University of Southampton

IBM Research

Predicting Extreme Weather Events

Augmenting AI Models to Improve Reliability

PyData Southampton

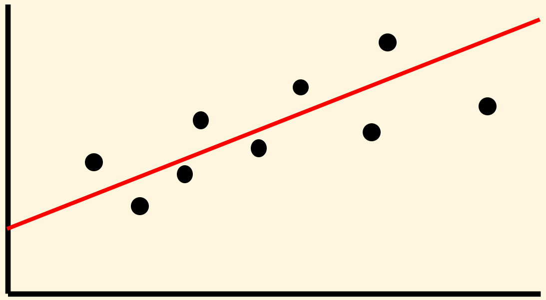
21st October 2025

Austen Wallis

What is a Machine Learning Model?

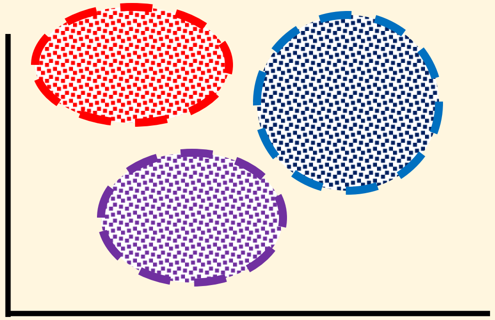
ML Model: A program created by training an algorithm on a dataset to find patterns and make predictions or decisions about new, unseen data.

Supervised Learning



Learns from labelled data

Unsupervised Learning

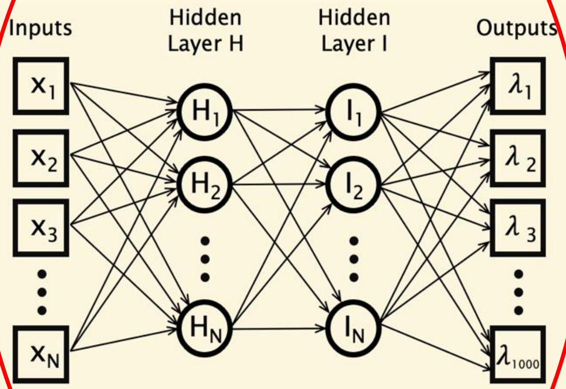


Learns from unlabeled data

Reinforcement Learning

Learns through trial and error/rewards

Deep Learning



Learns through artificial neural networks

Training Models

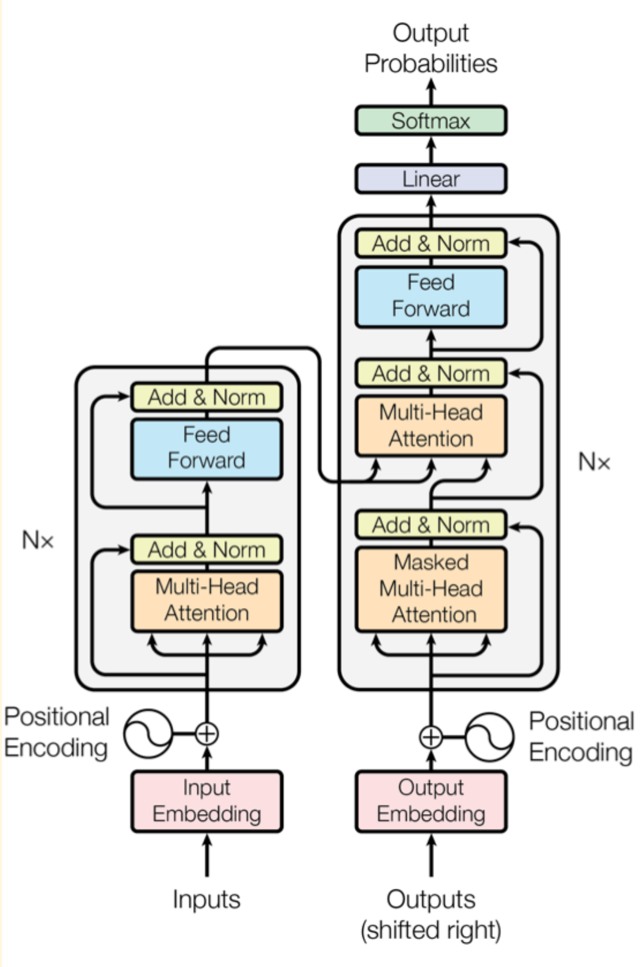
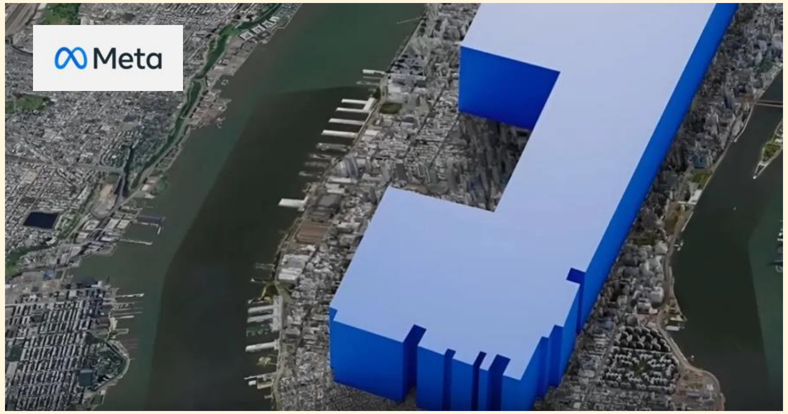


Figure 1: The Transformer - model architecture.

There are many deep learning architectures:

- Multi-Layer perceptrons
- Convolutional Neural Networks
- Recurrent Neural Networks
- Long Short-Term Memory
- Generative Adversarial Networks
- Autoencoders

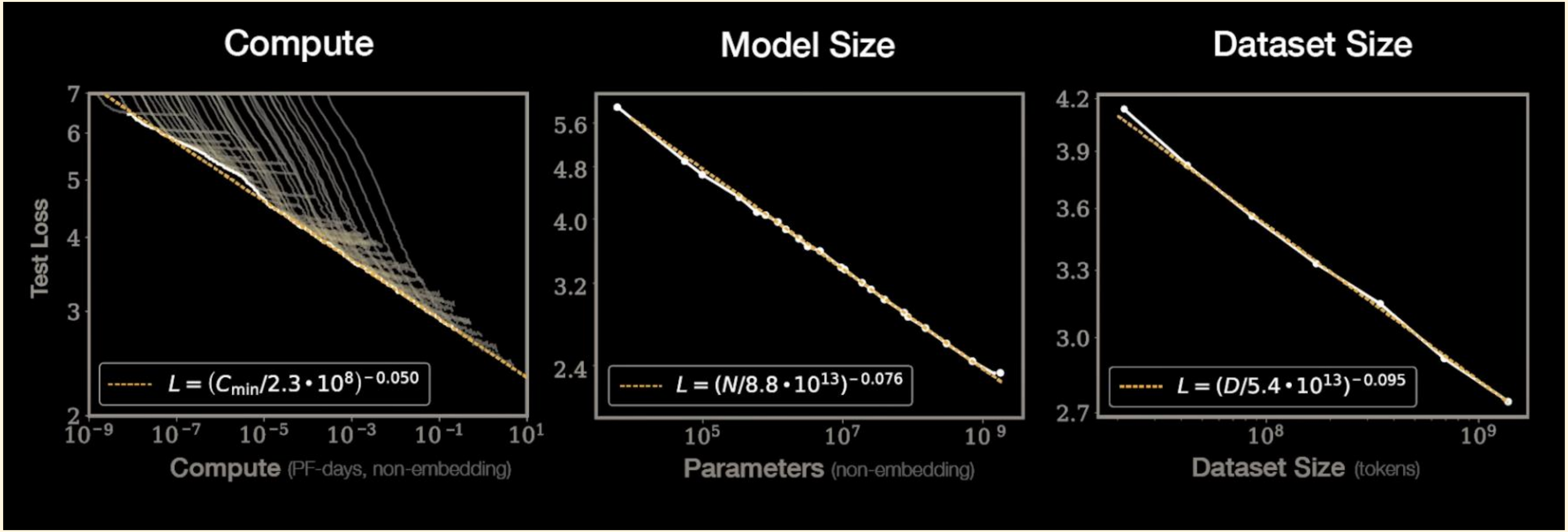
⋮



Today's most talked-about architecture is Transformers

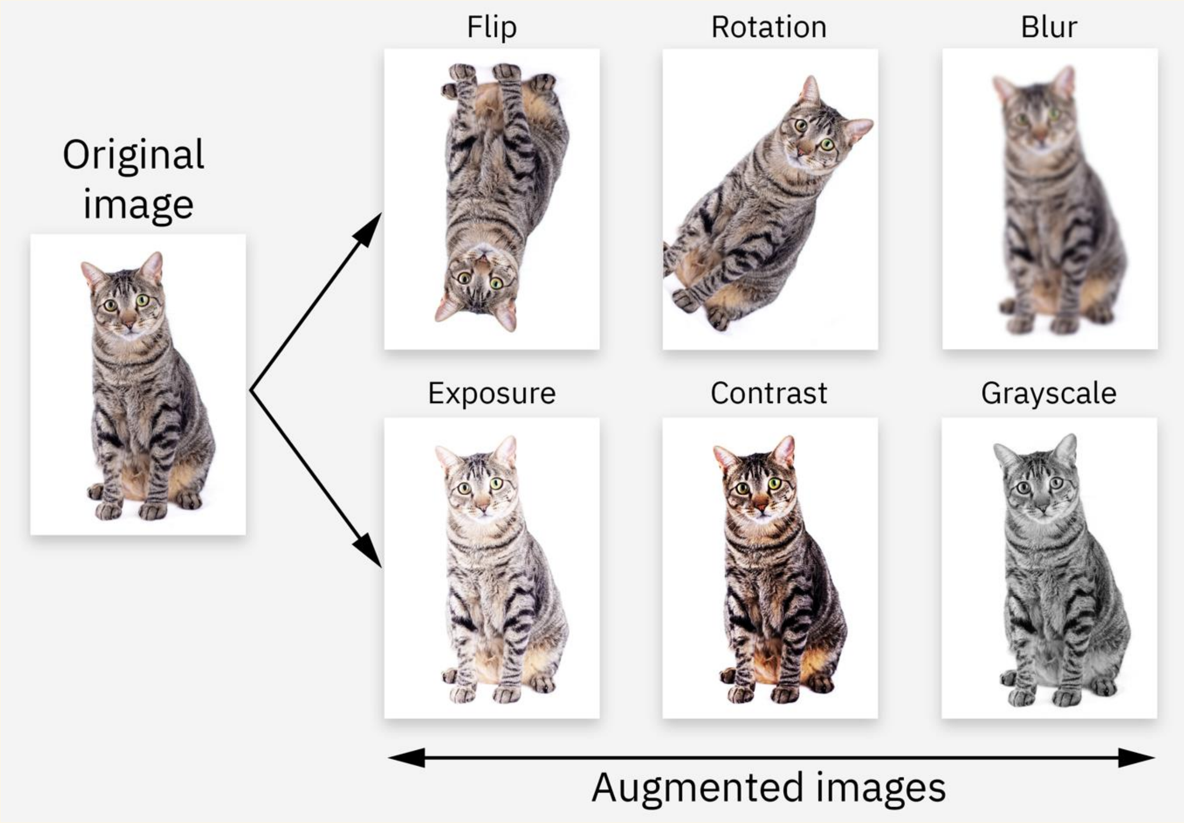
… and the larger, the better.

