

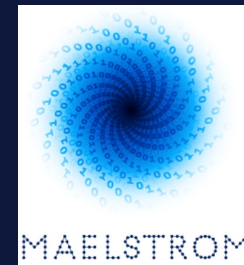
# Machine learning to make Earth observations ever more important for numerical weather and climate prediction

Peter Dueben

Head of the Earth System Modelling Section



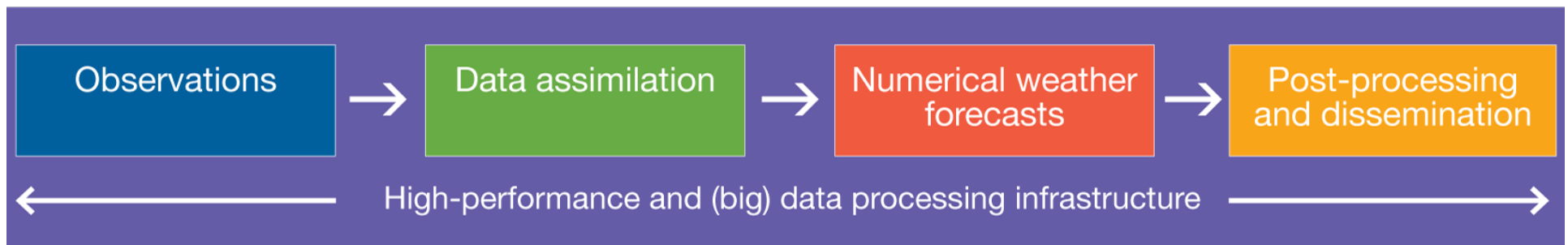
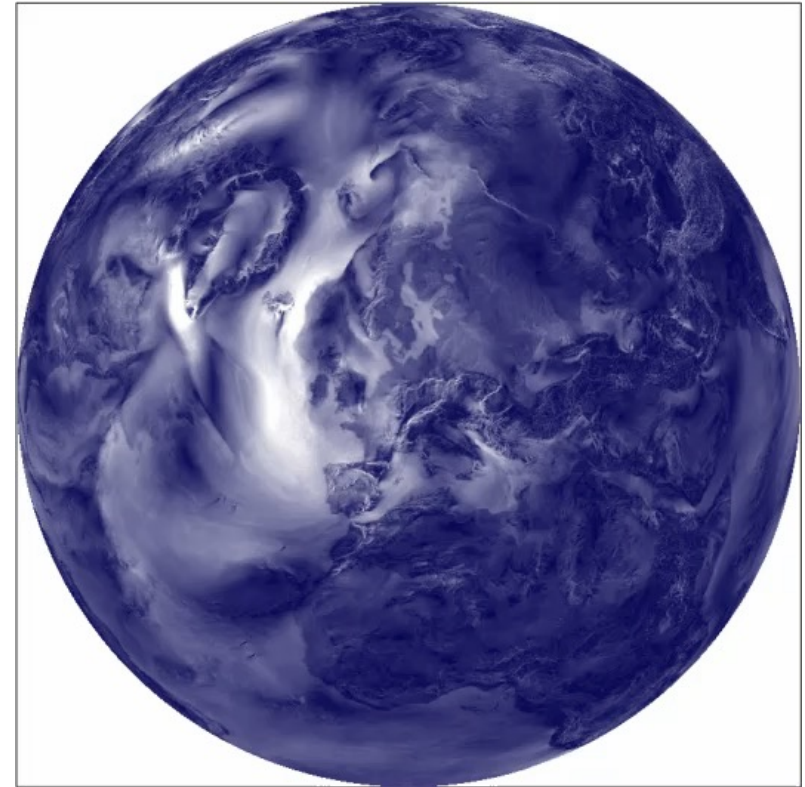
The strength of a common goal



The MAELSTROM and ESiWACE projects have received funding from the EuroHPC-Joint Undertaking under grant agreement No 955513 and 101093054.

