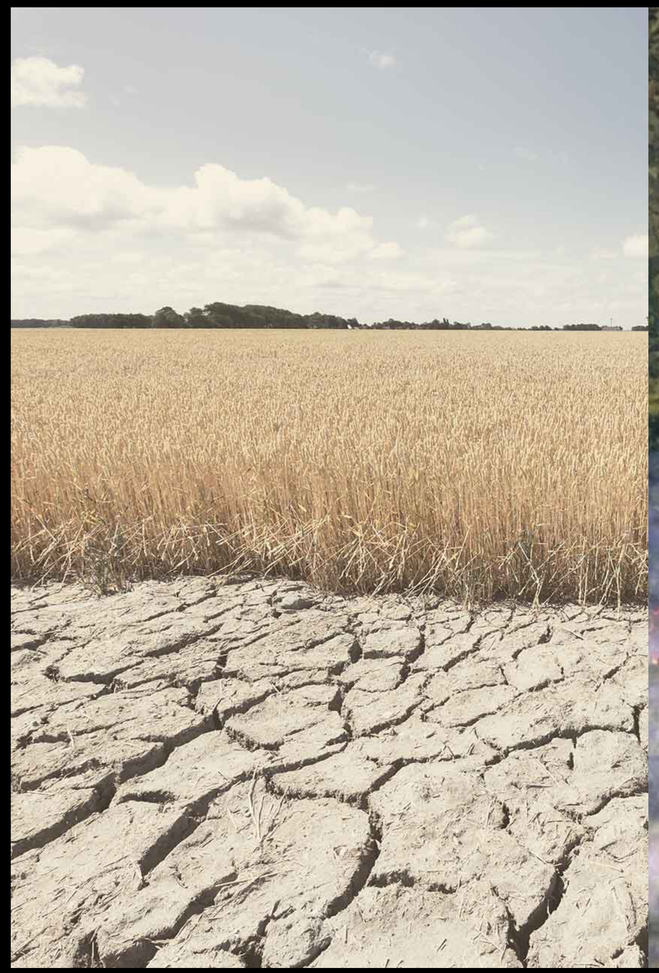


Imagine you could Select a Region of the Planet...



... Ask Questions about Climate Change's Impacts...

On Food, Health, Infrastructure, Energy systems, and more...




… and Receive Useful, Visual & Statistical Guidance?

From a Highly Interactive Future Climate Information System, at High Resolution, that Serves Society…

Earth-2 Mission #1

Interacting with Climate Predictions at Low Latency.

Imagine the System Evolving in Computational Fidelity and Ambition

Eventually fed by a new library of climate predictions so high-resolution they seem impossible today.

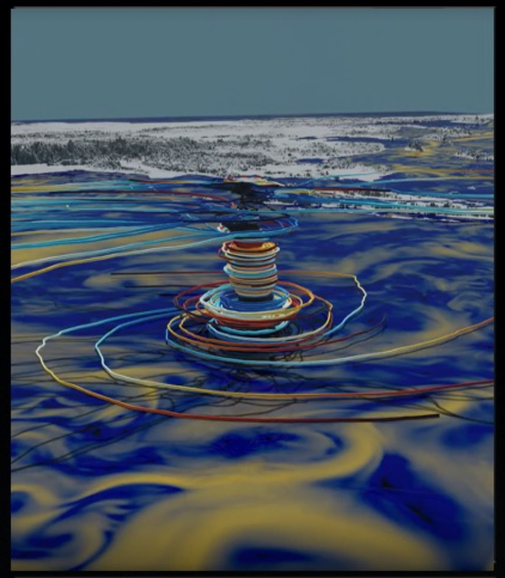
Earth-2 Mission #2

Achieving Next-Gen Climate Predictions using Hybrid Physics, Machine Learning & HPC.

NVIDIA's technical know-how can make a big difference

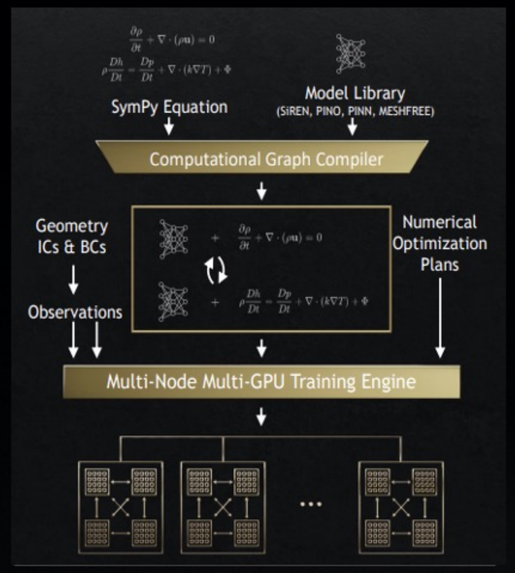
Earth-2 harnesses NVIDIA's full-stack technologies to make Earth digital twins a reality

Interactive Collaborative Platform



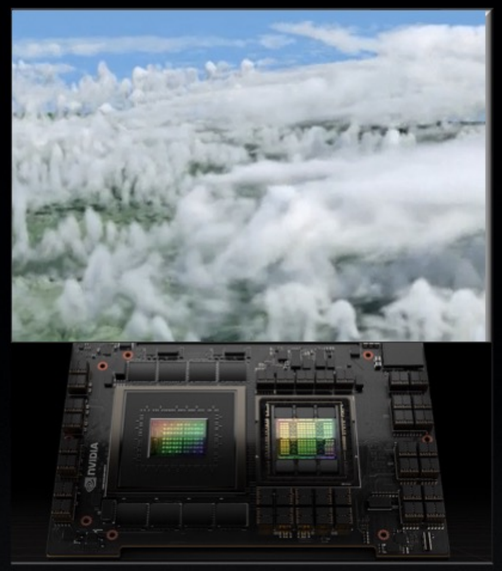
OMNIVERSE

Data-driven Models



PHYSICS-ML / MODULUS

Storm-resolving Models



GPU-ACCELERATION

Unified Observations

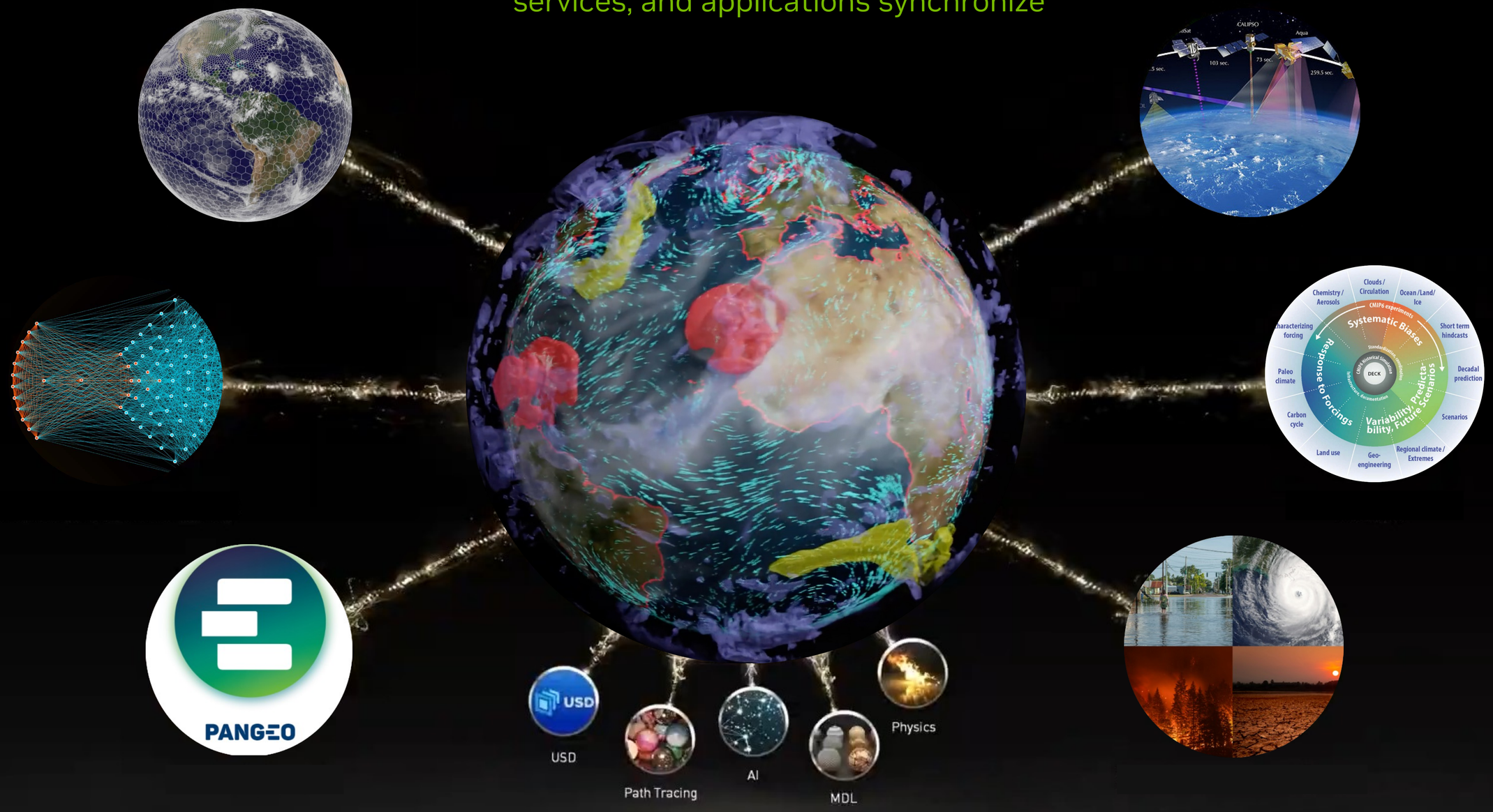
OMNIVERSE NUCLEUS

Exascale Compute

OVX SUPERPOD

Omniverse will enable scientists to create digital twins together

Nucleus: A shared space where models, data, tools, services, and applications synchronize



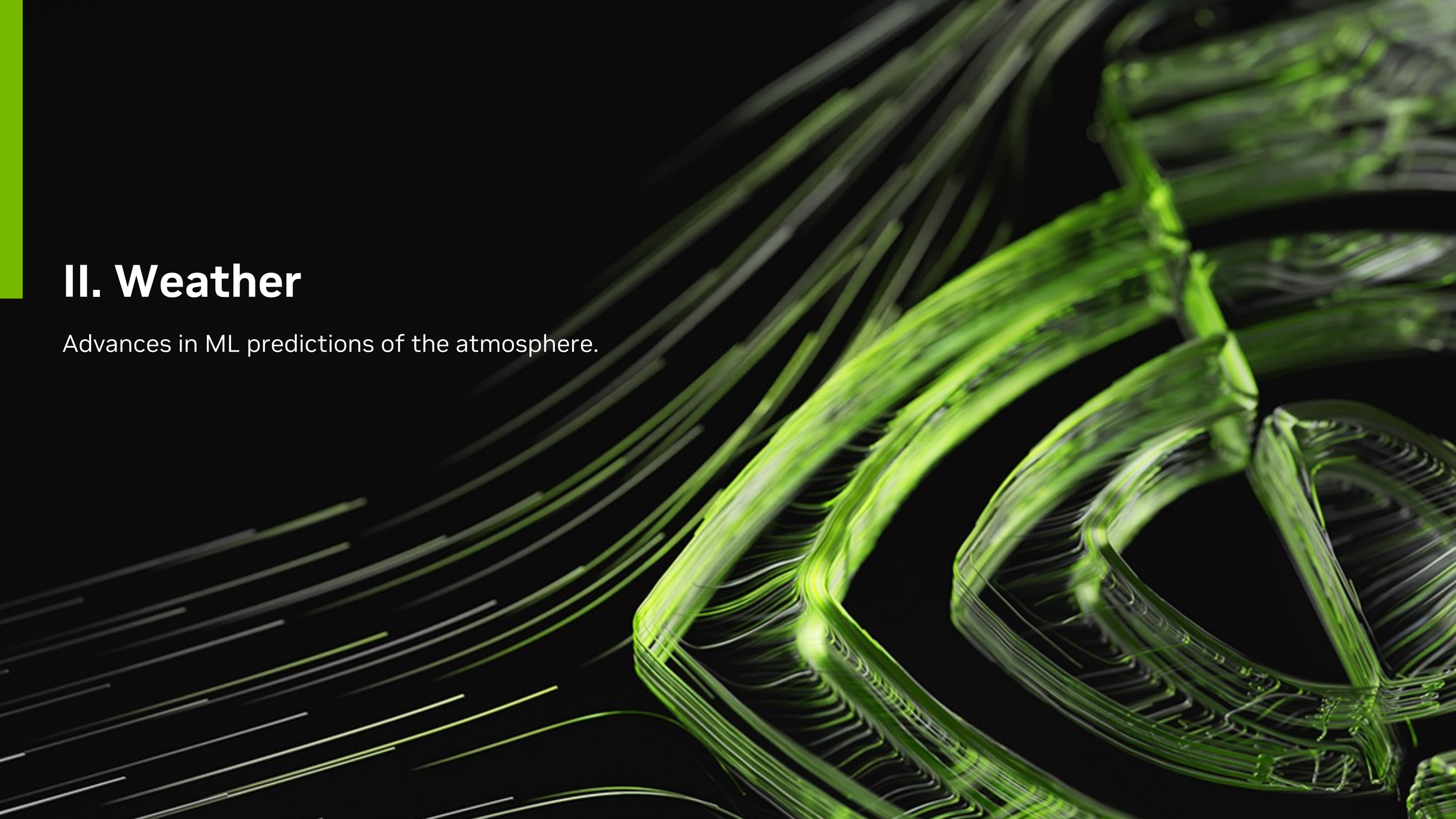
NVIDIA.