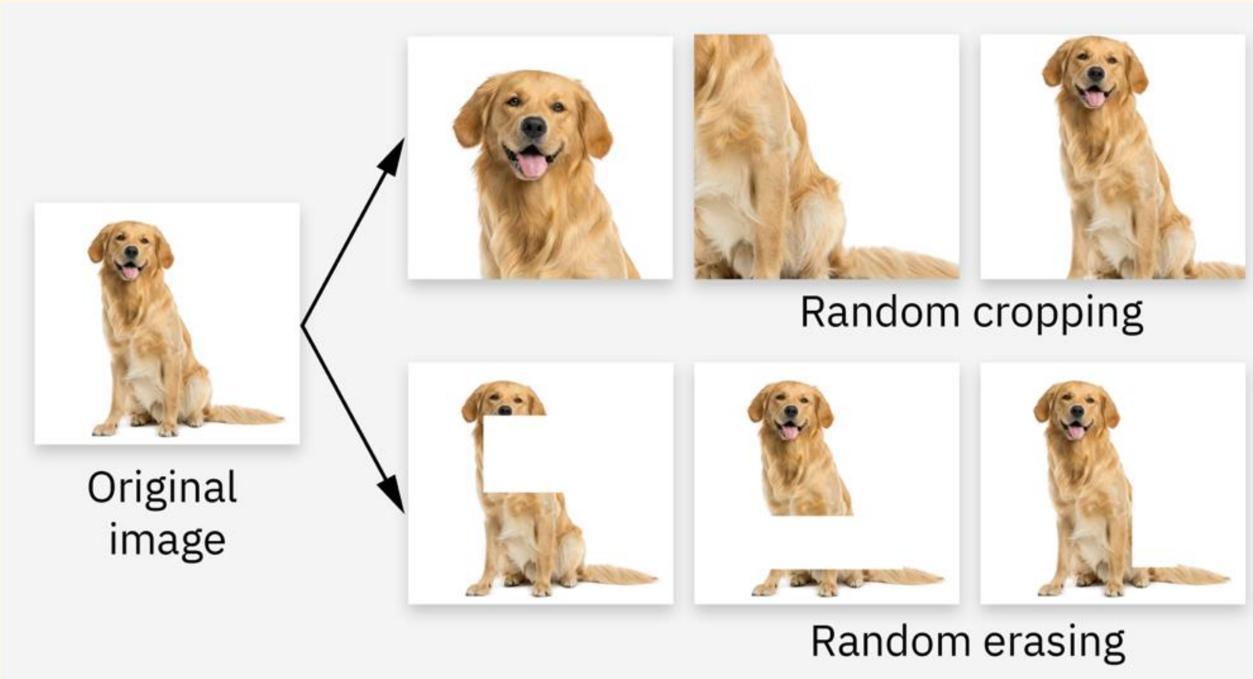
Transformer Scaling Laws



Adapting Data for Better Models

Data augmentation: uses pre-existing data to create new data samples that can improve model optimisation and generalizability.

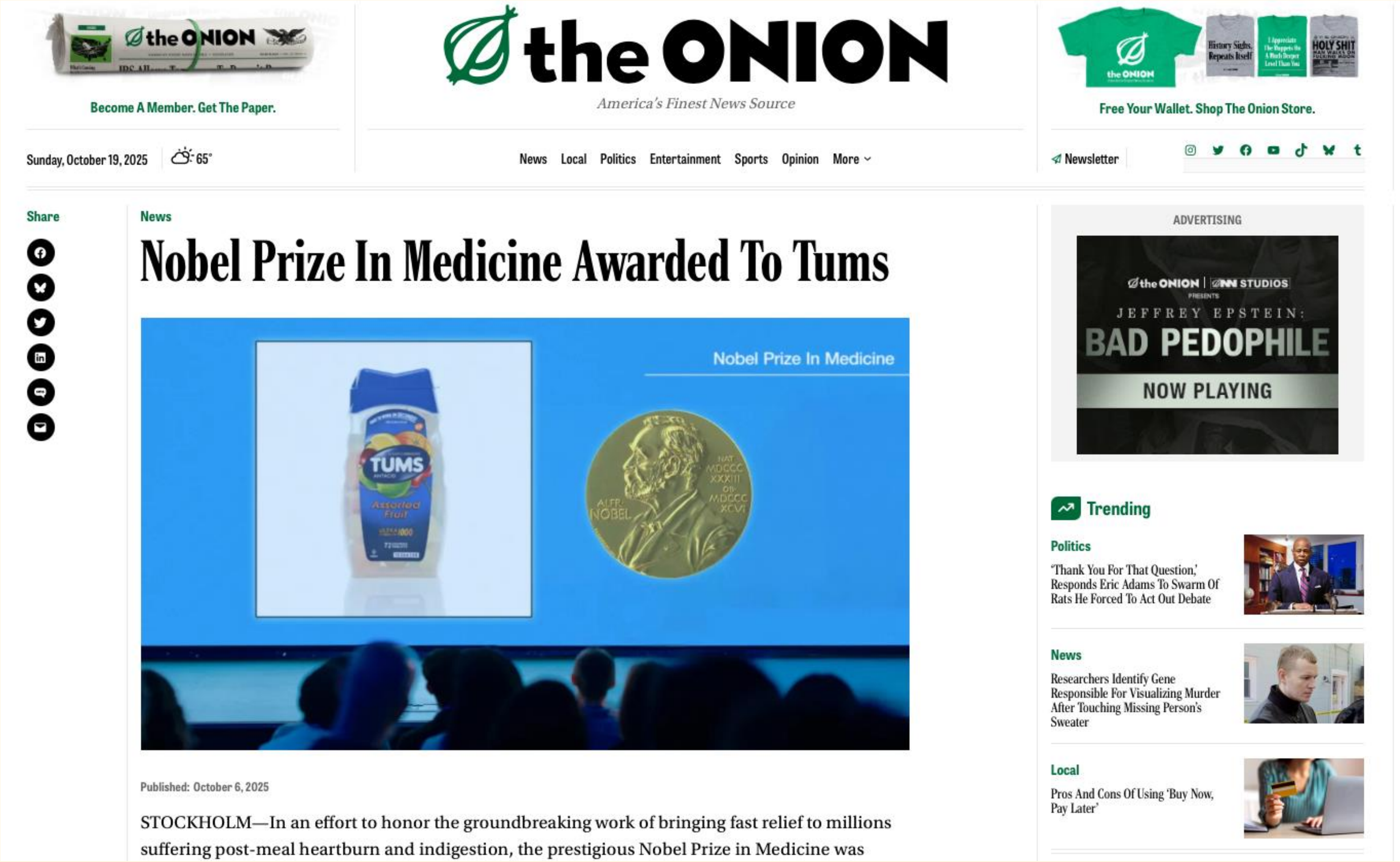




Noise Risks

Scaling datasets poses risks and noise into the models! For example:

Impossible Doggy



Become A Member. Get The Paper.

the ONION

America's Finest News Source

Free Your Wallet. Shop The Onion Store.

Sunday, October 19, 2025 65°

News Local Politics Entertainment Sports Opinion More

Newsletter

Share

News

Nobel Prize In Medicine Awarded To Tums

Nobel Prize In Medicine

Published: October 6, 2025

STOCKHOLM—In an effort to honor the groundbreaking work of bringing fast relief to millions suffering post-meal heartburn and indigestion, the prestigious Nobel Prize in Medicine was

ADVERTISING

the ONION | STUDIOS

JEFFREY EPSTEIN: BAD PEDOPHILE

NOW PLAYING

Trending

Politics

'Thank You For That Question,' Responds Eric Adams To Swarm Of Rats He Forced To Act Out Debate

News

Researchers Identify Gene Responsible For Visualizing Murder After Touching Missing Person's Sweater

Local

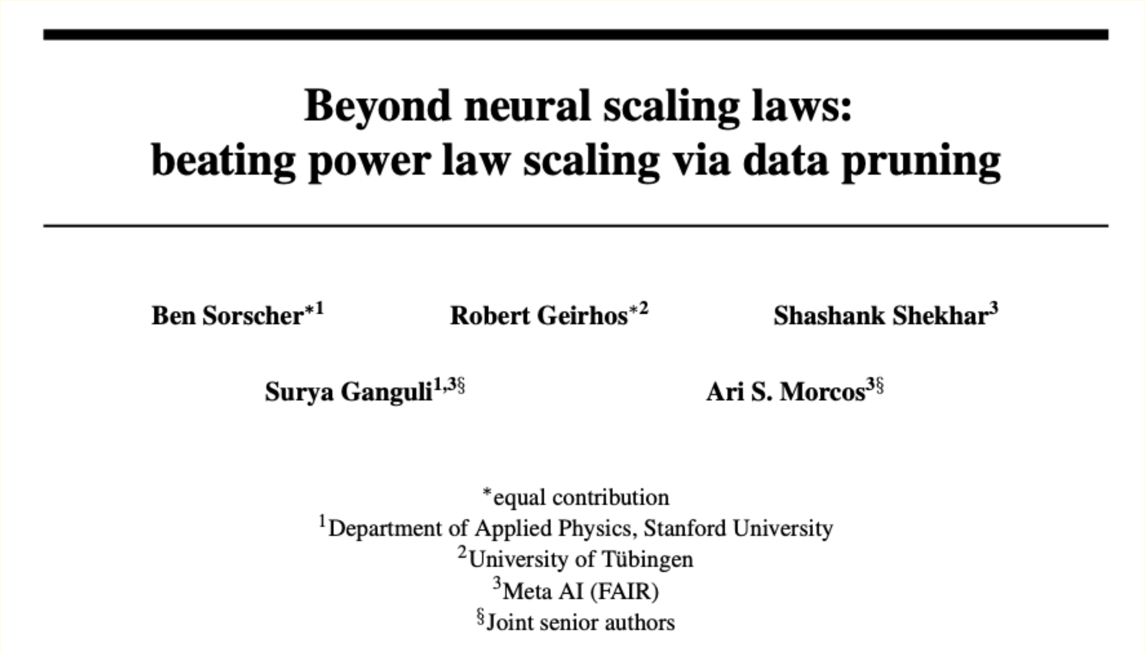
Pros And Cons Of Using 'Buy Now, Pay Later'

Misinformation

Pruning Data

Equally, not all data is equal!

Selective pruning can beat neural scaling laws.