# Numerical weather predications

Special thanks to Simon Lang



# The conventional approach to build a weather or climate model

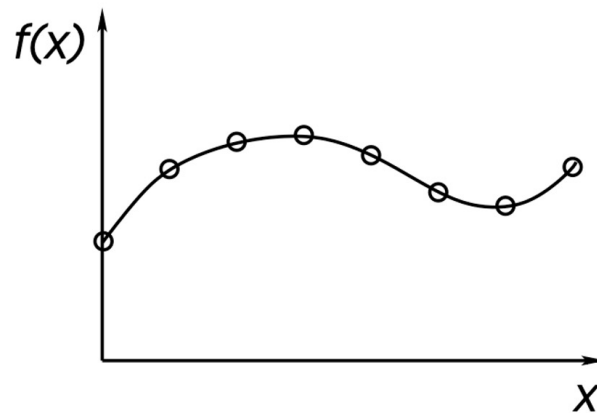
## 1. Derive the equations

The total mass inside the volume is given by

## 2. Define a grid

Atlas: a

## 3. Discretise the equations on the grid



$$\frac{\partial f(x_0, t)}{\partial x} \approx \frac{f(x_0 + \Delta x, t) - f(x_0 - \Delta x, t)}{2\Delta x},$$

$$\frac{\partial f(x_0, t)}{\partial x} \approx \frac{f(x_0 + \Delta x, t) - f(x_0, t)}{\Delta x},$$

$$\frac{\partial f(x, t_0)}{\partial t} \approx \frac{f(x, t_0) - f(x, t_0 - \Delta t)}{\Delta t}.$$

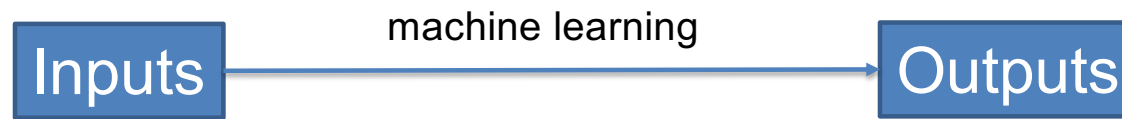
We discretise our function  $f(x)$  at specific grid points  $f(0)$ ,  $f(\Delta x)$ ,  $f(2\Delta x)$ ...

Derivatives are described by differential quotients  
→ There are plenty of different discretisation schemes

We need to discretise in both space and time

- The resulting model will produce realistic Earth systems at all lead times
- Observations are needed to tune model parameters and to generate initial conditions

So, how can machine learning help?



# Phase I: 2018-2021 – Explore the space

## Improve understanding

- Fuse information content from different datasources
- Unsupervised learning
- Causal discovery
- AI powered visualisation
- Uncertainty quantification
- ...

## Speed up simulations and green computing

- Emulate model components
- Port emulators to heterogeneous hardware
- Use reduced numerical precision and sparse machine learning
- Optimise HPC and data workflow
- Data compression
- ...

## Improve models

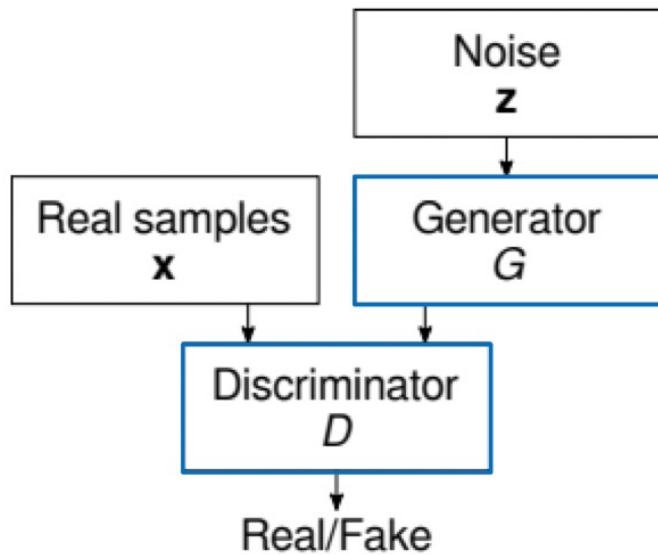
- Learn components from observations
- Correct biases
- Quality control of observations and observation operators
- Feature detection
- ...