Earth-2 will be in Collaboration with International Climate Science

NVIDIA's visualization, engineering, AI & full-stack expertise complement research capacity in academia & government.

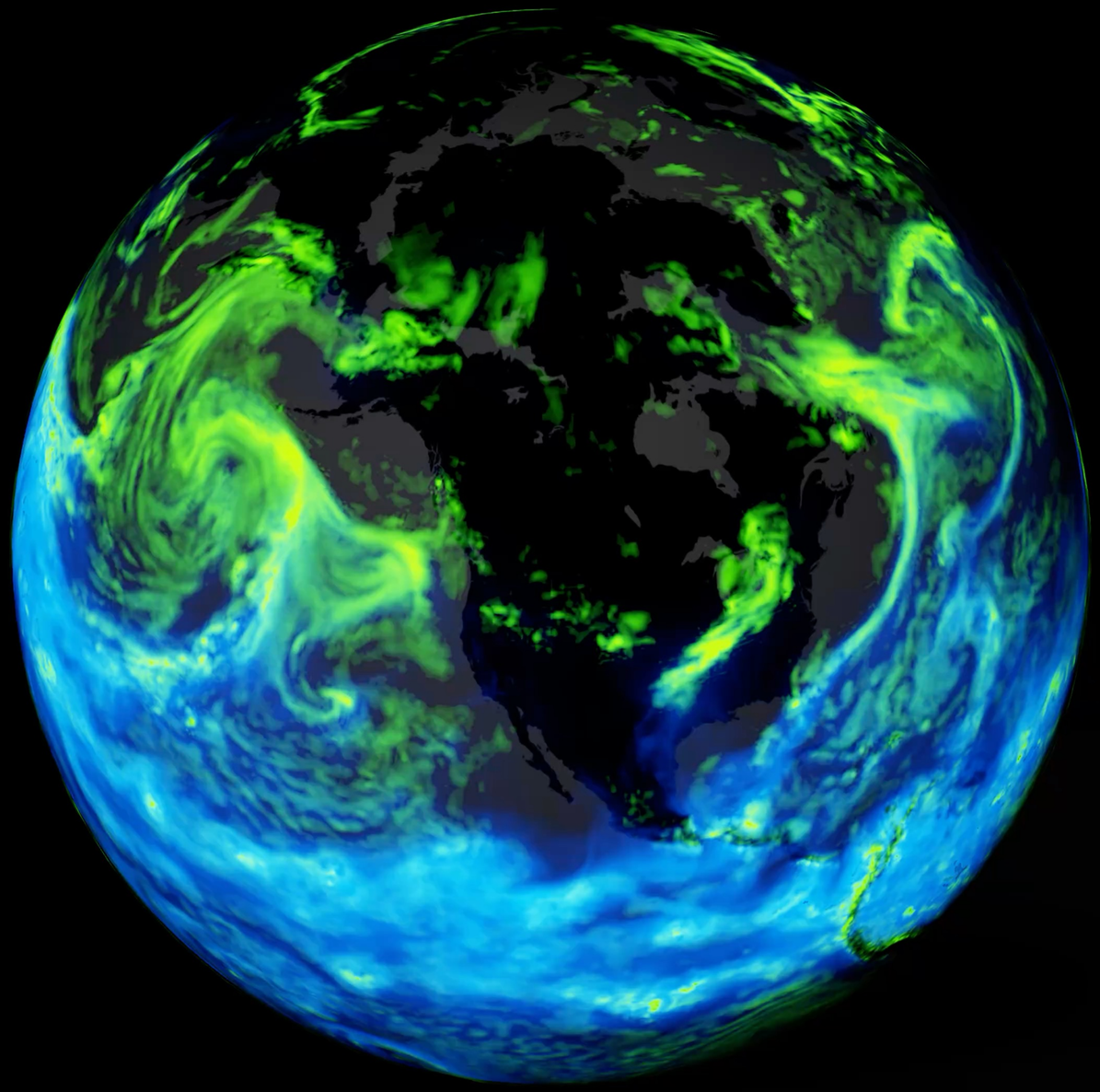
II. Weather

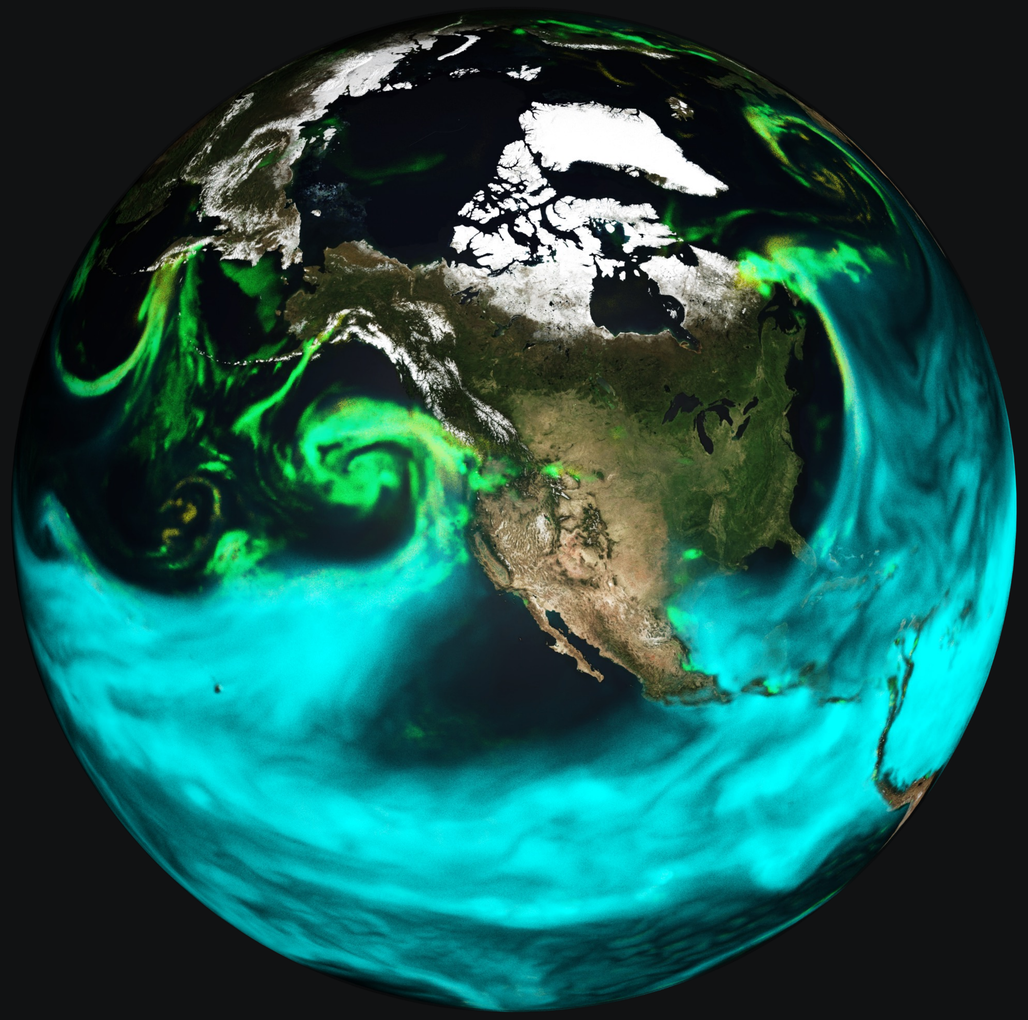
Advances in ML predictions of the atmosphere.



Earth-2 began by exploring the promise of ML weather prediction

ML shows remarkable prediction skill



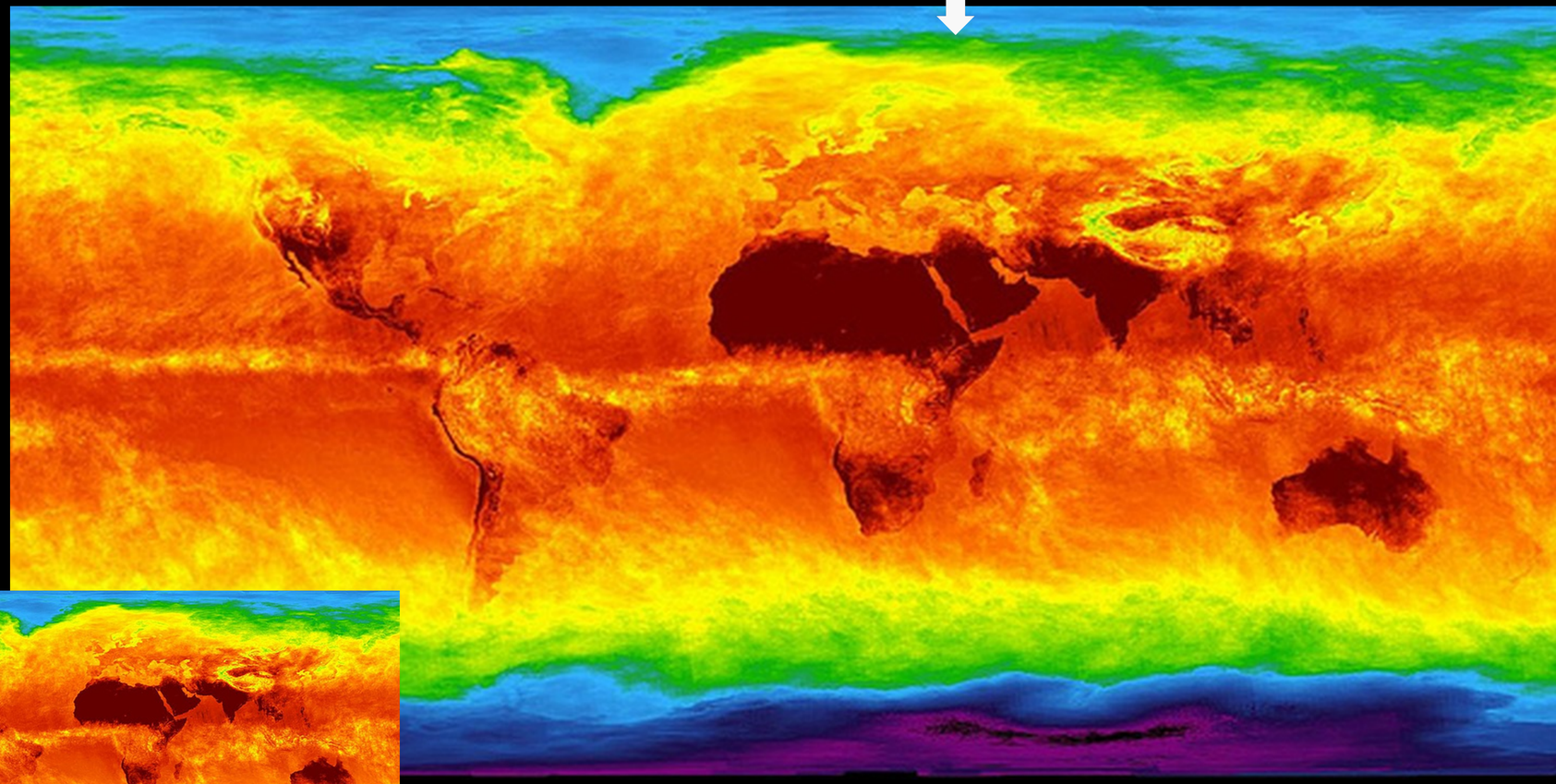


Earth-2 Began by Exploring Data-Driven Weather Prediction

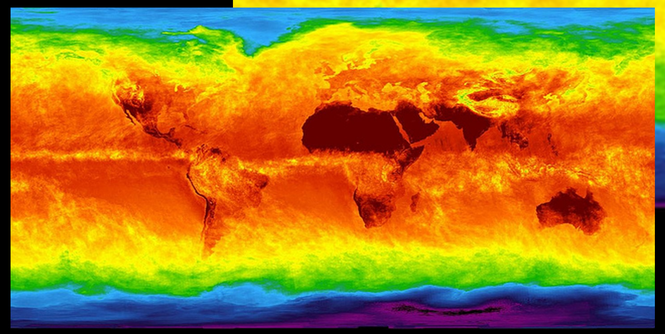
FourCastNet

• Scope	Global, Medium Range
• Model Type	Full-Model AI Surrogate
• Architecture	AFNO (Adaptive Fourier Neural Op.)
• Resolution:	25km
• Training Data:	ERA5 Reanalysis
• Initial Condition	GFS / UFS
• Inference Time	0.25 sec (2-week forecast)
• Speedup vs NWP	$O(10^4\text{-}10^5)$
• Power Savings	$O(10^4)$