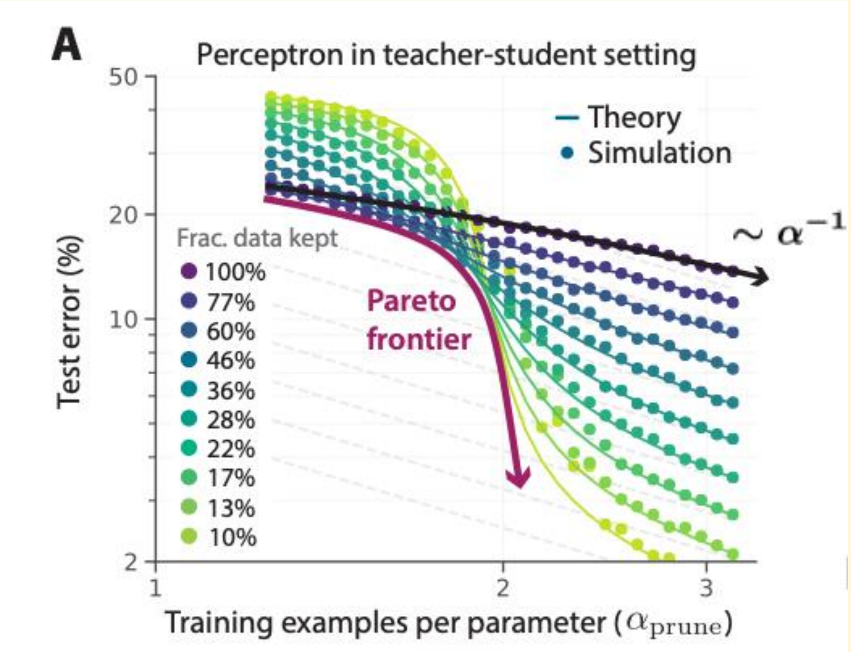


Beyond neural scaling laws: beating power law scaling via data pruning

Ben Sorscher*[1] Robert Geirhos*[2] Shashank Shekhar[3]

Surya Ganguli[1,3§] Ari S. Morcos[3§]

*equal contribution
[1]Department of Applied Physics, Stanford University
[2]University of Tübingen
[3]Meta AI (FAIR)
[§]Joint senior authors

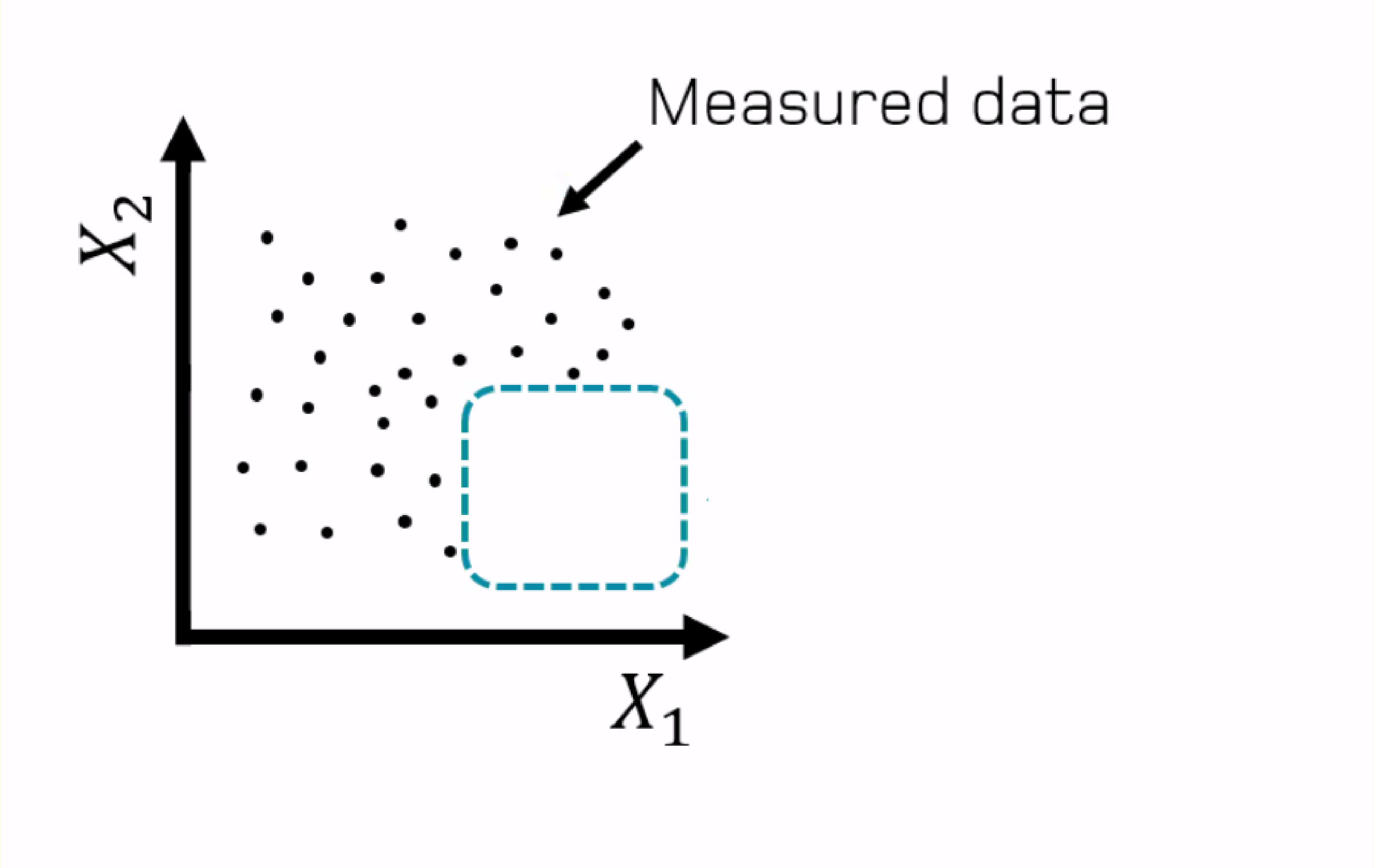


Improving Results for Science

This works great for existing datasets.

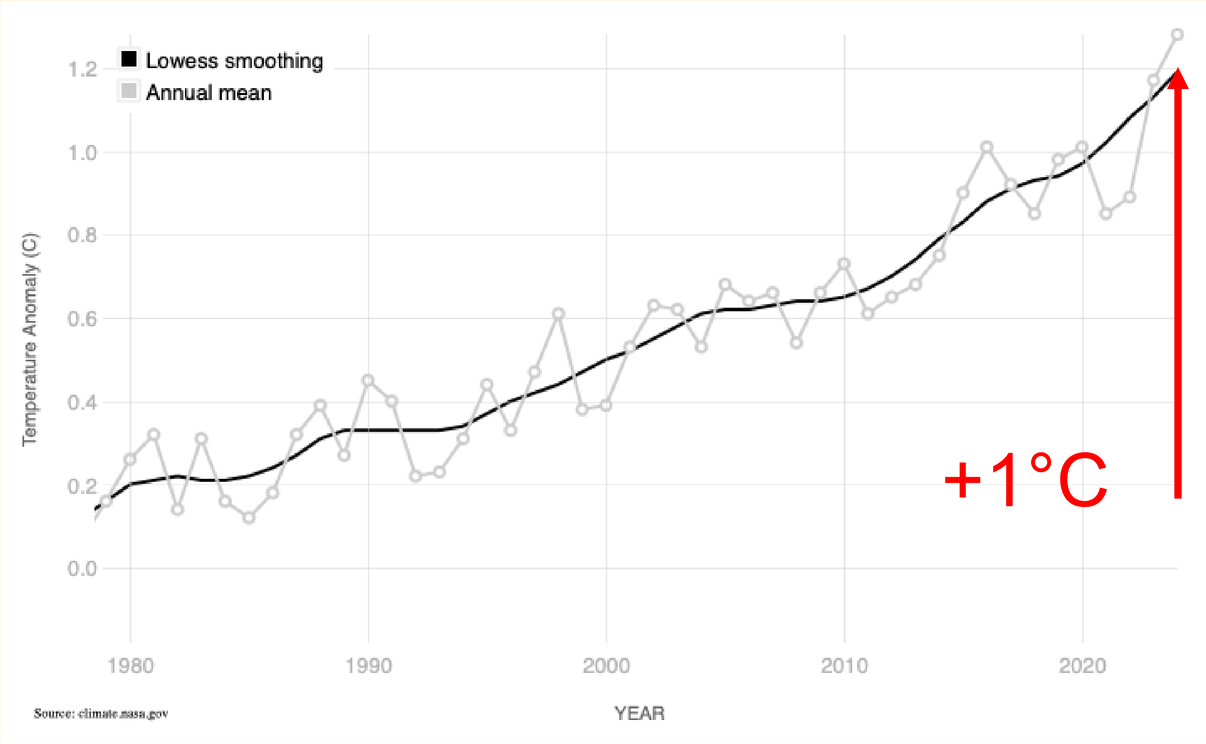
However, what happens if your data **doesn't** exist?

For example, how do extreme weather events react with warmer global temperatures?