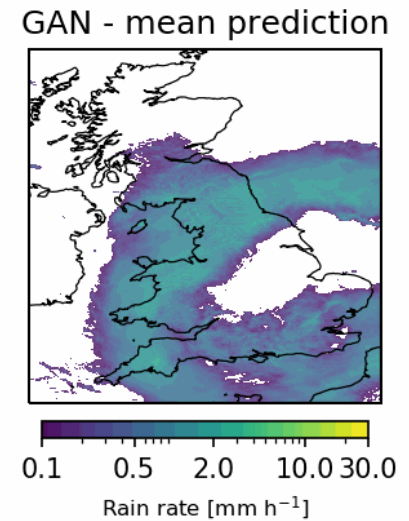
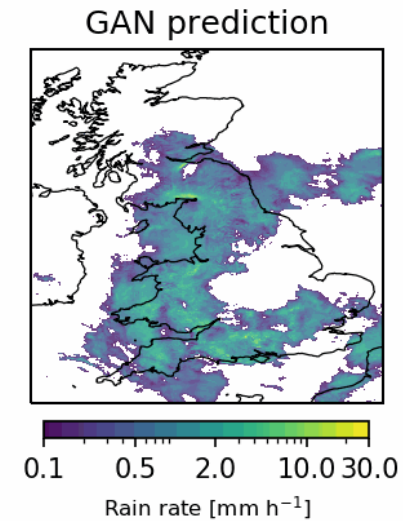
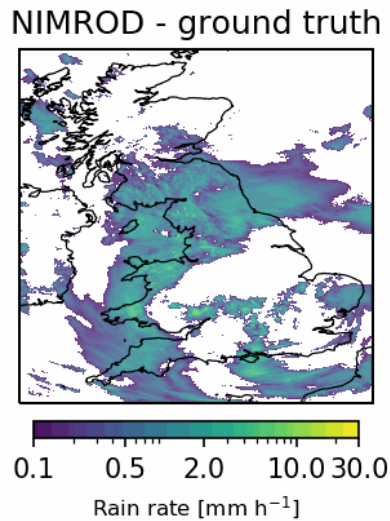
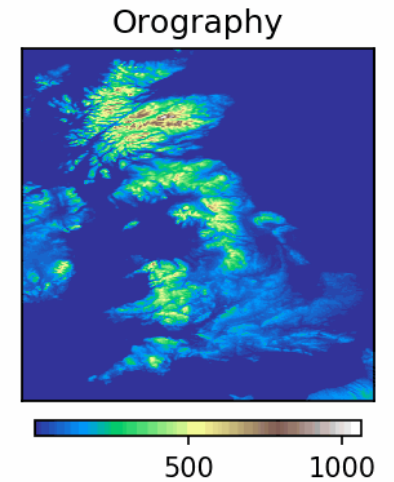
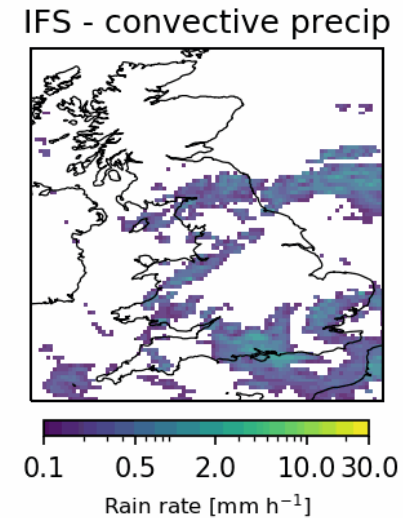
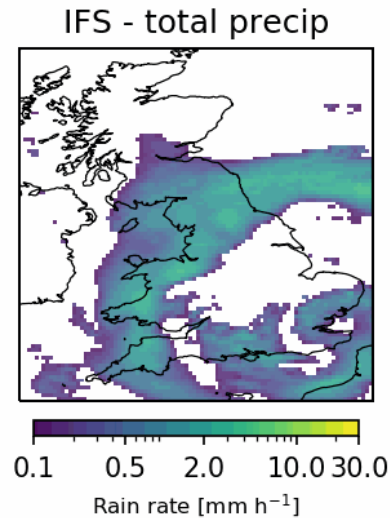
## Link communities