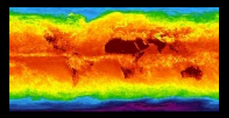
FourCastNet: A new data-driven weather predictor of unprecedented resolution



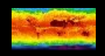
FourCastNet, Pathak et al. (2022), 0.25°, ~1,000,000 Pixels, ViT+AFNO



GNN, Keisler et al. (2022), 1°, 64,000 Pixels, Graph Neural Networks



DLWP, Weyn et al. (2020). 2°, 16K pixels, Deep CNN on Cubesphere/(2021) ResNet



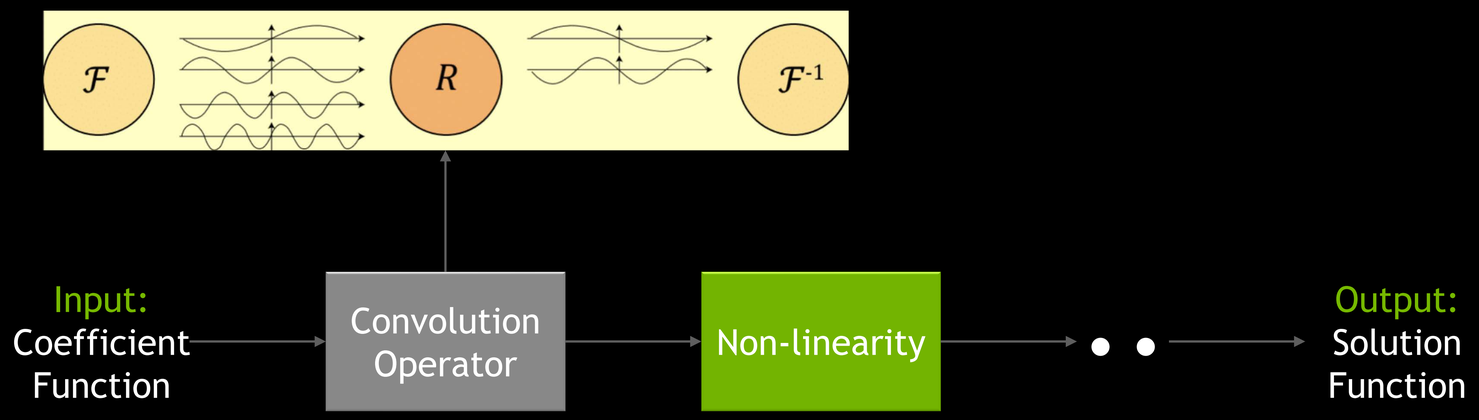
Weyn et al. (2019), 2.5° N.H only, 72x36, 2.6k pixels, ConvLSTM

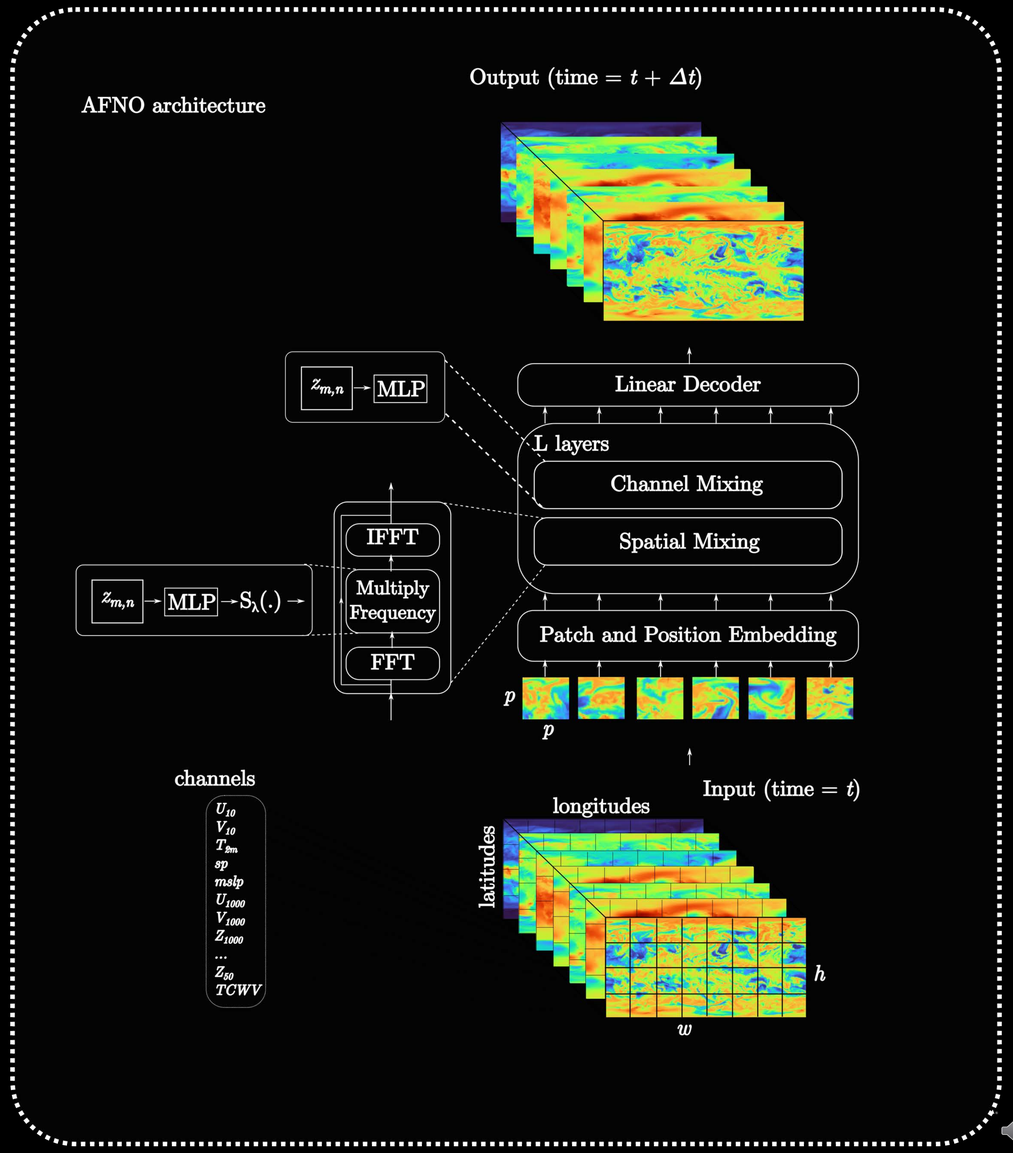
WeatherBench, Rasp et al. (2020). 5.625°, 64x32, 2K pixels, CNN

Deuben & Bauer (2018), 6° , 60x30, 1.8K pixels, MLP

FCN USES A NOVEL TRANSFORMER ARCHITECTURE

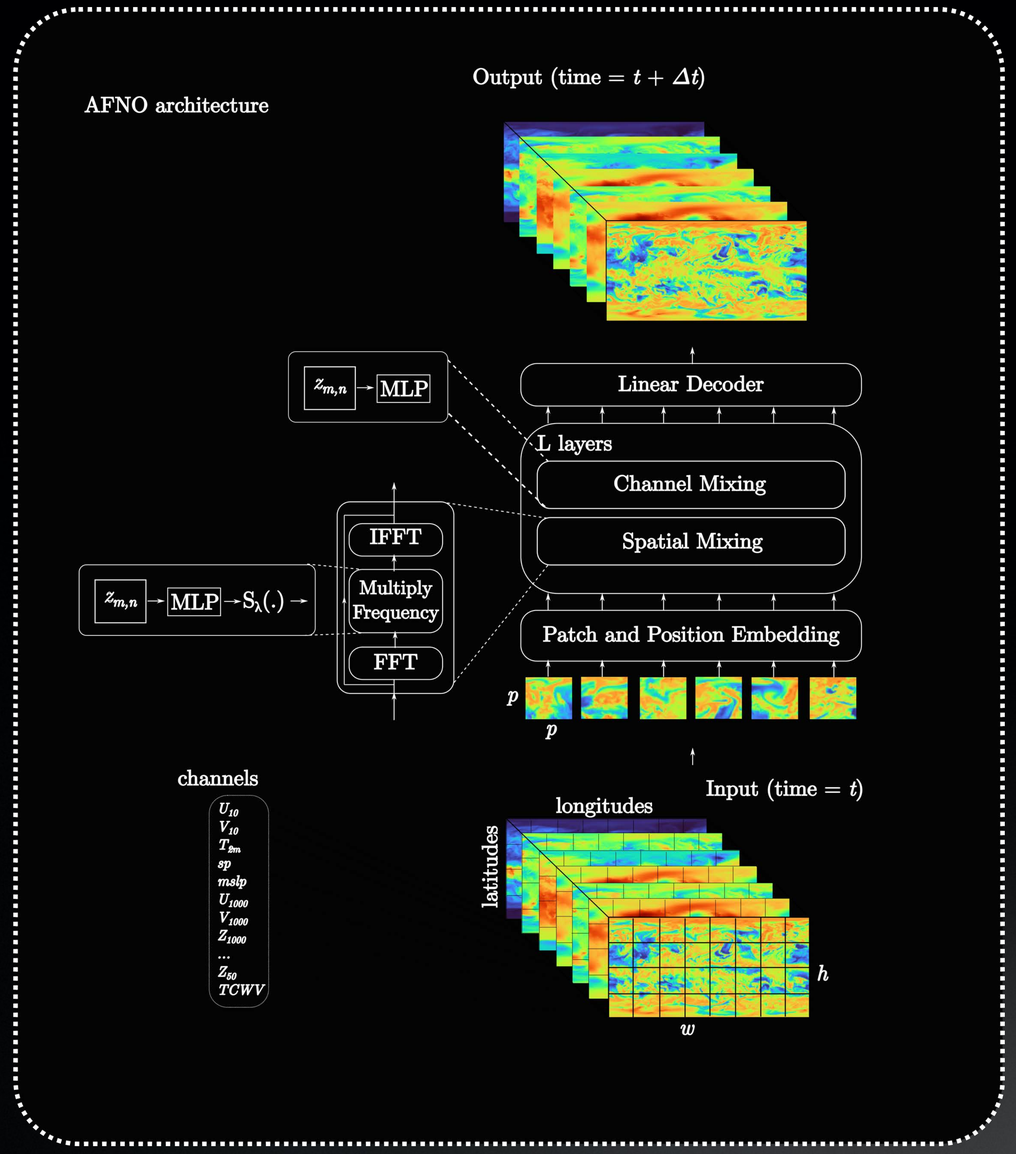
With Fourier Neural Operator Blocks - in search of grid-free, high-resolution, machine-learnt simulations.