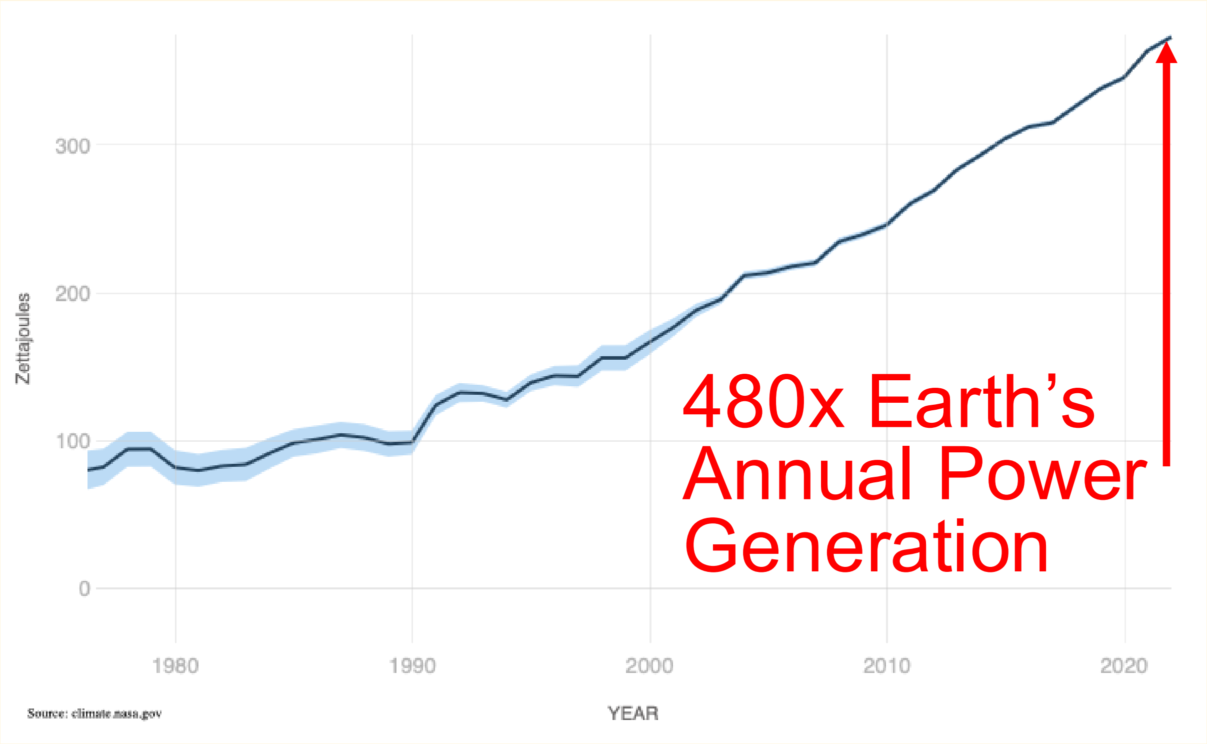


Energy absorbed by Atmosphere and Oceans

Atmospheric Temperature



Ocean Temperature



From 1980 to Today

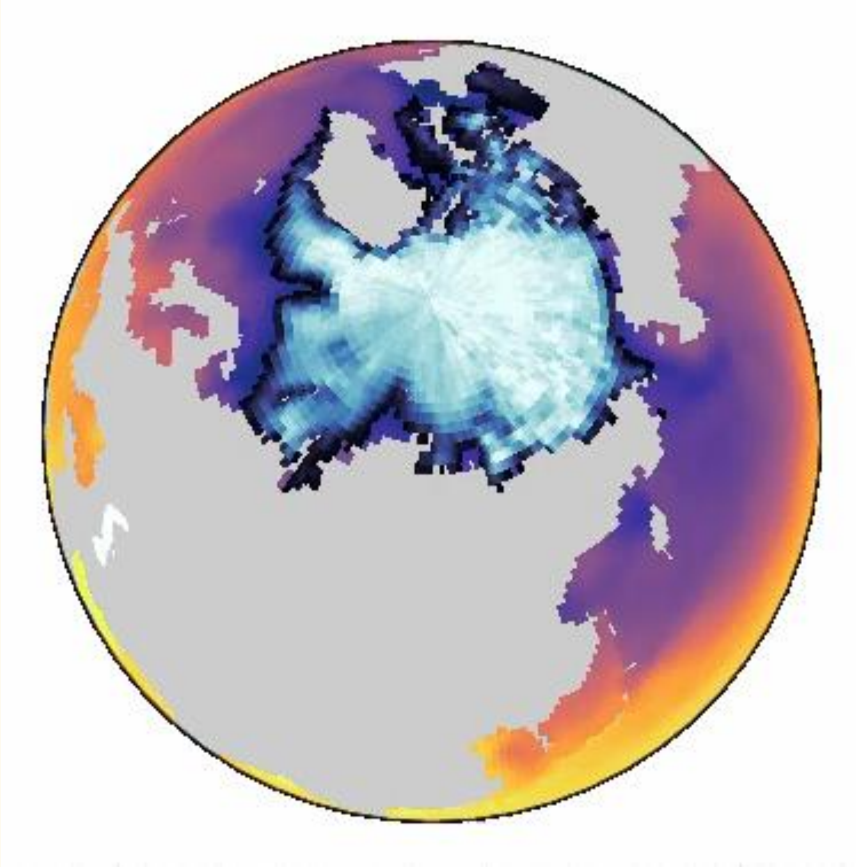
Albedo, Ice Melt and CO_2

- Ice melt is decreasing Earth's albedo

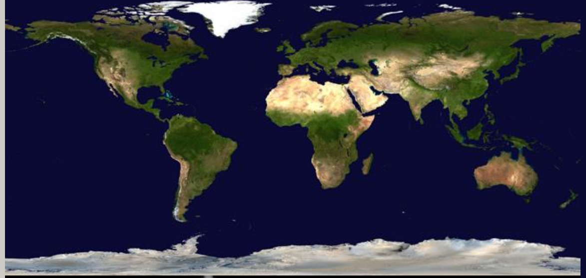
$$Albedo = \frac{Reflected\ Light}{Absorbed\ Light}$$

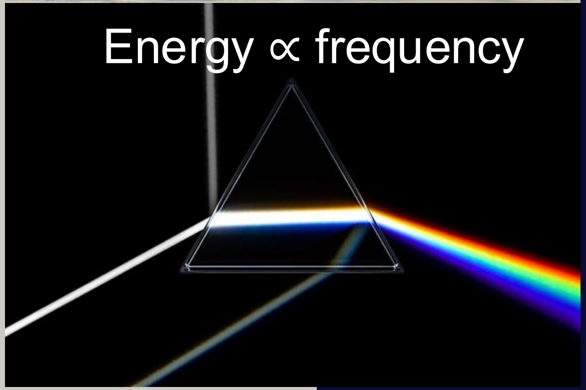
- Albedo change equivalent to a quarter of CO_2 emission in last 30yrs (Pistone et al. 2014)
- Expected ice-free summers in the next 70 yrs

Sea temperatures and ice coverage - 40-year rollout



Credit: SamudrACE





Ice Colour

Ocean Colour

Credit: NASA/Goddard Space Flight Center Scientific Visualization Studio
The Blue Marble Next Generation data is courtesy of Reto Stockli (NASA/GSFC) and NASA's Earth Observatory.

Out-of-distribution events

New Records Yearly!

How do we model and predict these events?

NEWS / HURRICANE

Erick, the strongest hurricane to hit Mexico so early in the year, triggers deadly flooding

By Mary Gilbert, CNN

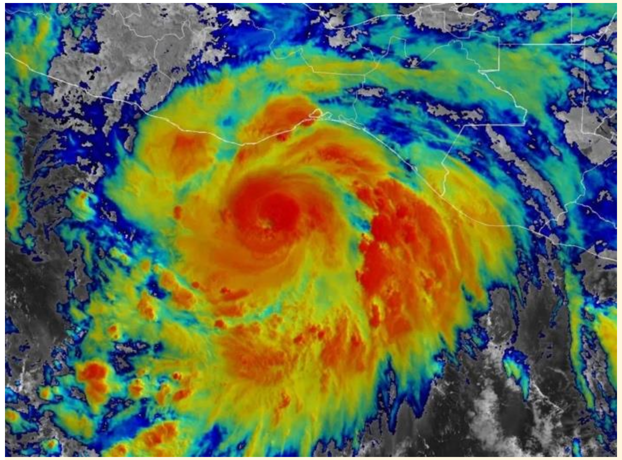
20 Jun, 2025 2:57 PM BST | Updated 20 Jun, 2025 2:57 PM BST

TROPICAL CYCLONES

Record-setting Atlantic storm Andrea dissipates within 12 hours of forming

By Rishav Kothari • Wednesday, June 25, 2025 • Follow us on Google News

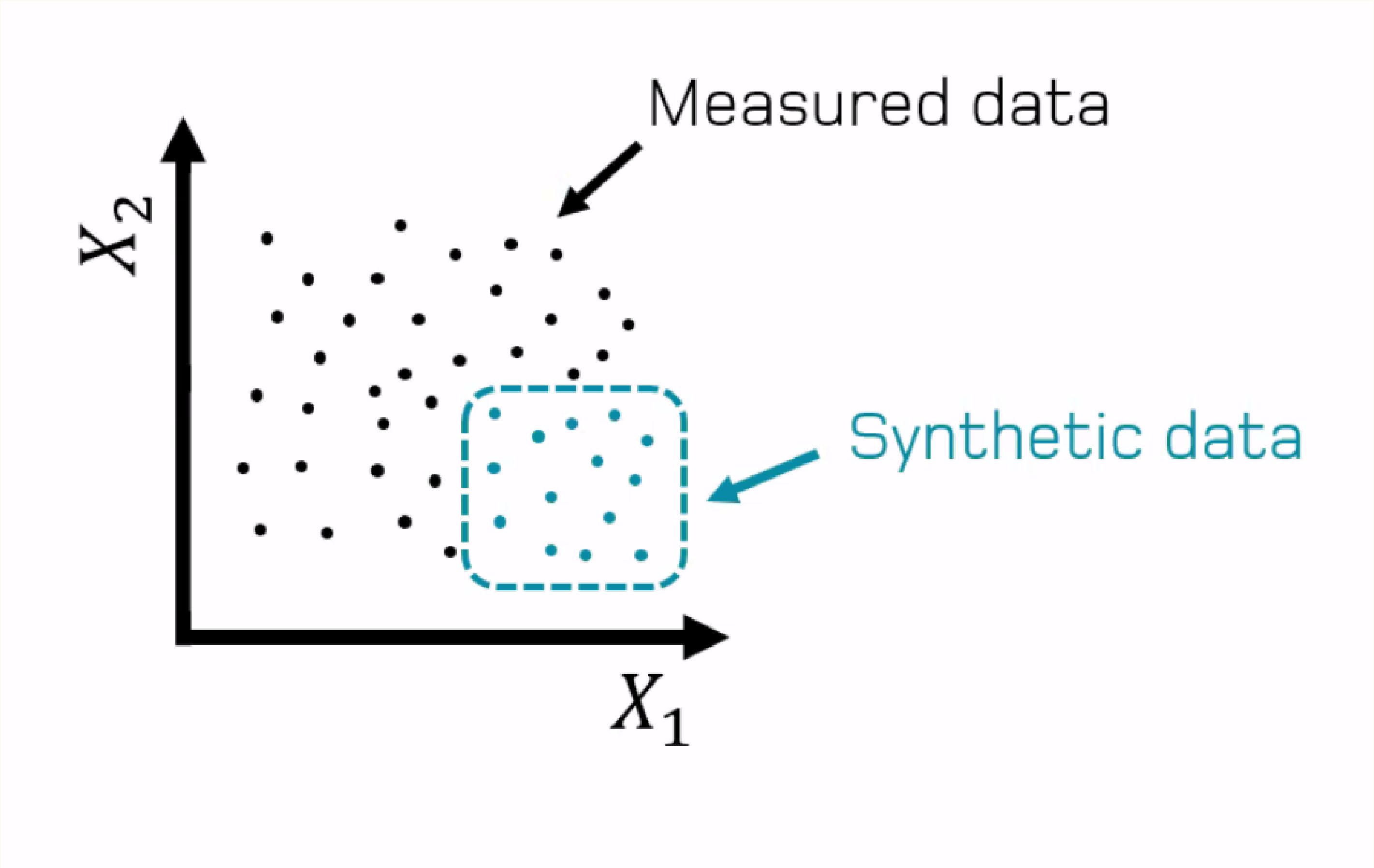
Tropical Storm Andrea weakened into a post-tropical cyclone less than 12 hours after formation as the first named storm of the 2025 Atlantic hurricane season on June 24, 2025, and is forecast to soon dissipate on June 25. Andrea has set a new record as the farthest north and east a tropical system has formed in the Atlantic during June since 1851.