- Health – e.g. for predictions of risks
- Energy – e.g. for local downscaling
- Transport – e.g. to combine weather and IoT data
- Pollution – e.g. to detect sources
- Extremes – e.g. to predict wild fires
- ...

# Application example: Downscaling with Generative Adversarial Networks



**Input:** IFS Model Simulation fields on coarse (9 km) grid  
**Output:** Precipitation observation on fine (1 km) grid



# Will machine learning replace conventional modelling?



If Hamlet were a musical comedy, rather than a tragedy, how would it have ended? Describe the plot.



If Hamlet were a musical comedy, rather than a tragedy, the ending would likely be much happier and more lighthearted. Instead of ending in a violent and tragic showdown, with the deaths of several major characters, the play would likely have a more comedic and satisfying resolution.



For example, the play could end with Hamlet successfully avenging his father's death, but then choosing to forgive and reconcile with his uncle, the new king. This would allow Hamlet to put aside his grudges and bitterness, and to find peace and happiness.

As part of this resolution, there could be a big, energetic finale song, where the characters sing and dance together, expressing their newfound sense of joy and unity. This could be a lively and upbeat number, with a catchy and infectious melody, and lyrics that reflect the themes and events of the play, while also expressing a sense of hope and optimism for the future.

Overall, a musical comedy version of Hamlet would likely have a more uplifting and optimistic ending, with the main character achieving a sense of closure and resolution, and moving on with his life in a positive and fulfilling way.

The New York Times

SUBSCRIBE

THE SHIFT

## An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy.

"I won, and I didn't break any rules," the artwork's creator says.

Give this article 1.9K



Jason Allen's A.I.-generated work, "Théâtre D'opéra Spatial," took first place in the digital category at the Colorado State Fair. via Jason Allen

# Can we replace conventional weather forecast systems by deep learning?

## **We could base the entire model on neural networks and trash the conventional models.?**

There are limitations for existing models and ECMWF provides access to hundreds of petabytes of data