Pathak et al. (2022), FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators, arXiv:2202.11214

FOURCASTNET (FOURIER FORECASTING NETWORK)



- Purely data-driven ML surrogate weather model
- Trained on ERA5 reanalysis data at the native resolution of 0.25 degrees

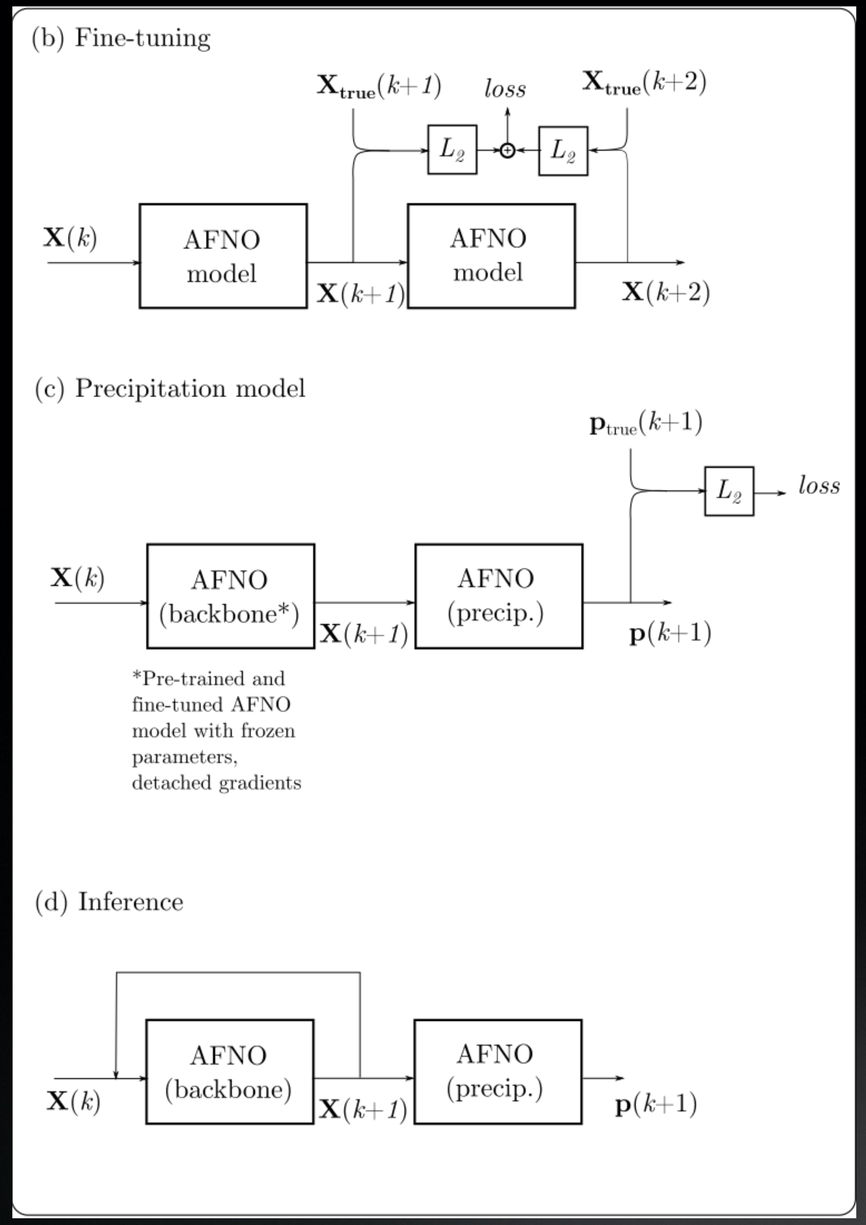
Vertical Level	Variables
Surface	$U_{10}, V_{10}, T_{2m}, sp, mslp$
$1000hPa$	U, V, Z
$850hPa$	T, U, V, Z, RH
$500hPa$	T, U, V, Z, RH
$50hPa$	Z
Integrated	$TCWV$

Extending to include radiation processes, vapor transport, clouds

Training set: 1979 to 2015
Validation set: 2016, 2017
Held out: 2018 onwards

Pathak et al. (2022), FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators, arXiv:2202.11214

FOURCASTNET (FOURIER FORECASTING NETWORK)



- Purely data-driven ML surrogate weather model
- Trained on ERA5 reanalysis data at the native resolution of 0.25 degrees

Vertical Level	Variables
Surface	$U_{10}, V_{10}, T_{2m}, sp, mslp$
$1000hPa$	U, V, Z
$850hPa$	T, U, V, Z, RH
$500hPa$	T, U, V, Z, RH
$50hPa$	Z
Integrated	$TCWV$

Extending to include radiation processes, vapor transport, clouds

Training set: 1979 to 2015
Validation set: 2016, 2017
Held out: 2018 onwards

Pathak et al. (2022), FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators, arXiv:2202.11214

COMPUTATIONAL PERFORMANCE

Latency and Energy consumption for a 24-hour 100-member ensemble forecast				
	IFS	FCN - 30km (actual)	FCN - 18km (extrapolated)	IFS / FCN(18km) Ratio
Nodes required	3060	1	2	**1530**
Latency (Node-seconds)	984000	7	22	**44727**
Energy Consumed (kJ)	271000	7	22	**12318**

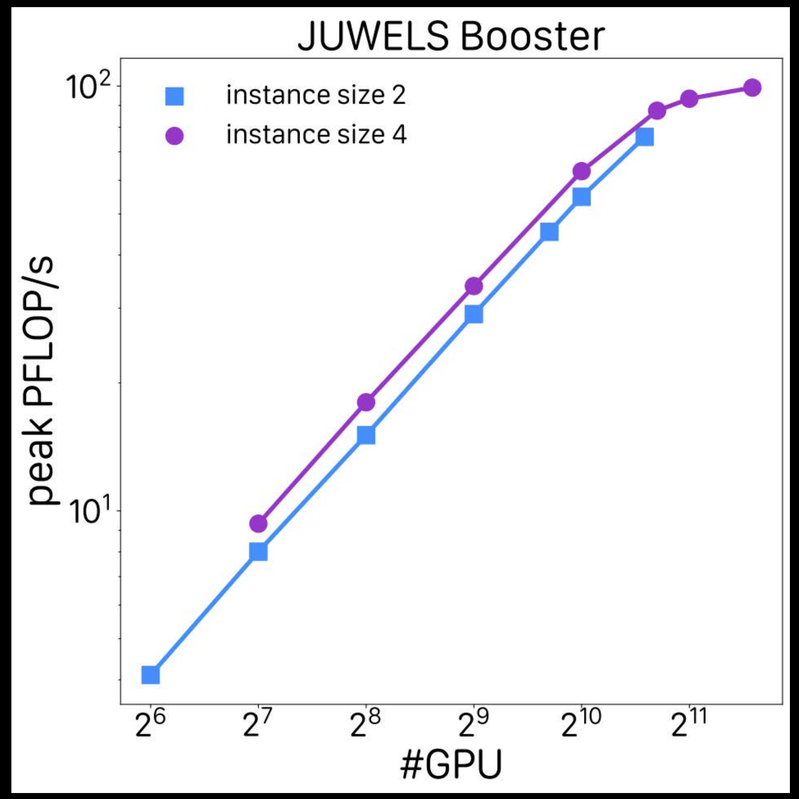
- 100-member ensemble forecast in 7 seconds
- 100-member ensemble forecast consumes 7 kJ
- 4 to 5 orders-of-magnitude speedup over NWP
- 4 orders-of-magnitude smaller energy footprint

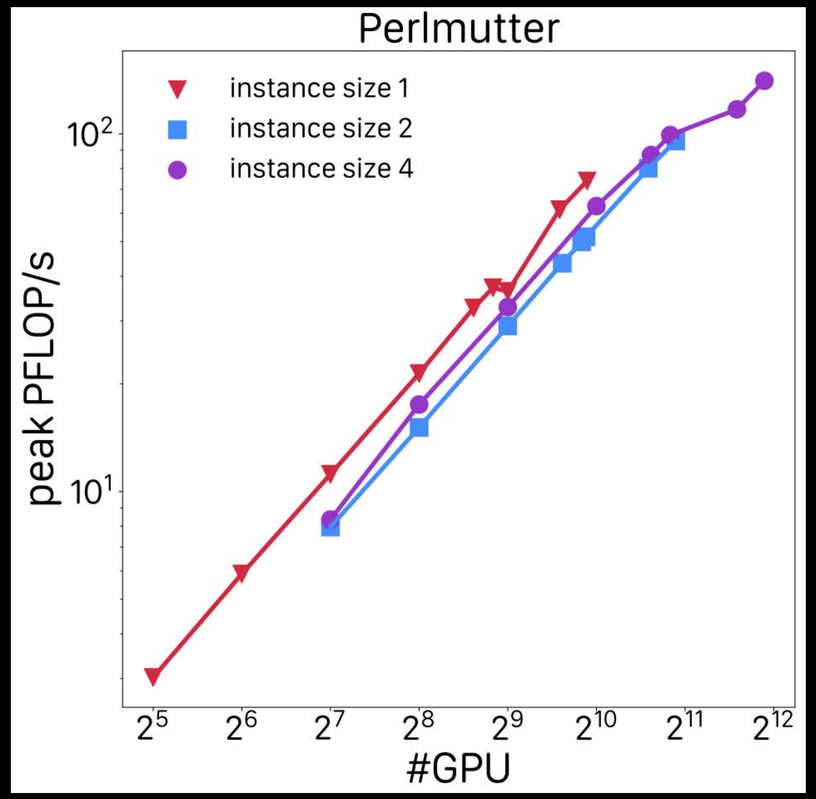
Caveats

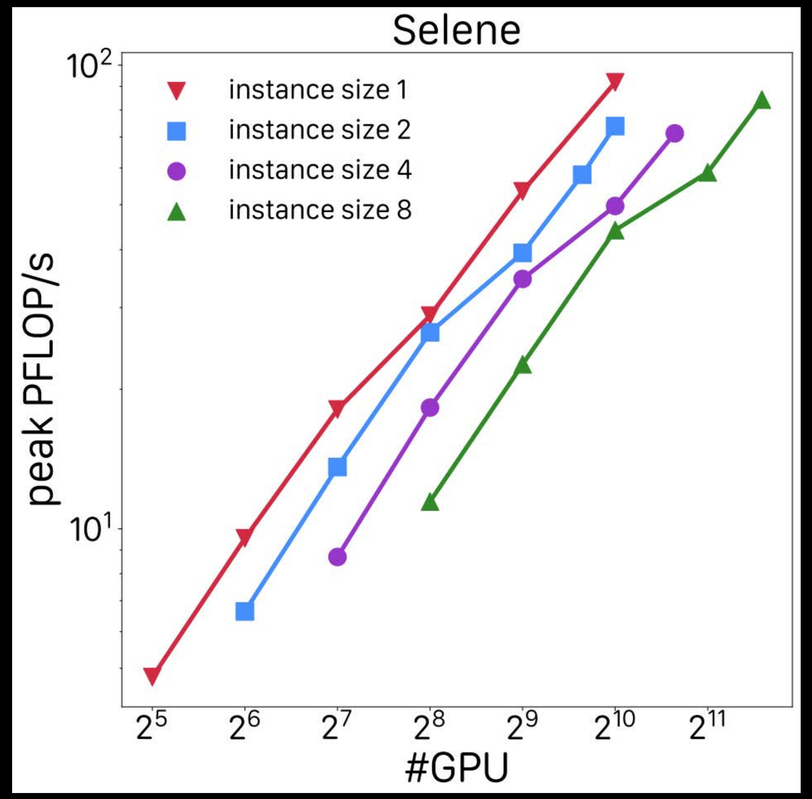
- FourCastNet is *not* physics constrained

We train FCN on ambitious amounts of data on large machines
FCN scales efficiently up to ~ 4000 GPUs on three supercomputing systems

Thanks to full-stack AI + HPC expertise we train on a growing amount of the world's petabytes of past weather data.