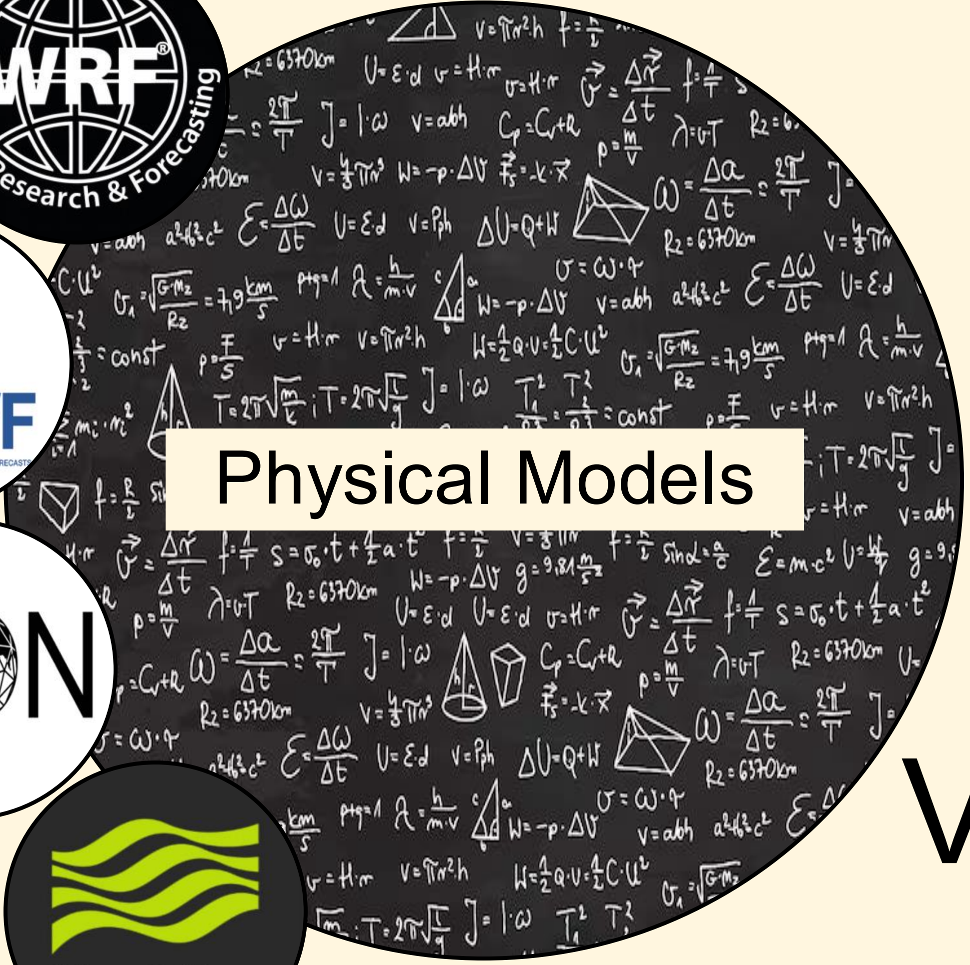
Blending Physical and ML Modelling Methods

- We want to create synthetic data that embodies the underlying physics to add **observation bias** into historic datasets.

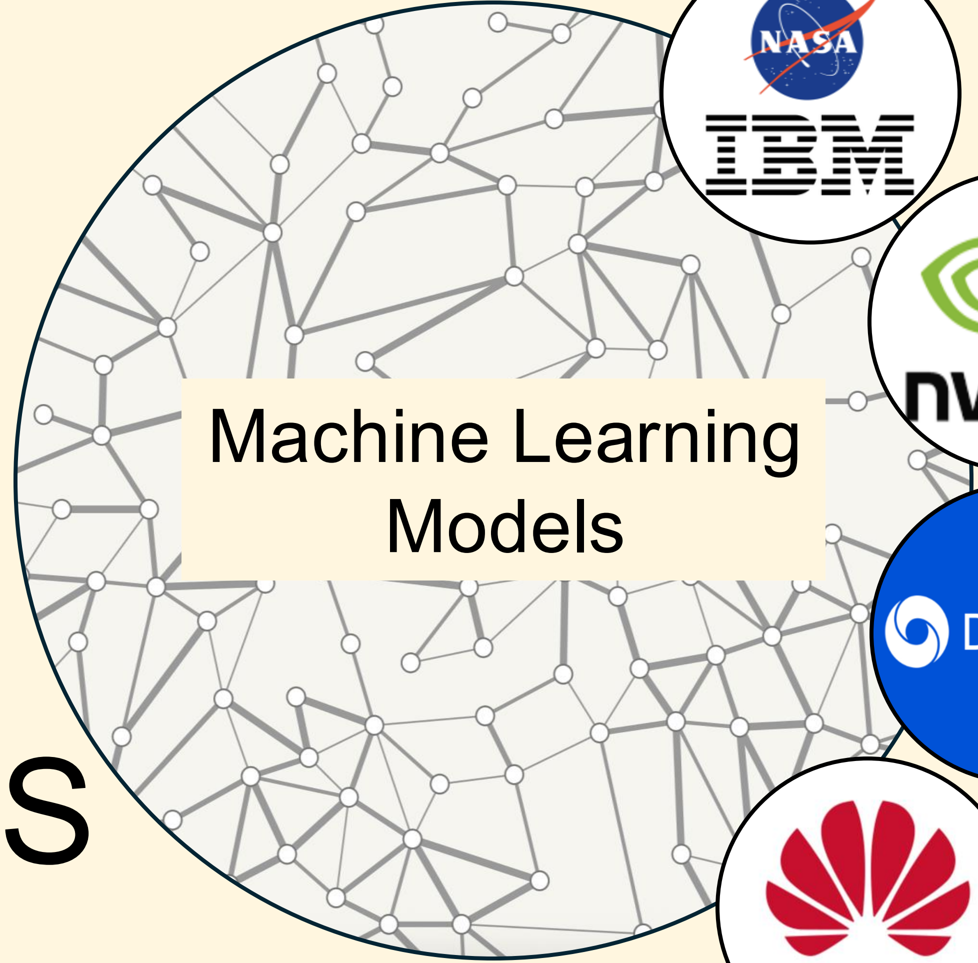
- This leads to robuster ML models.



Weather Modelling Toolkits



VS

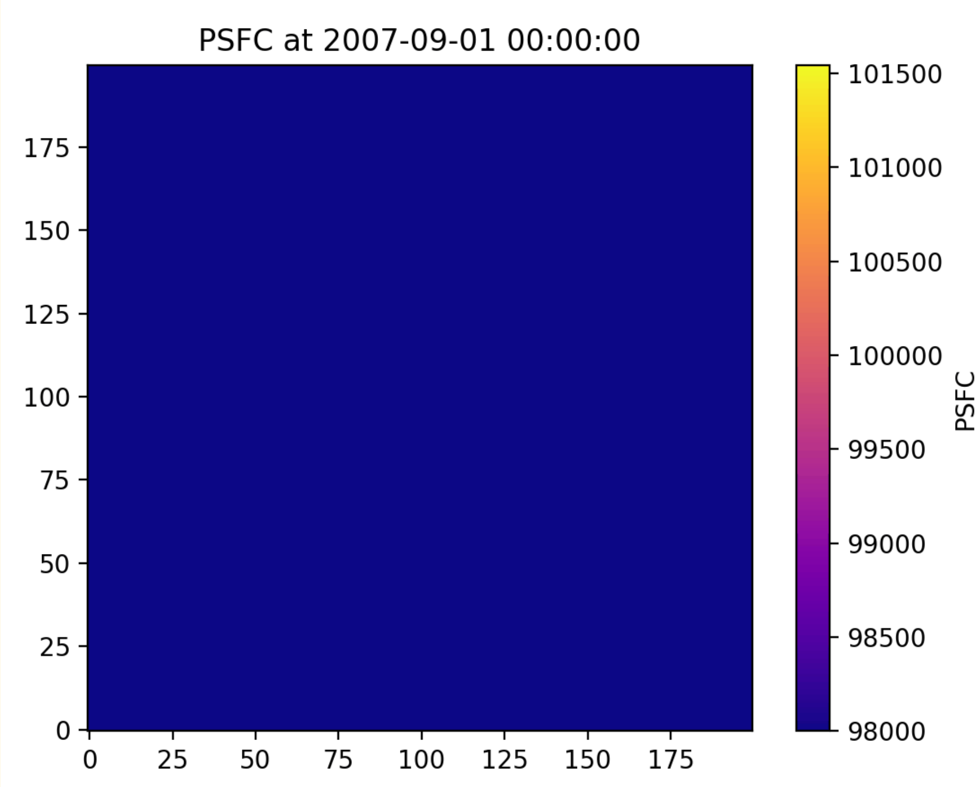


…

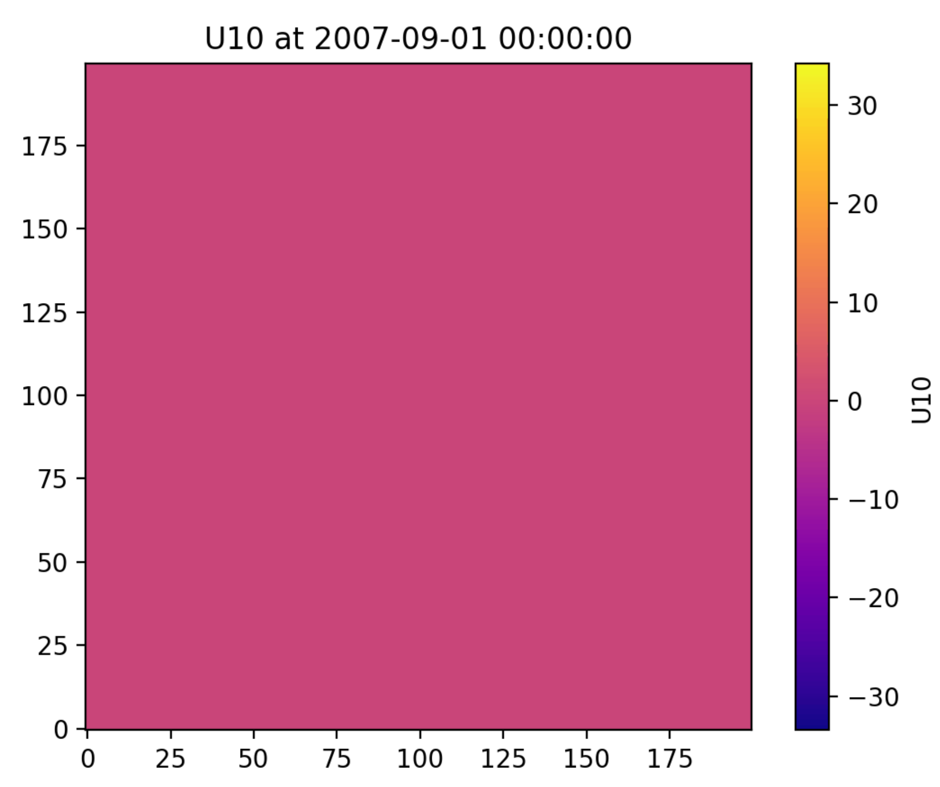
…

Physical Modelling - WRF

The weather research and forecasting (WRF) tool is a deterministic, physics-based forecasting tool to predict bounded weather phenomena.