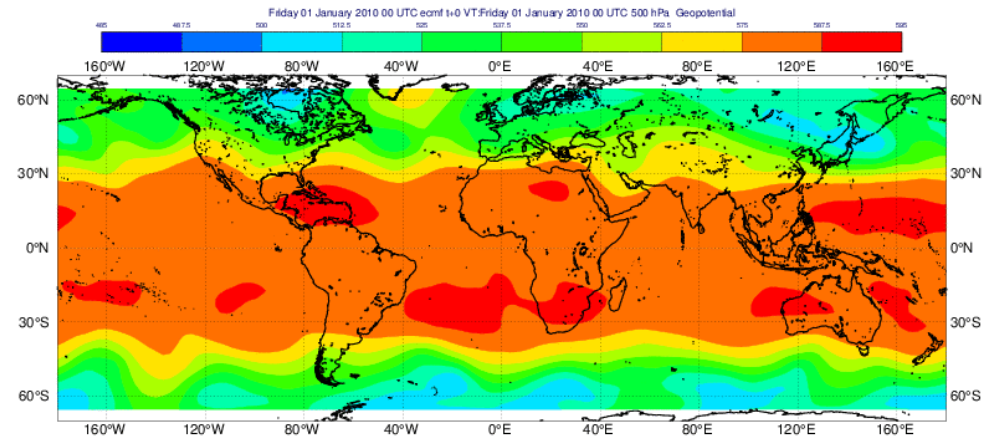
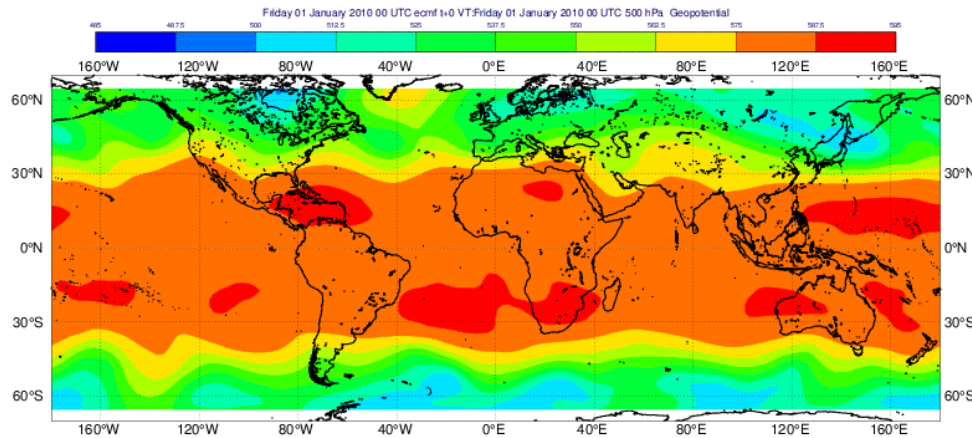
### **A simple test configuration:**

- We retrieve historical data (ERA5) for geopotential at 500 hPa (Z500) for the last decades (>65,000 global data sets)
- We map the global data to a coarse two-dimensional grid (60x31)
- We learn to predict the update of the field from one hour to the next using deep learning
- Once we have learned the update, we can perform predictions into the future