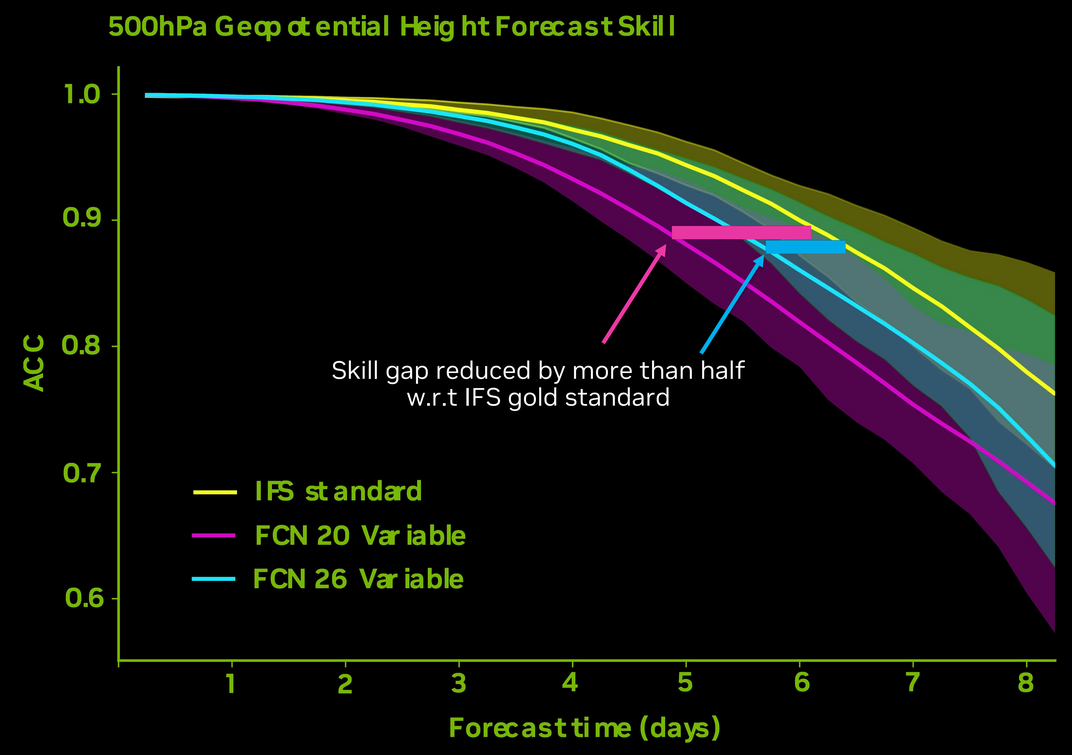
Peak performance is 140.8 petaFLOPS in mixed precision (averaged over a full epoch)

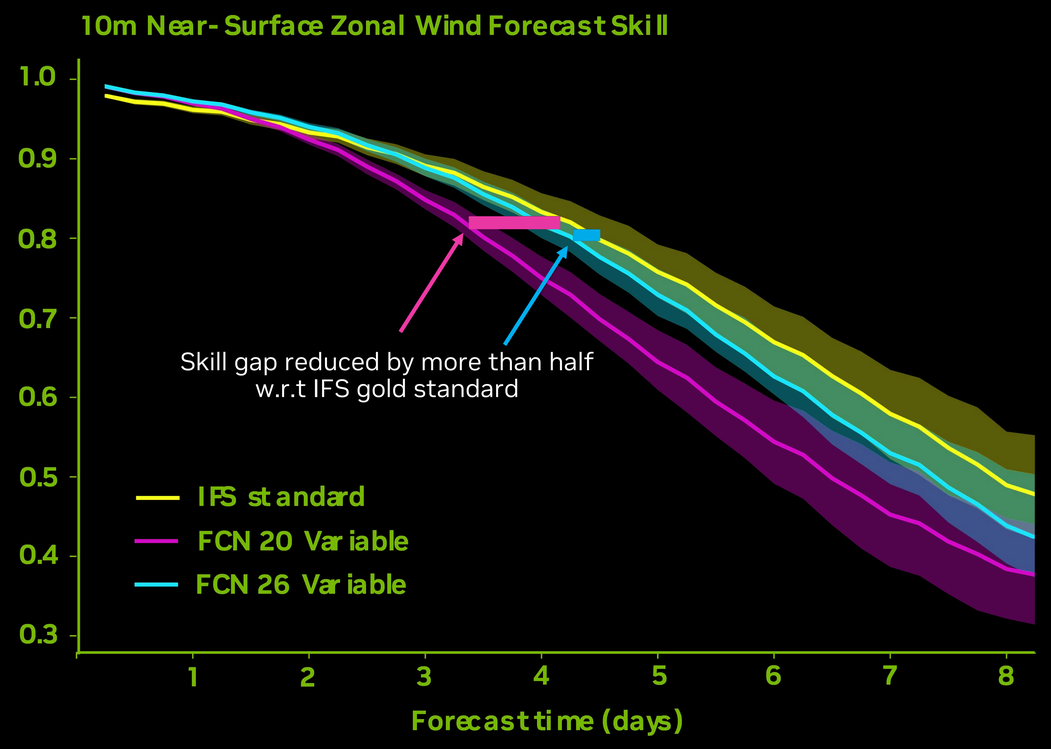
Time to solution decreased from 24+ hours to 67 minutes with model and data parallelism

FourCastNet: Accelerating Global High-Resolution Weather Forecasting using Adaptive Fourier Neural Operators, Kurth et al. (2022), https://arxiv.org/abs/2208.05419

FCN skill improving with training ambition.

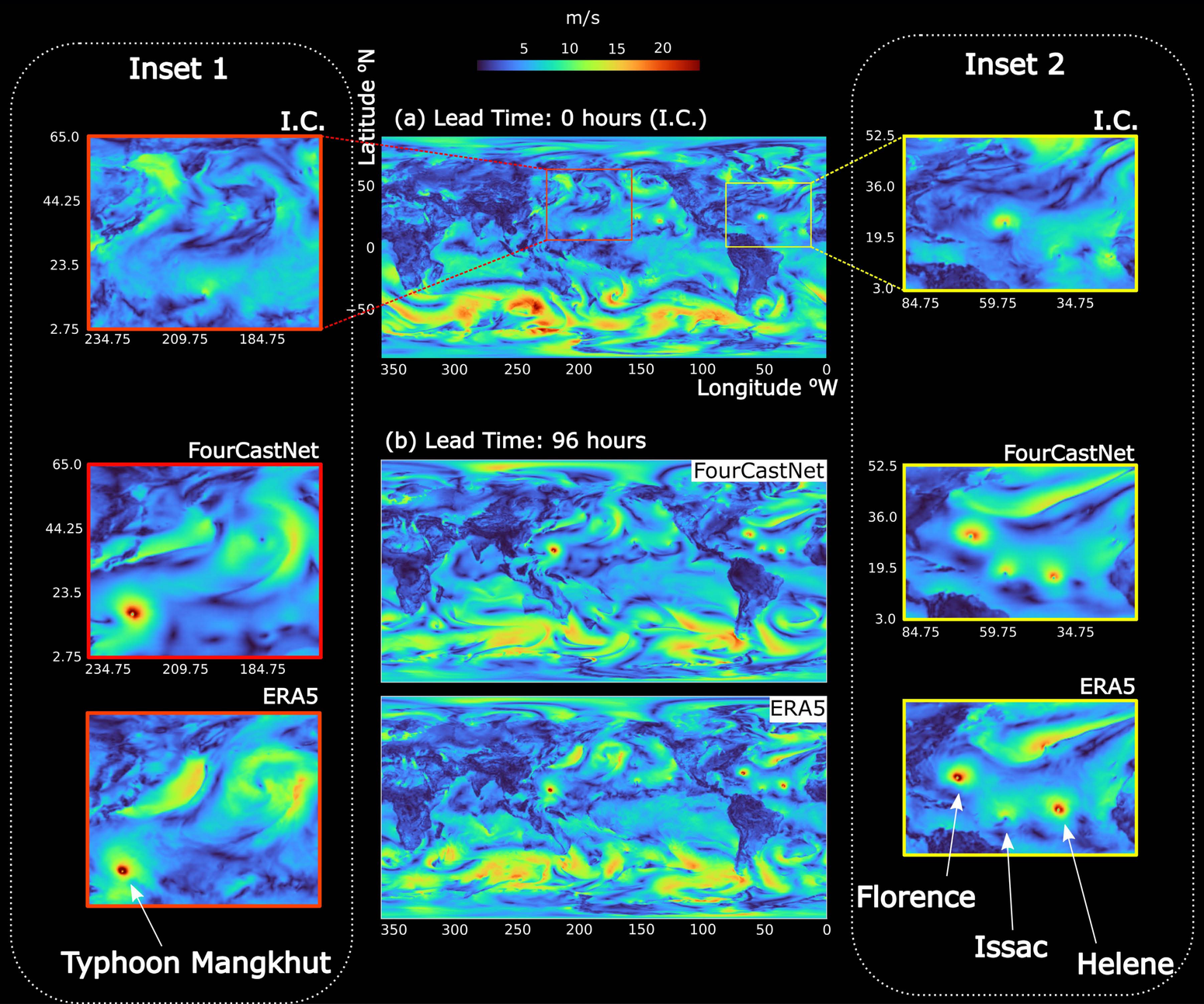
Could it one day outperform deterministic models? We don't yet know the limit.





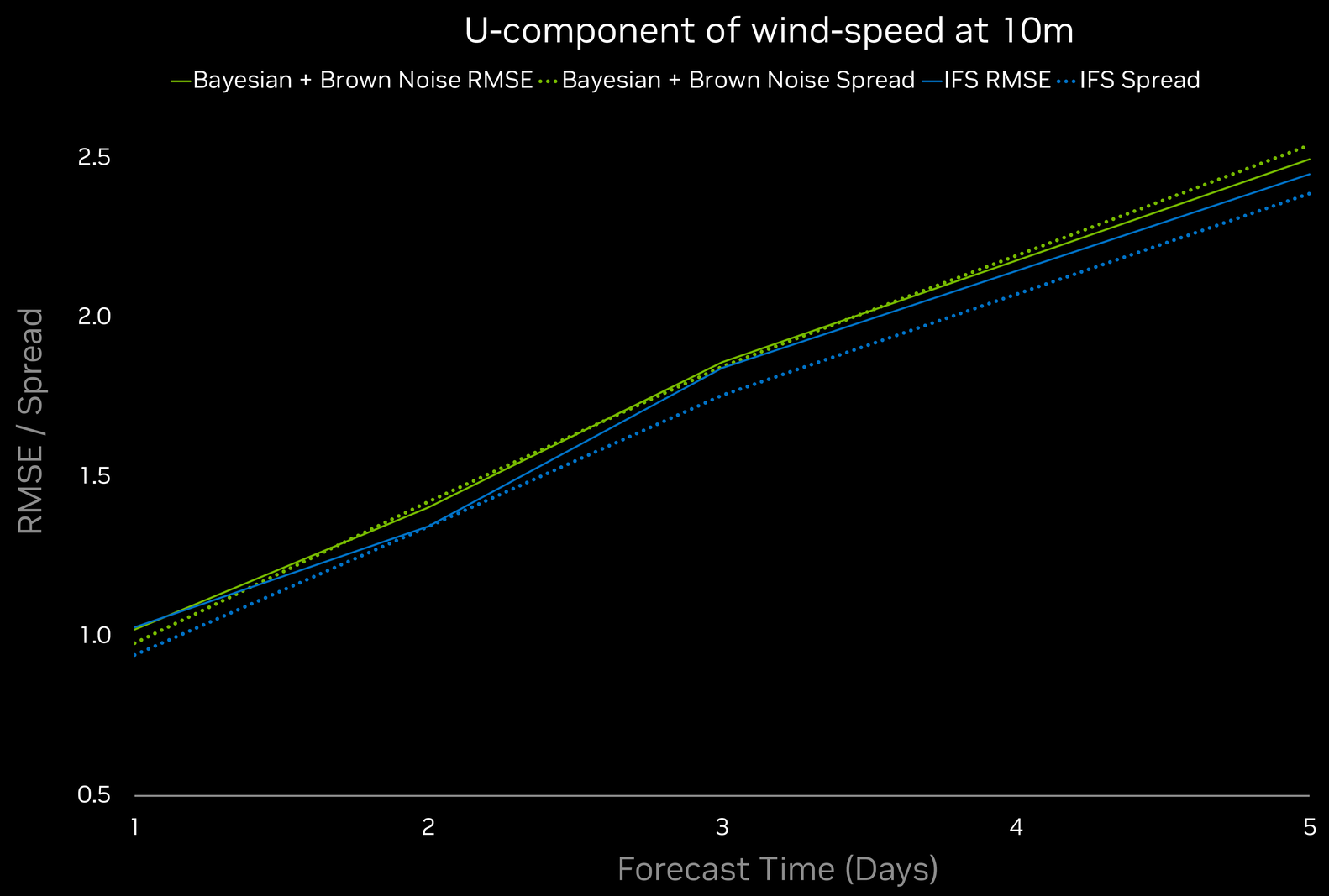
Acronym Alert:
ACC: Anomaly Correlation Coefficient (metric of weather skill)
IFS: The Integrated Forecast System, a gold standard weather model
FCN: FourCastNet, our digital twin of weather.