Pressure



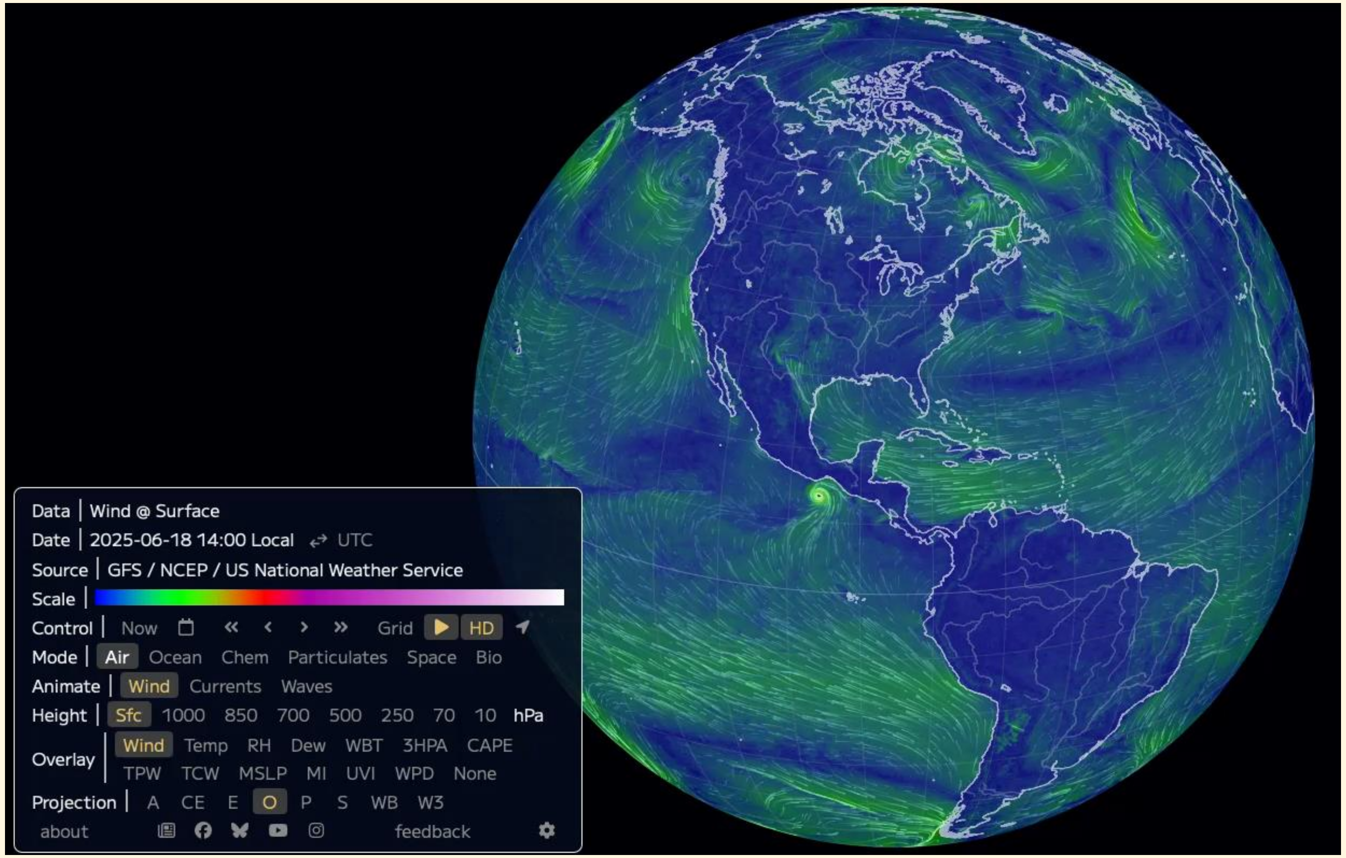
Easterly Wind Speed

Machine Learning Models

Program	Company	Architecture
GraphCast	Google DeepMind	Graph Neural Network
GenCast	Google DeepMind	Diffusion Model
FourCastNet v3	NVIDIA	Adaptive Fourier Neural Operators
Pangu-Weather	HUAWEI	3D Earth-specific transformer
Aurora AI	Microsoft	Autoencoder/Transformer
Prithvi	IBM NASA	Temporal Vision Transformer

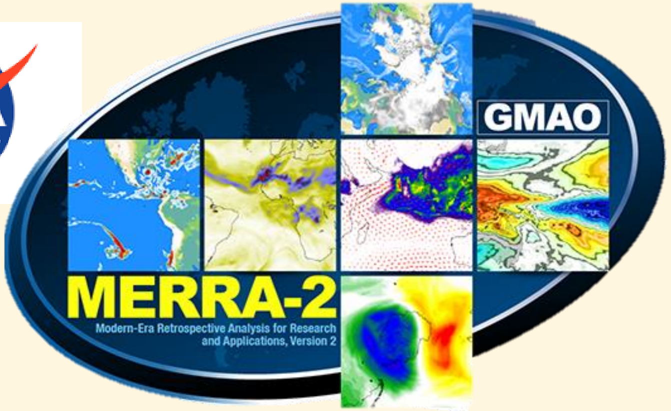
Data Sources

Global Forecast System (GFS) Data i.e. Live Data



earth.nullschool.net

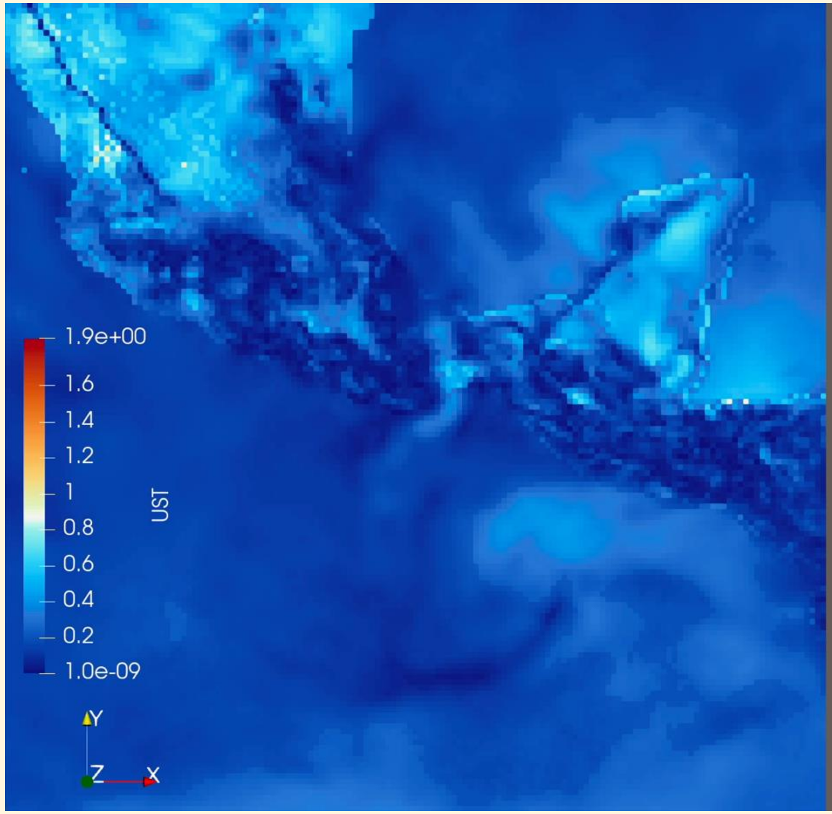
Historic Data - From 1940 to Today



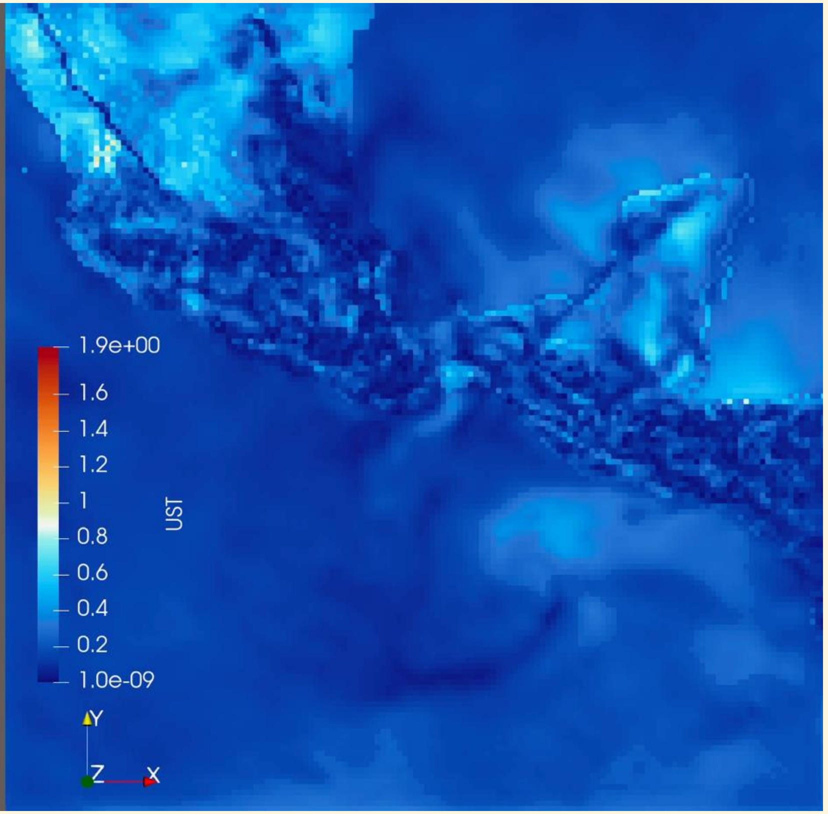
ECMWF Reanalysis v5 (ERA5)

Running WRF Models

Model different parameter scenarios but not impossible scenarios



Average Temperatures

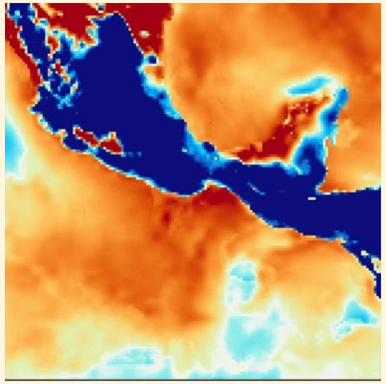


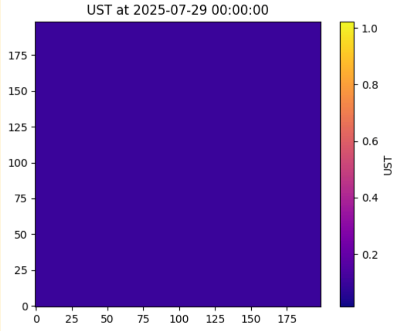
+2°C Atmosphere, +1°C Ocean

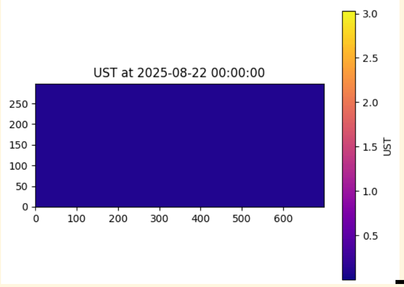
(UST = Friction Velocity)