**No physical understanding is required!**

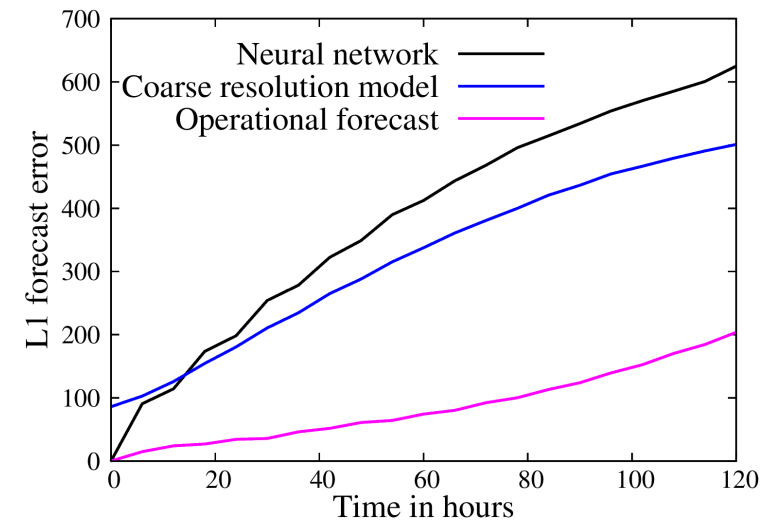


# Can we replace conventional weather forecast systems by deep learning?



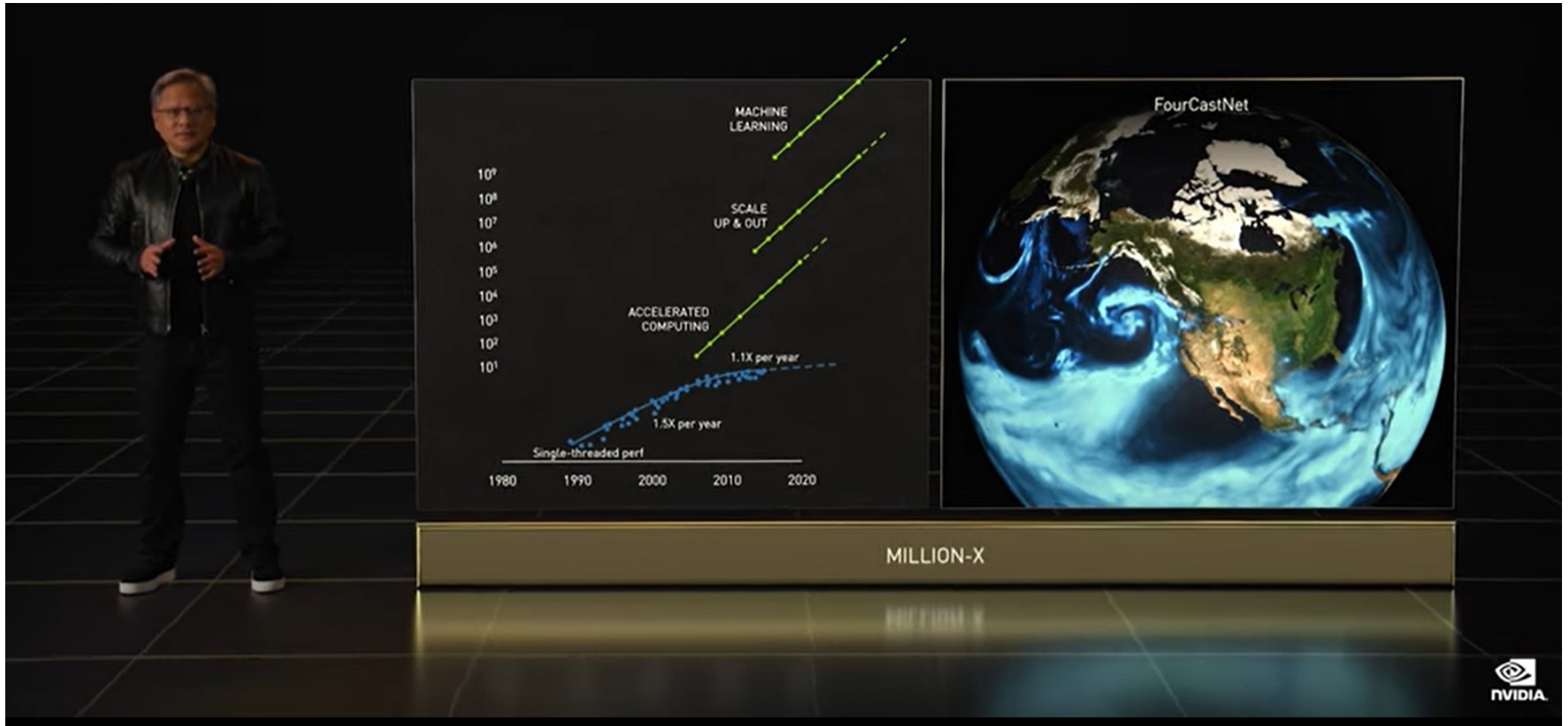
Time evolution of Z500 for historic data and a neural network prediction.  
**Can you tell which one is the neural network?**

- The neural network is picking up the dynamics nicely.
- Forecast errors are comparable if we compare like with like.