EXCELLENT SKILL ON FORECASTING SURFACE WINDS, PRECIPITATION, …



Probabilities: Spread matters as much as Skill

FCN's ensembles calibrated using initial condition uncertainty and model uncertainty (Bayesian SWA-G).

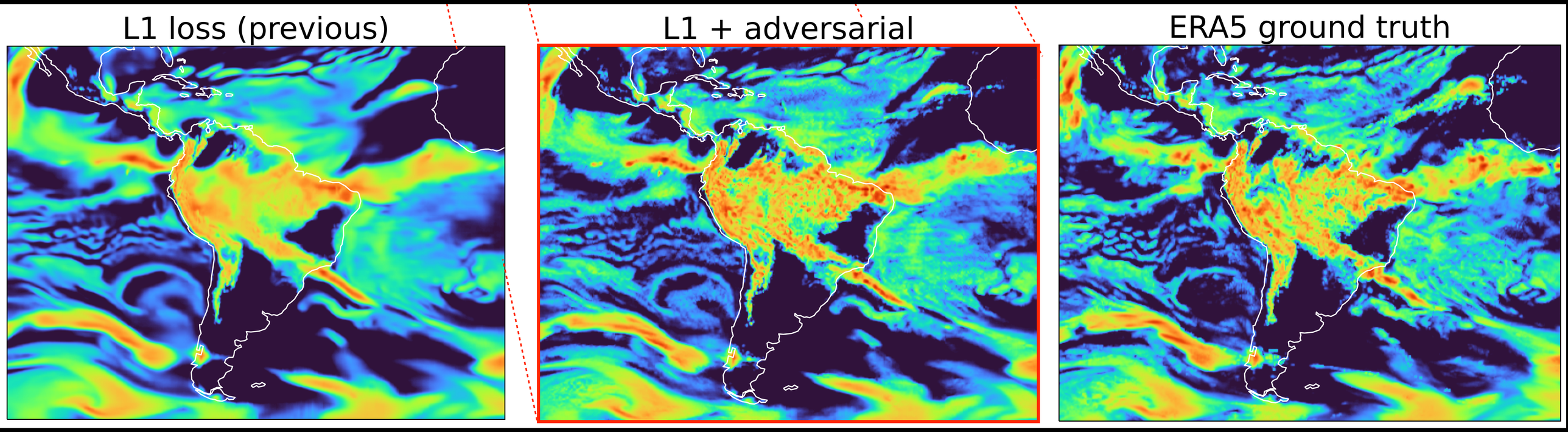


Continuous Ranked Probability Score competitive with IFS standard

Can generative approaches improve regional extremes?

Yes. Adversarial loss improves fine scale detail skillfully.

Forecast lead time of 18 hours



Credits: James Duncan, Shashank Subramaniam, and Peter Harrington, collaborators at NERSC, LBNL

BERKELEY LAB

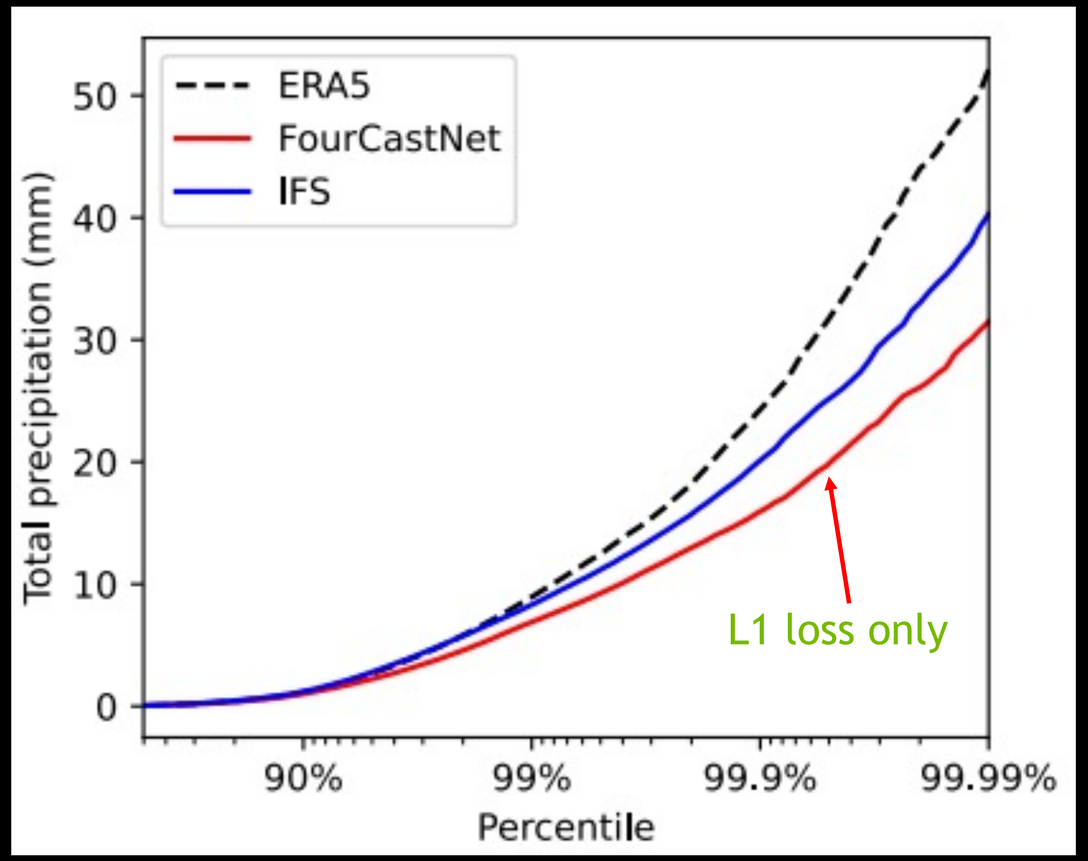
NERSC
National Energy Research Scientific Computing Center

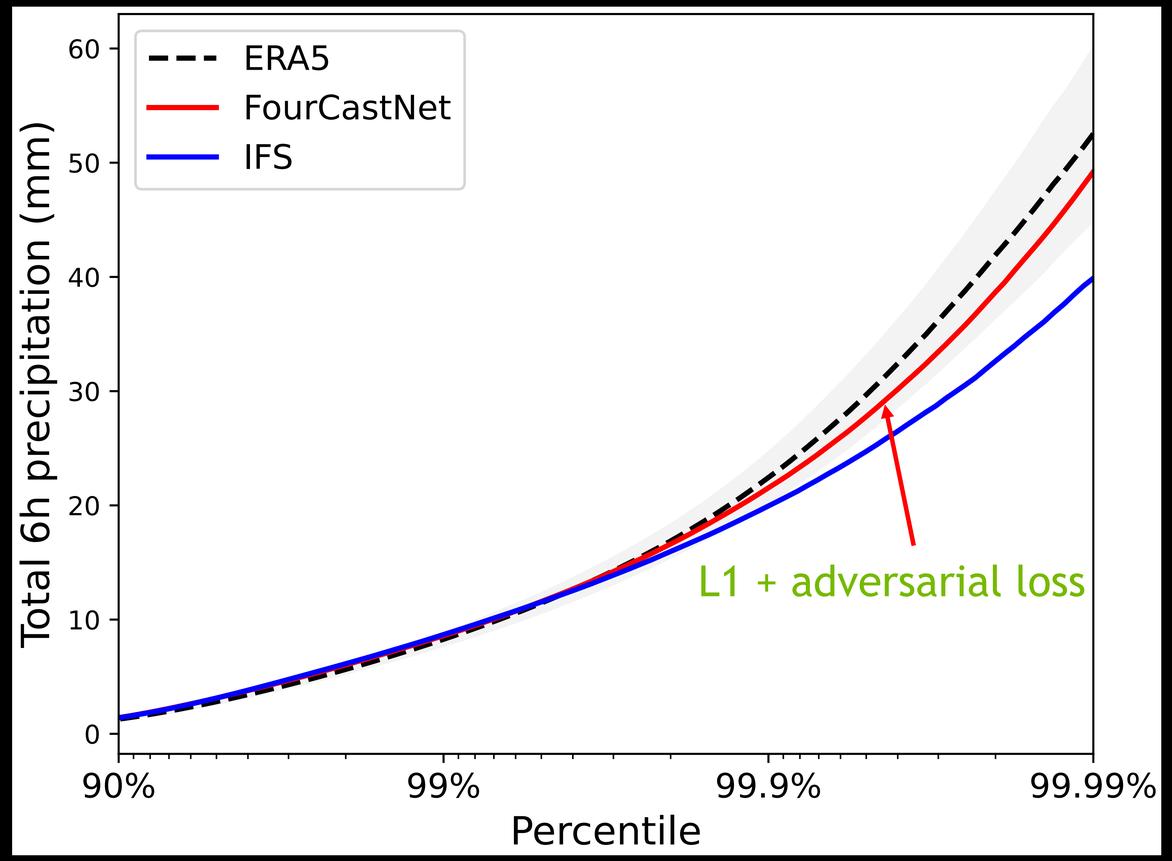
NVIDIA

Progress in capturing extreme precipitation statistics

Adding generative adversarial loss improves predictions of rarest, most intense rainfall events.