Generate High Volume of Synthetic Data

Different parameter scenarios

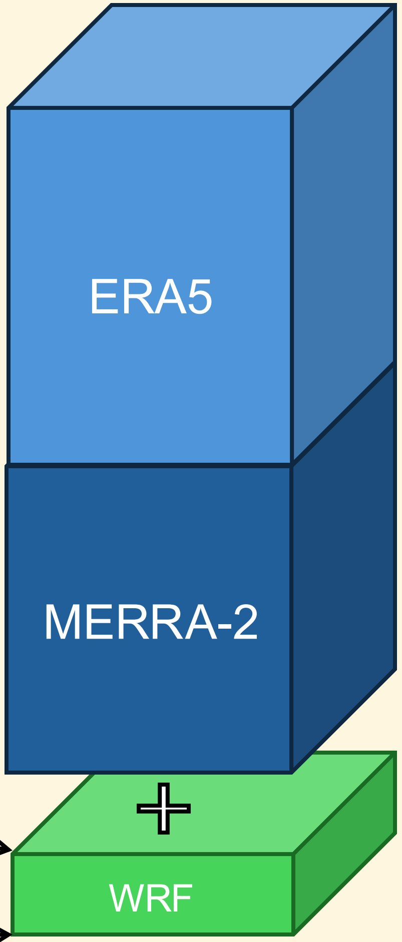






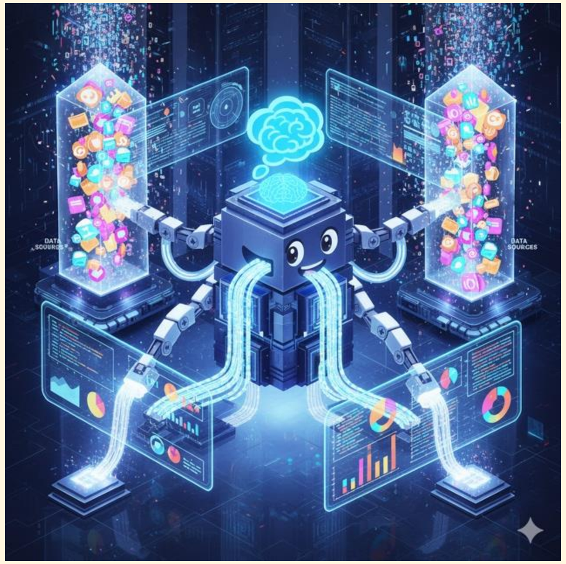
1000+ TB's

1-10 TB's



Historic Data

Synthetic Data

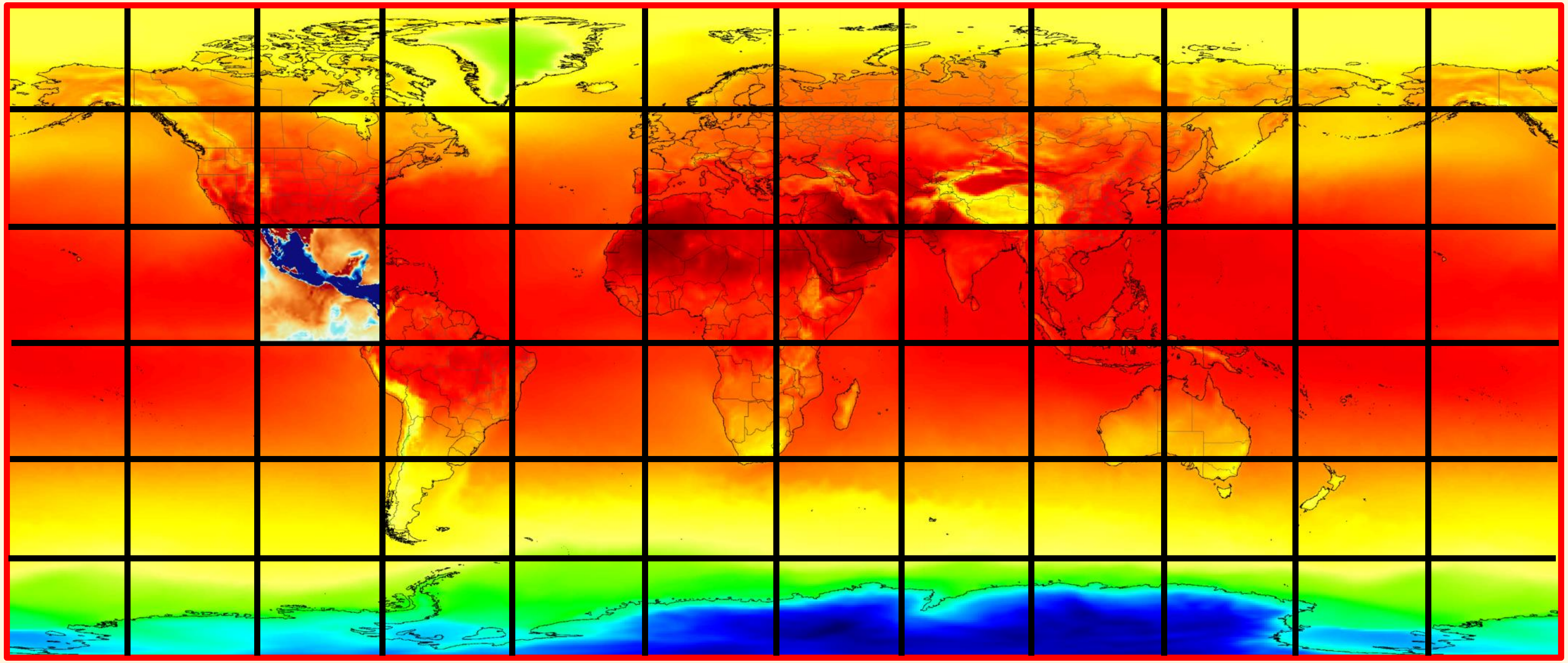


Merge or Fine-tune ML Model

- Different parameter scenarios

Temperature

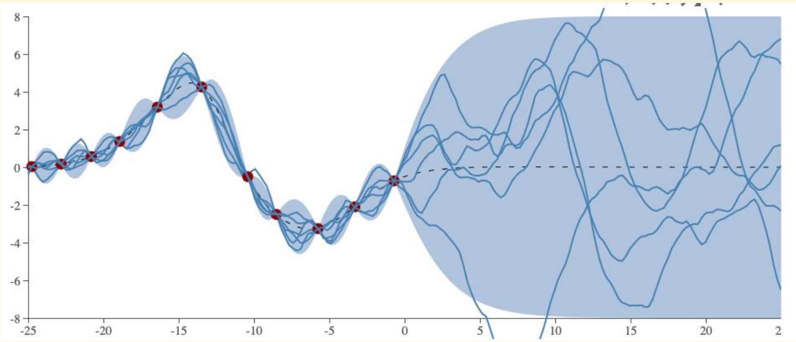
The Merra-2 Data Cube



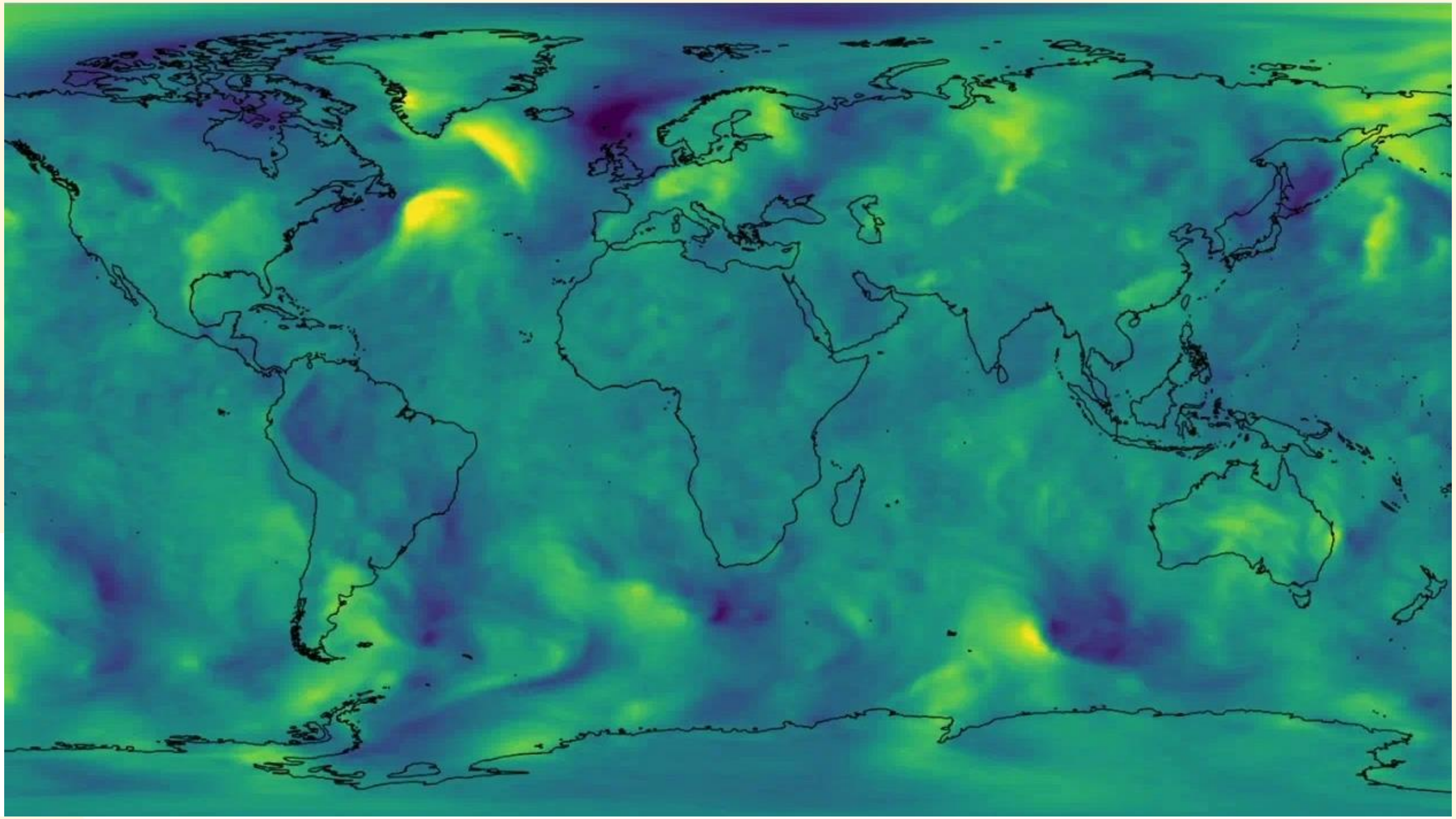
Improving Mid-Range Forecasts

Extrapolation is hard!

True representation lost at longer rollouts, so mean of ensembles used.