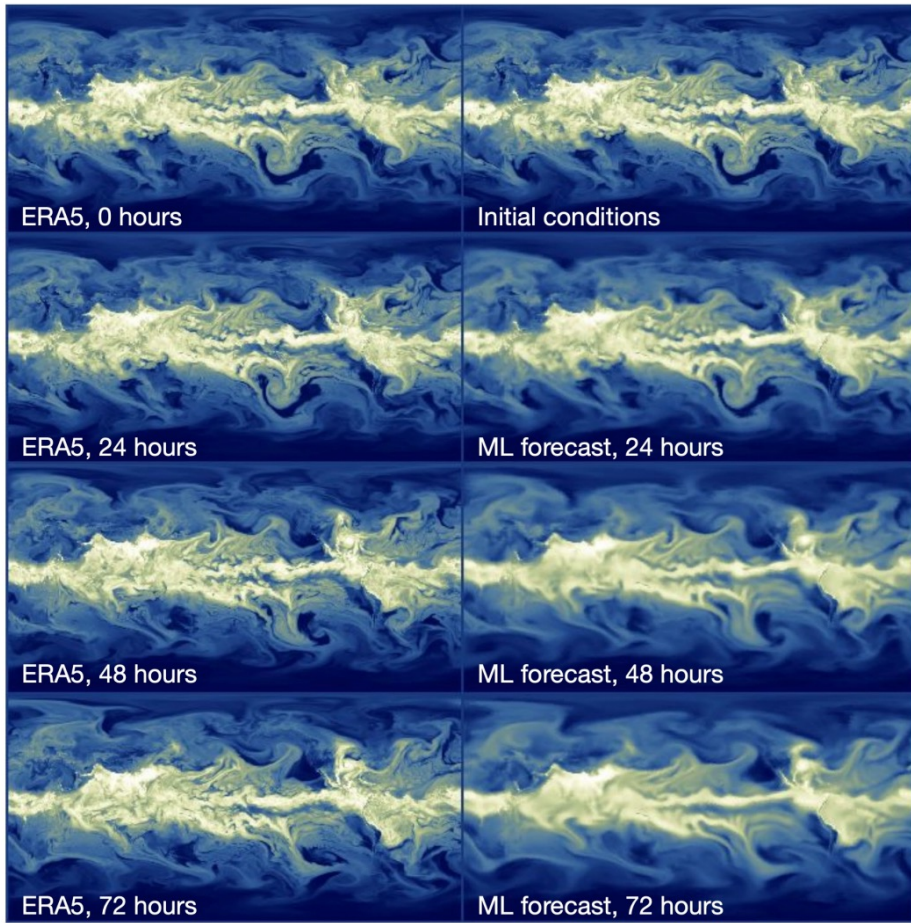
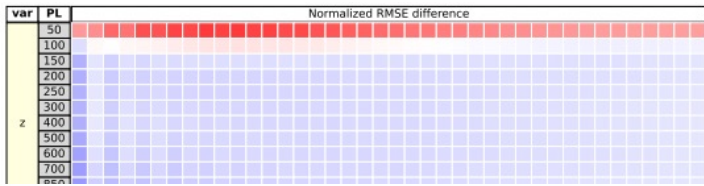
Dueben and Bauer GMD 2018

## Phase X: A full machine learning model for weather and climate



NVIDIA's Earth-2 is coming with FourCastNet

# Phase X: A full machine learning model for weather and climate



GraphCast from Google/Deepmind is beating conventional weather forecast model in deterministic scores.

But how do these models actually work?

They get the best results when using very large timesteps (6h vs. 600s) and a couple of the previous timesteps as input.