Forecast lead time of 18 hours

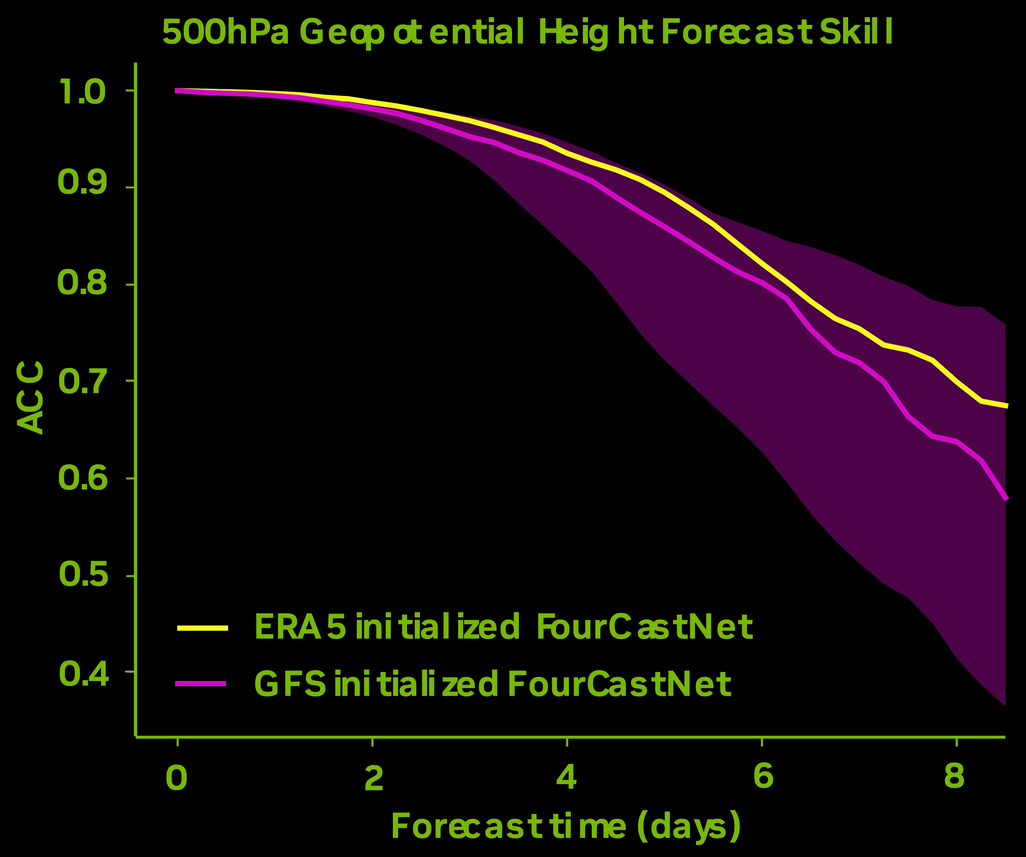


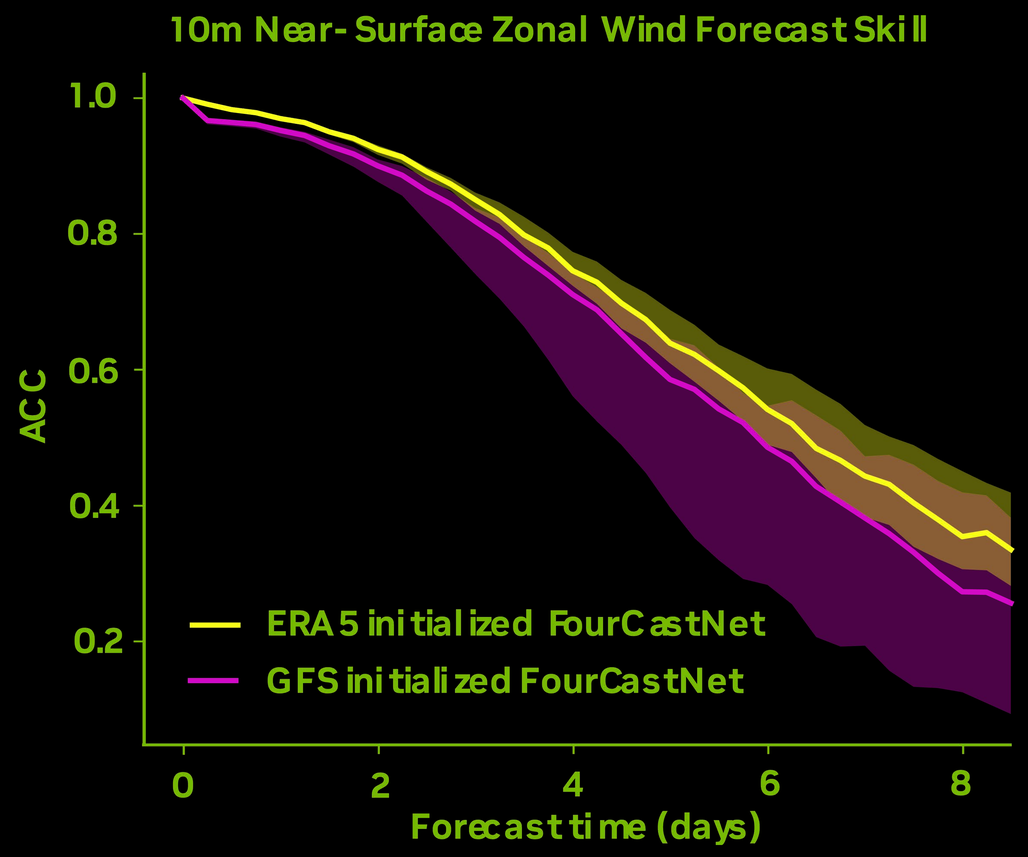


Credits: James Duncan, Shashank Subramaniam, and Peter Harrington, collaborators at NERSC, LBNL

Can FCN be initialized with real-time conditions?

Yes. Zero-shot skill transfer using initial conditions from a separate US dataset that FCN was not directly trained on.





Introducing open-source FCN.

Join us in pushing the frontiers of data-driven numerical weather prediction.

README.md

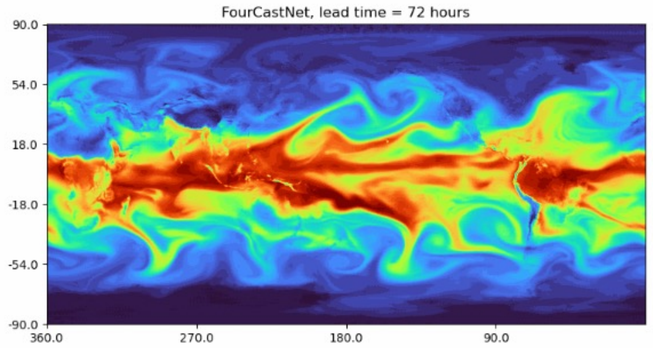
FourCastNet

This repository contains the code used for "FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators" [paper]

The code was developed by the authors of the preprint: Jaideep Pathak, Shashank Subramanian, Peter Harrington, Sanjeev Raja, Ashesh Chattopadhyay, Morteza Mardani, Thorsten Kurth, David Hall, Zongyi Li, Kamyar Azizzadenesheli, Pedram Hassanzadeh, Karthik Kashinath, Animashree Anandkumar

FourCastNet, short for Fourier Forecasting Neural Network, is a global data-driven weather forecasting model that provides accurate short to medium-range global predictions at 0.25∘ resolution. FourCastNet accurately forecasts high-resolution, fast-timescale variables such as the surface wind speed, precipitation, and atmospheric water vapor. It has important implications for planning wind energy resources, predicting extreme weather events such as tropical cyclones, extra-tropical cyclones, and atmospheric rivers. FourCastNet matches the forecasting accuracy of the ECMWF Integrated Forecasting System (IFS), a state-of-the-art Numerical Weather Prediction (NWP) model, at short lead times for large-scale variables, while outperforming IFS for variables with complex fine-scale structure, including precipitation. FourCastNet generates a week-long forecast in less than 2 seconds, orders of magnitude faster than IFS. The speed of FourCastNet enables the creation of rapid and inexpensive large-ensemble forecasts with thousands of ensemble-members for improving probabilistic forecasting. We discuss how data-driven deep learning models such as FourCastNet are a valuable addition to the meteorology toolkit to aid and augment NWP models.

FourCastNet is based on the vision transformer architecture with Adaptive Fourier Neural Operator (AFNO) attention proposed in Guibas-Mardani et al. [paper], [code].

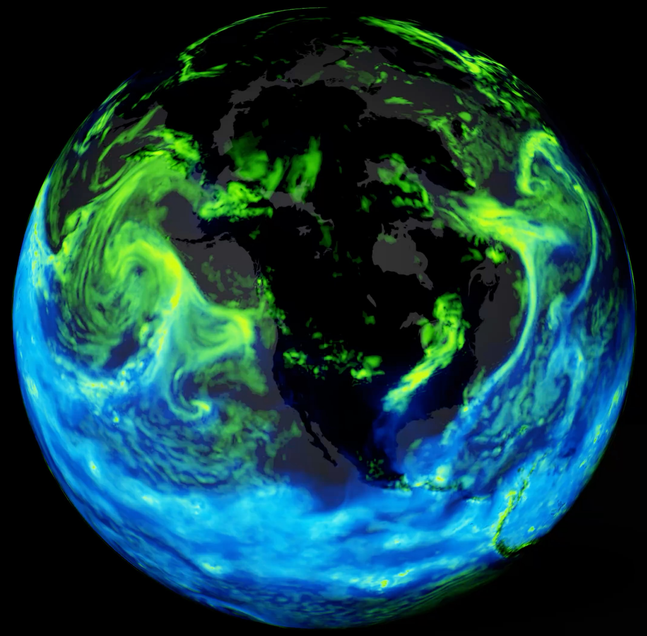


https://github.com/NVlabs/FourCastNet

Our Full AI Weather Model

FourCastNet has Progressed Rapidly this Year.

- Skill improving with training ambition
- Improving ensemble spread calibration
- Near-real time initialization, zero-shot skill transfer
- Generative approaches leading to improved extremes
- Available **Open Source**
- Join us in pushing the frontiers of data-driven weather prediction!



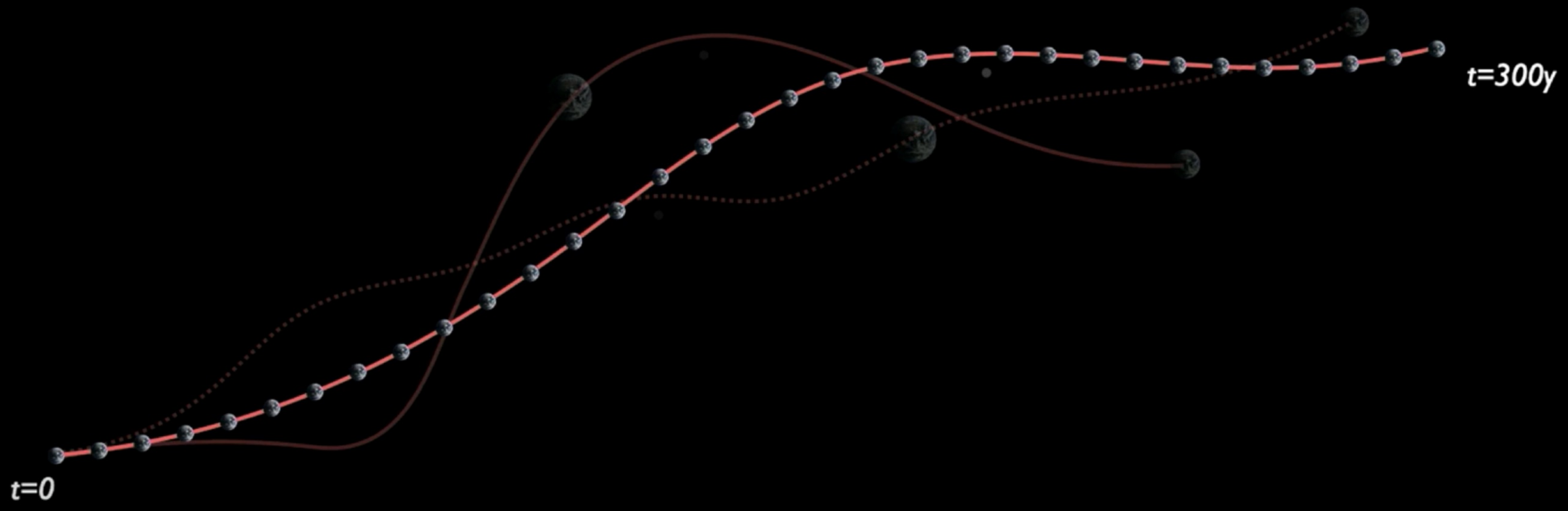
III. Beyond Today's Weather

Our Climate Simulation Research Strategy for Earth-2.

Given Future Data, FourCastNet's Speed Allows Fast Tethering

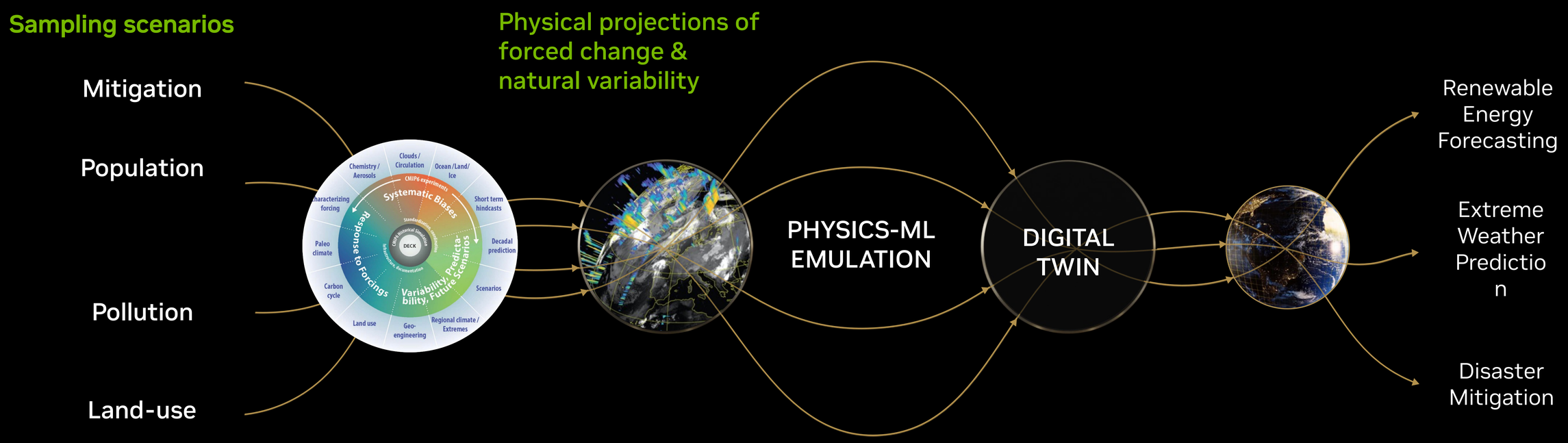
AI nimbly generates details between "checkpoints" saved only infrequently from physics-based climate simulations