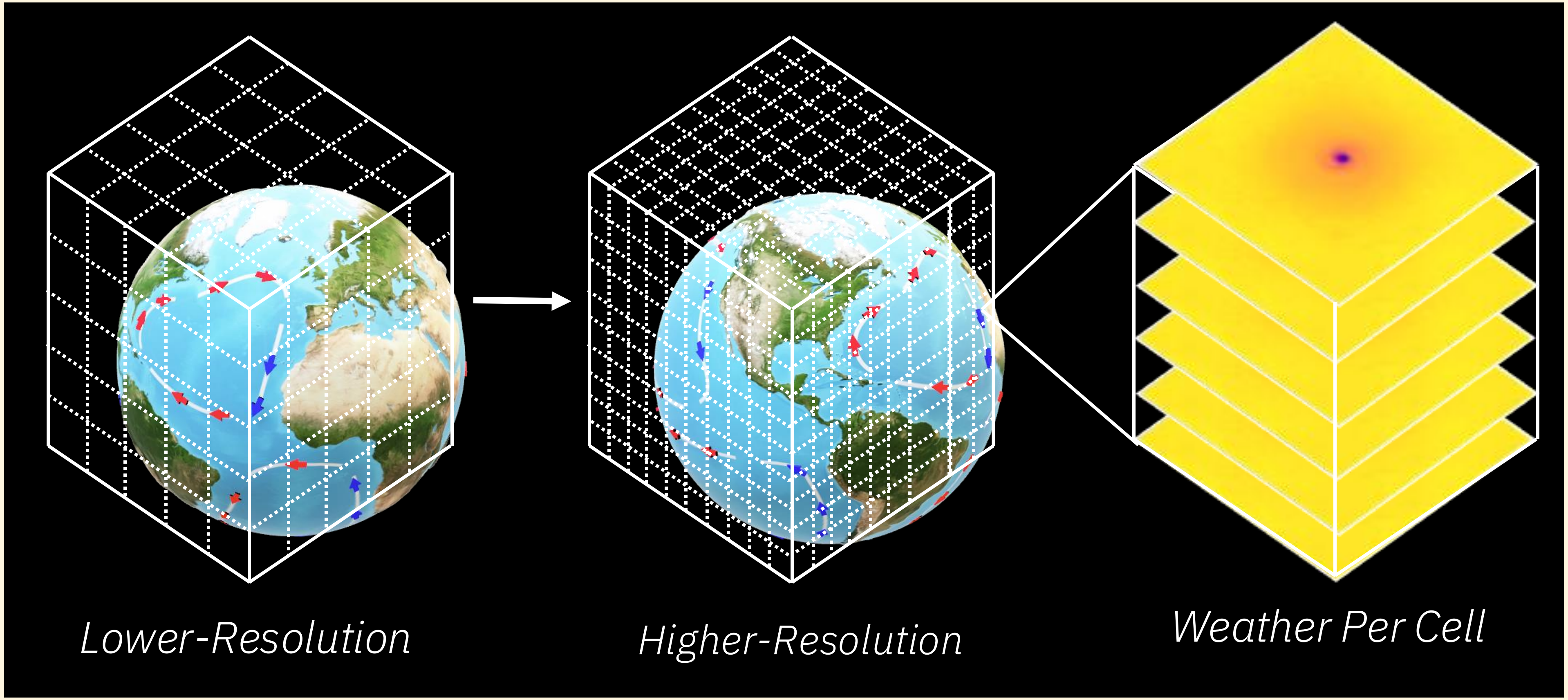
A Gaussian Process Extrapolating



A 10-day wind forecast made with the European Artificial Intelligence Forecasting System. EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

High Fidelity Resolutions

- A more chaotic nature at smaller distances
- Much more computationally intensive

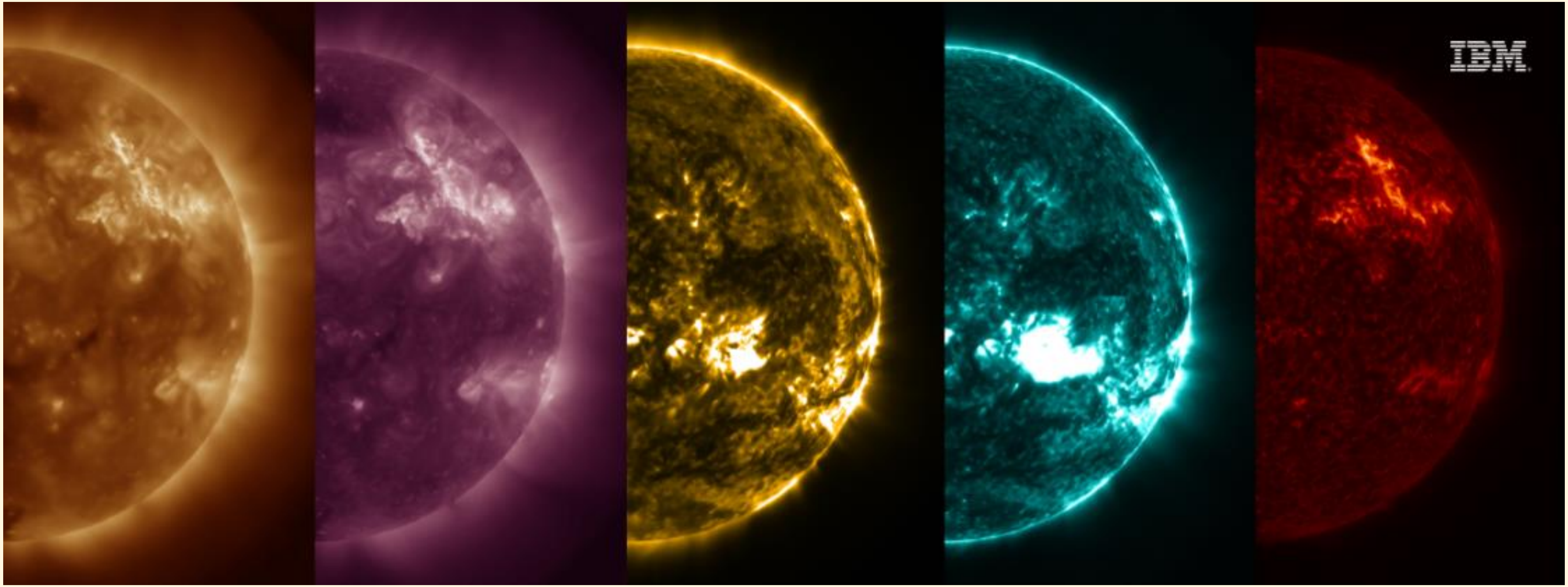


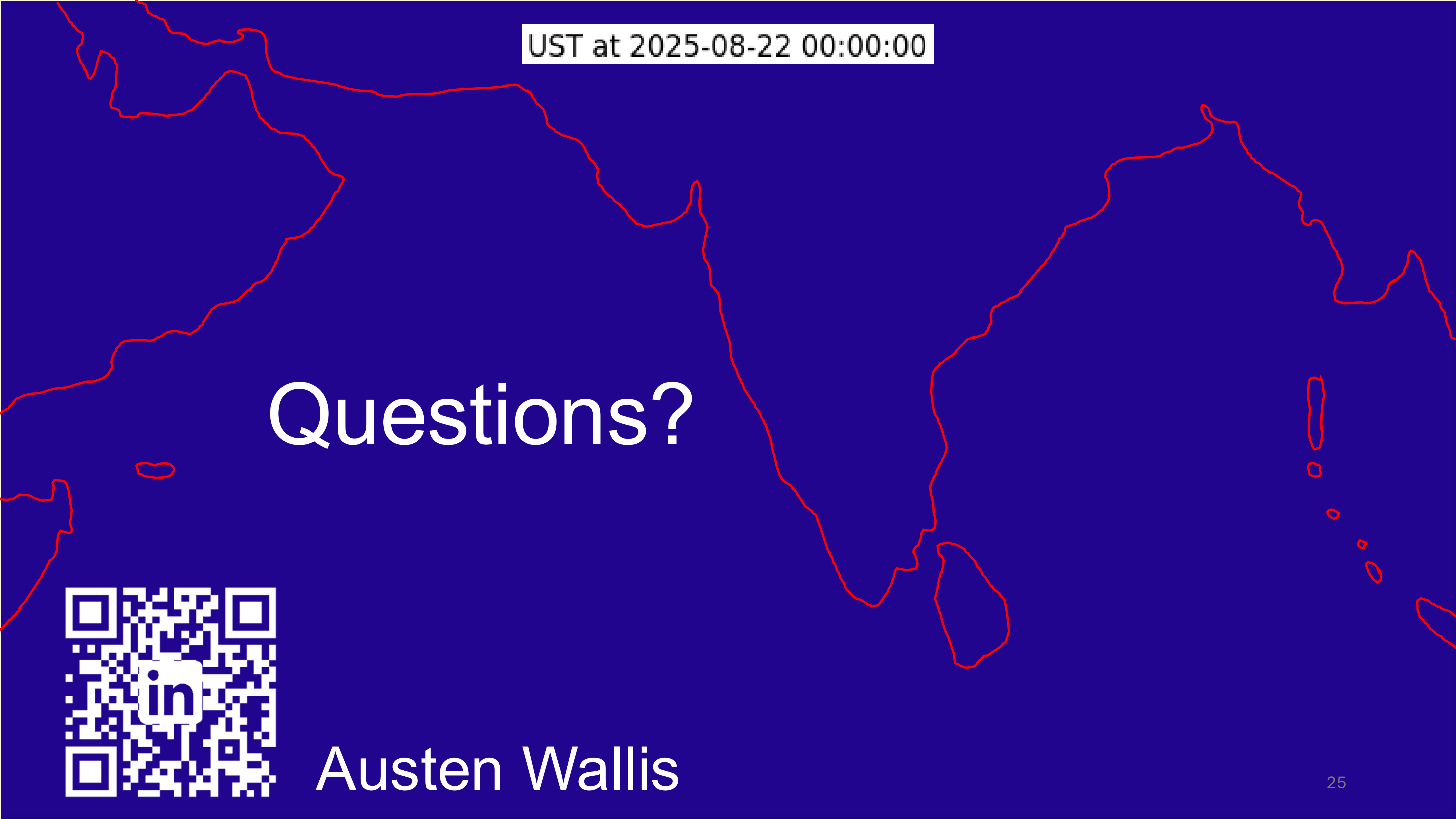
Benchmarking Unseen Events

How

Solar Flare Predictions

Space weather predictions with IBM's Suyra model





Questions?

Austen Wallis

Abstract

Machine learning thrives on data, and the common wisdom is simple: *the more, the better*. A widely used approach to expand training datasets is **data augmentation**, which can enhance model robustness. Yet, augmentation is not without risk—it can also degrade performance by introducing noise. More critically, there are cases where the required training data simply does not exist, especially when models are applied to previously unseen regimes.

Nowhere is this more relevant than in the context of improving weather model predictions for extreme events, such as hurricanes. Climate change is driving our atmosphere into uncharted territory, producing extremes absent from recorded history. How, then, can we trust model predictions under such conditions? One promising avenue is through the generation of synthetic data. Hence, in this talk, we will explore how weather data can be simulated, the ML model architectures we use that are designed to forecast extreme events, and the growing competition to build the most powerful foundational weather model.