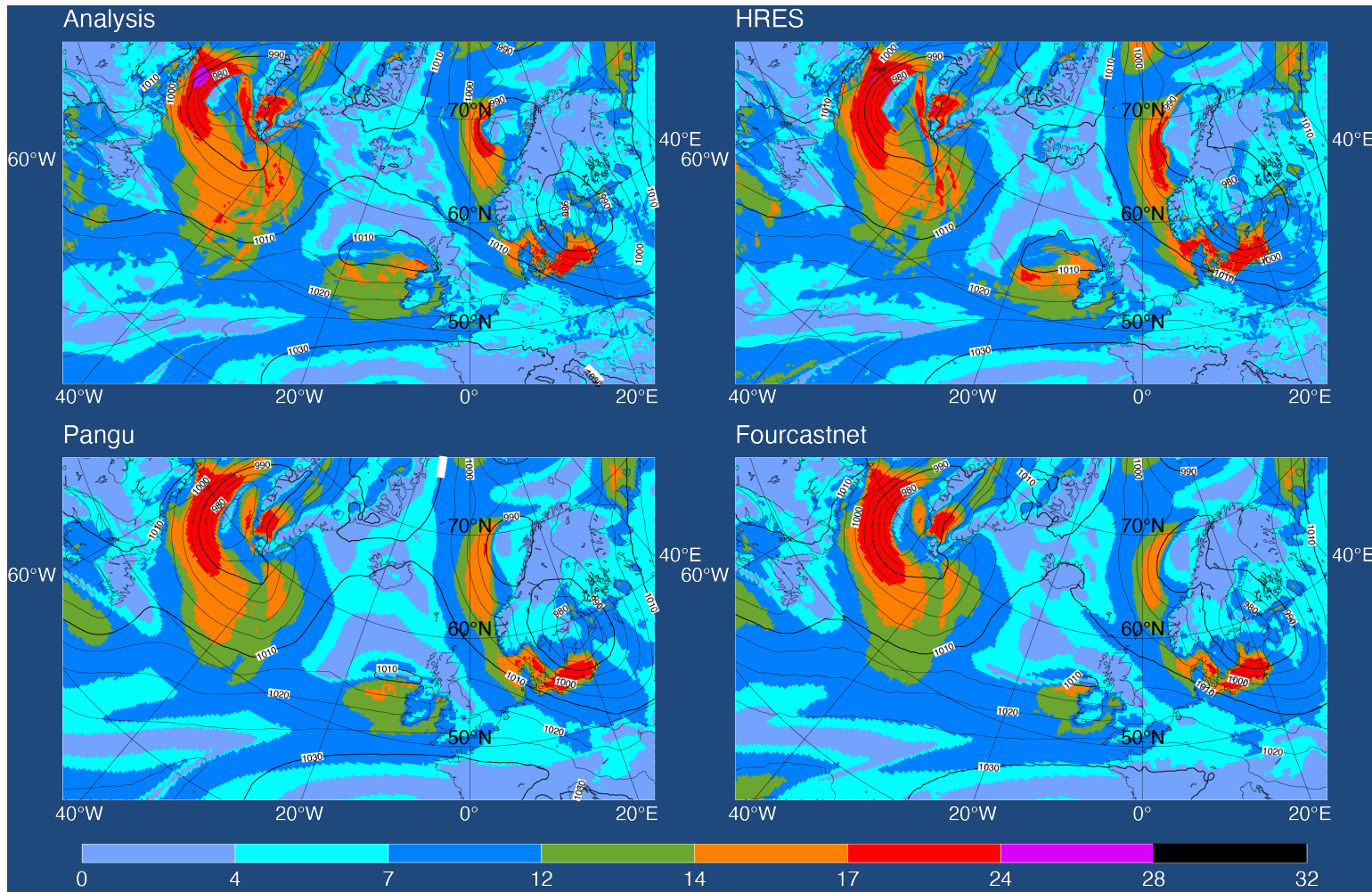
They are trained for a small Root Mean Square Error.  
→ They smear out for large lead times.

Can they extrapolate? Learn uncertainty? Learn from observations? Fill the state vector? Learn all important processes?

These models will not always represent a realistic Earth system and cannot be better than the training data.

Images from Keisler (2022)

# The forecasts are working very well: Storm Otto (18<sup>th</sup> Feb 2023 00UTC)



Linus Magnusson and Baudouin Raoult

# What is the number one need for machine learning for weather and climate prediction?

## It's the data!

You need the best data available to train.

At the moment, the most important datasets for training come from reanalysis (mainly ERA5) products that combine models and observations in the optimal way.

→ ECMWF and EUMETSAT need to further improve these products for the optimal machine learning predictions in the future

→ Every observation that is not taken today will be missing in the future

Many thanks!

Peter.Dueben@ecmwf.int

@PDueben



The strength of a common goal