-- Bjorn Stevens, GTC 2021



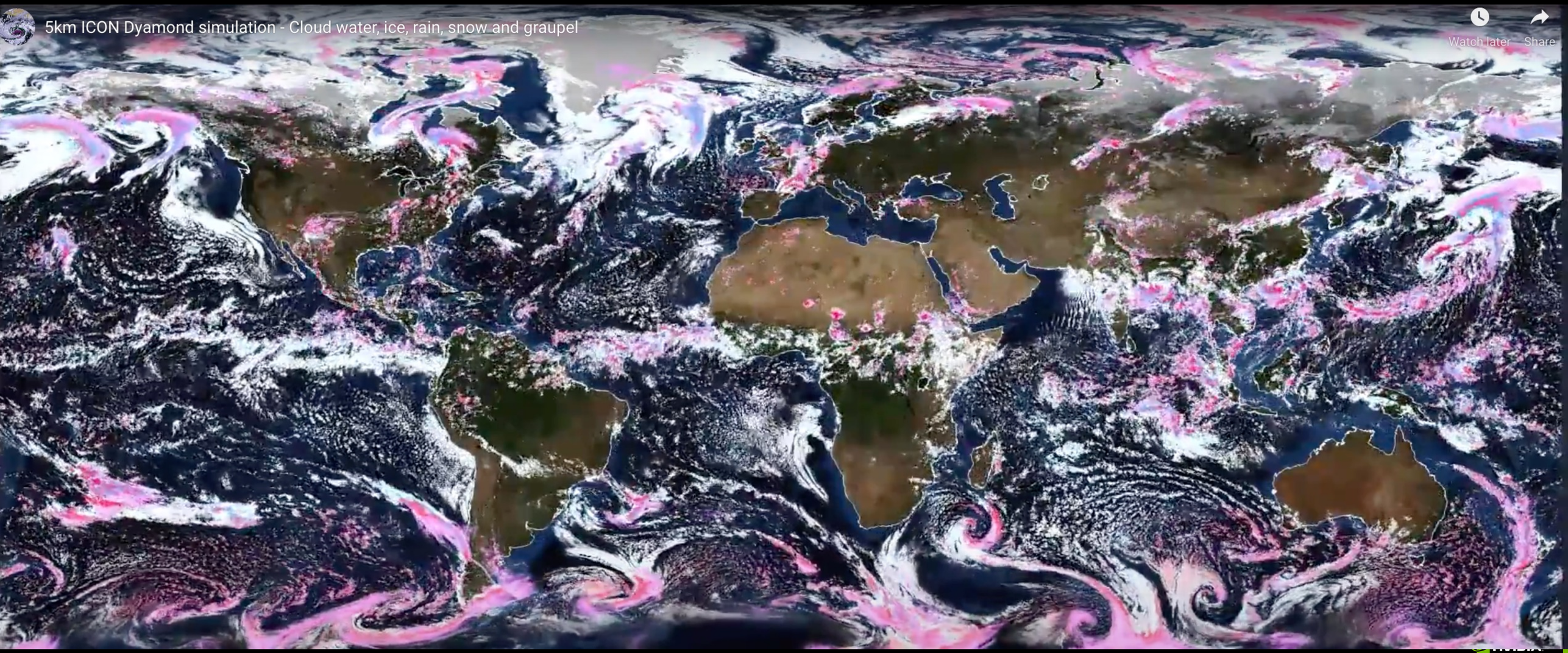
TO BEGIN, WE CAN TETHER TO EXISTING CLIMATE PREDICTIONS

Using the world's current data library of 100-km resolution intergovernmental climate predictions.



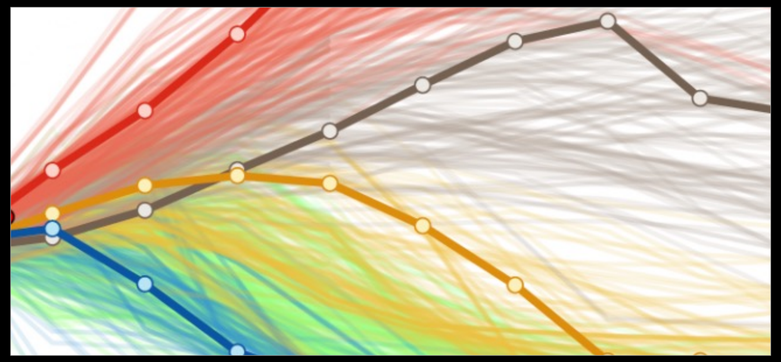
Eventually we want to Tether to KM-SCALE Predictions.

Because credible cloud feedbacks and storm dynamics from km-scale simulators matter to predicting regional risk.



EARTH-2: LOOKING AHEAD

We are pursuing various strategies to improve our climate digital twin

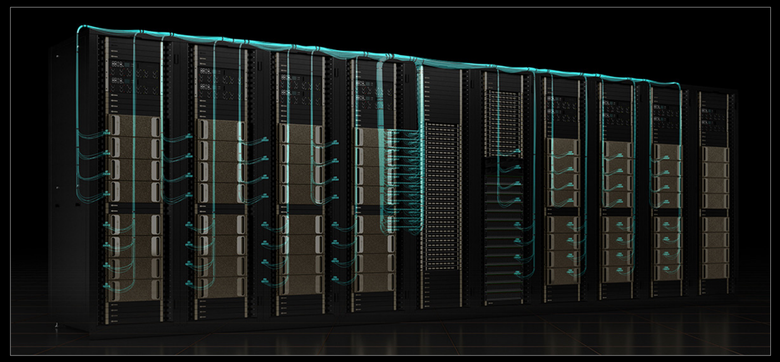


CMIP-6 Initialization or Surrogate

Source: Fuss et al., 2014

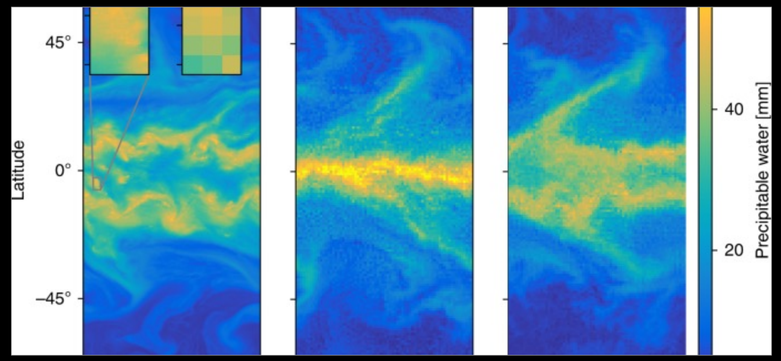
ICON 1km tethering

source: dyamond proj



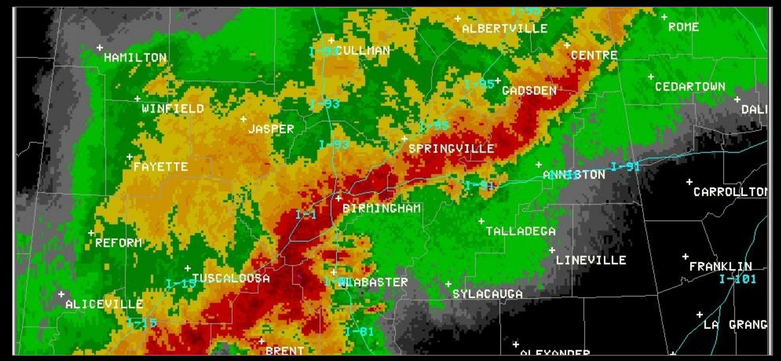
OVX hardware / software co-design

ovx superpods



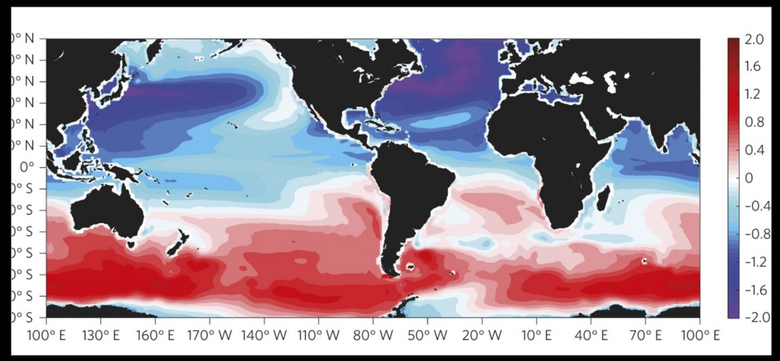
1km → 5 km AI sub-grid emulation

source: nature.com



Regional Fine-Tuning and Downscaling

source

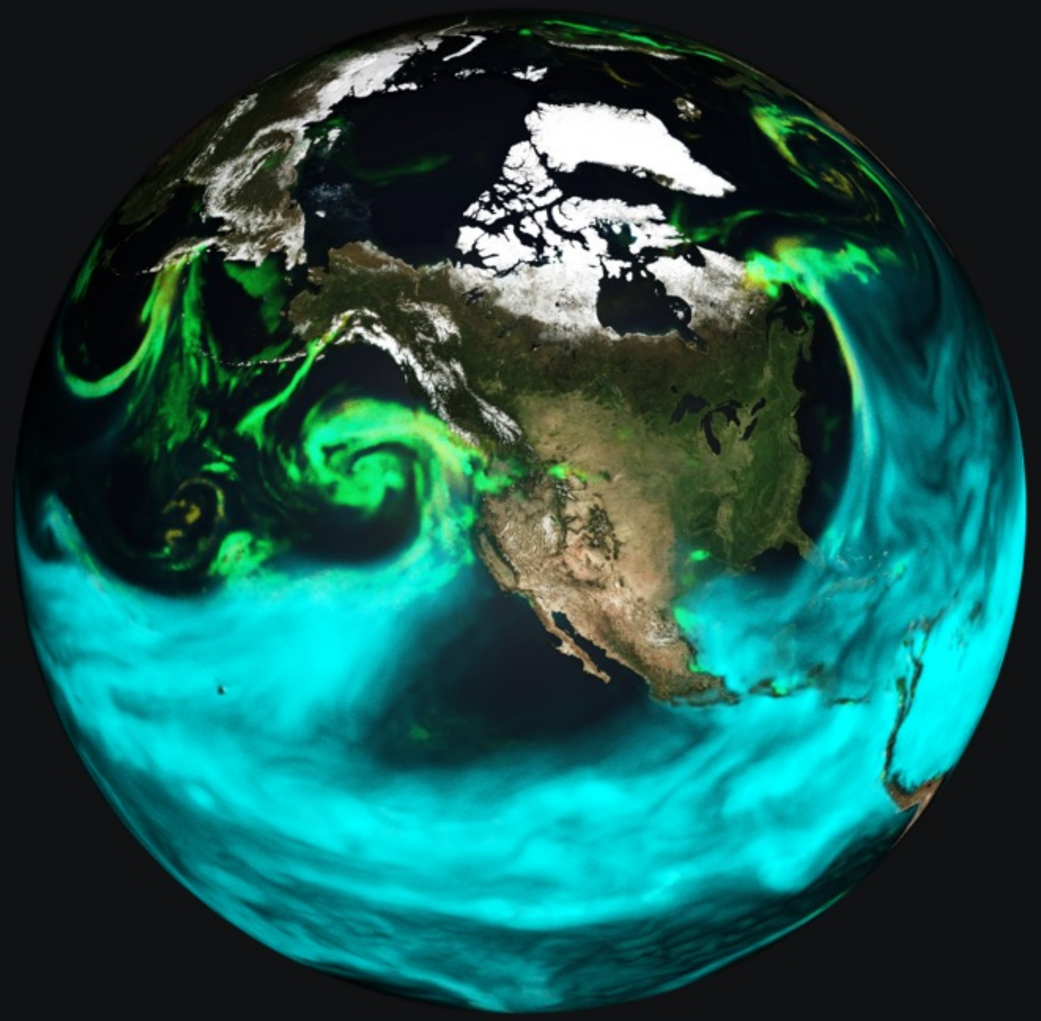


Model Auto-Calibration via RL

source: nature

The Vision of Earth-2

Is Beginning to Take Shape



Acknowledging: Mike Pritchard, Anima Anandkumar, David Hall, Jaideep Pathak, Thorsten Kurth, Andre Graubner, Peter Messmer, Stan Posey, Akshay Subranamian, Sanjay Choudhry, Farah Hariri, Niklas Roeber, Ram Cherukuri, Nicholas Geneva, Mathias Hummel, Christopher Lamb, Mike Houston, Kamyar Azizzadenesheli, Jean Kossaifi, Steffen Roemer, Marius Koch & David Appelhans, many more NV staff & our **generous external climate science advisors Bjorn Stevens, Peter Deuben, Peter Bauer, Nils Wedi, and Francisco Doblas-Reyes**.

NVIDIA


NVIDIA