Advances in machine learning for Earth system modelling

Peter Dueben

Head of the Earth System Modelling Section

The strength of a common goal

The MAELSTROM, ESiWACE and WeatherGenerator projects have received funding from Horizon Europe and the EuroHPC-Joint Undertaking under grant agreement No 955513, 101093054. and 101187947.

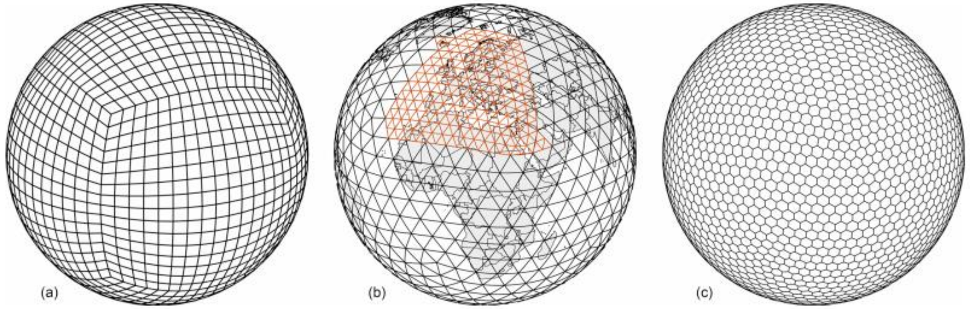
How do we build weather and climate models that are based on physics?

Find equations

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\,\mathbf{v} = -\frac{\nabla p}{\rho} + \nu \nabla^2 \mathbf{v} + \frac{\mathbf{F}}{\rho} - 2\boldsymbol{\Omega} \times \mathbf{v}$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0$$

Define grid



Discretise equations

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

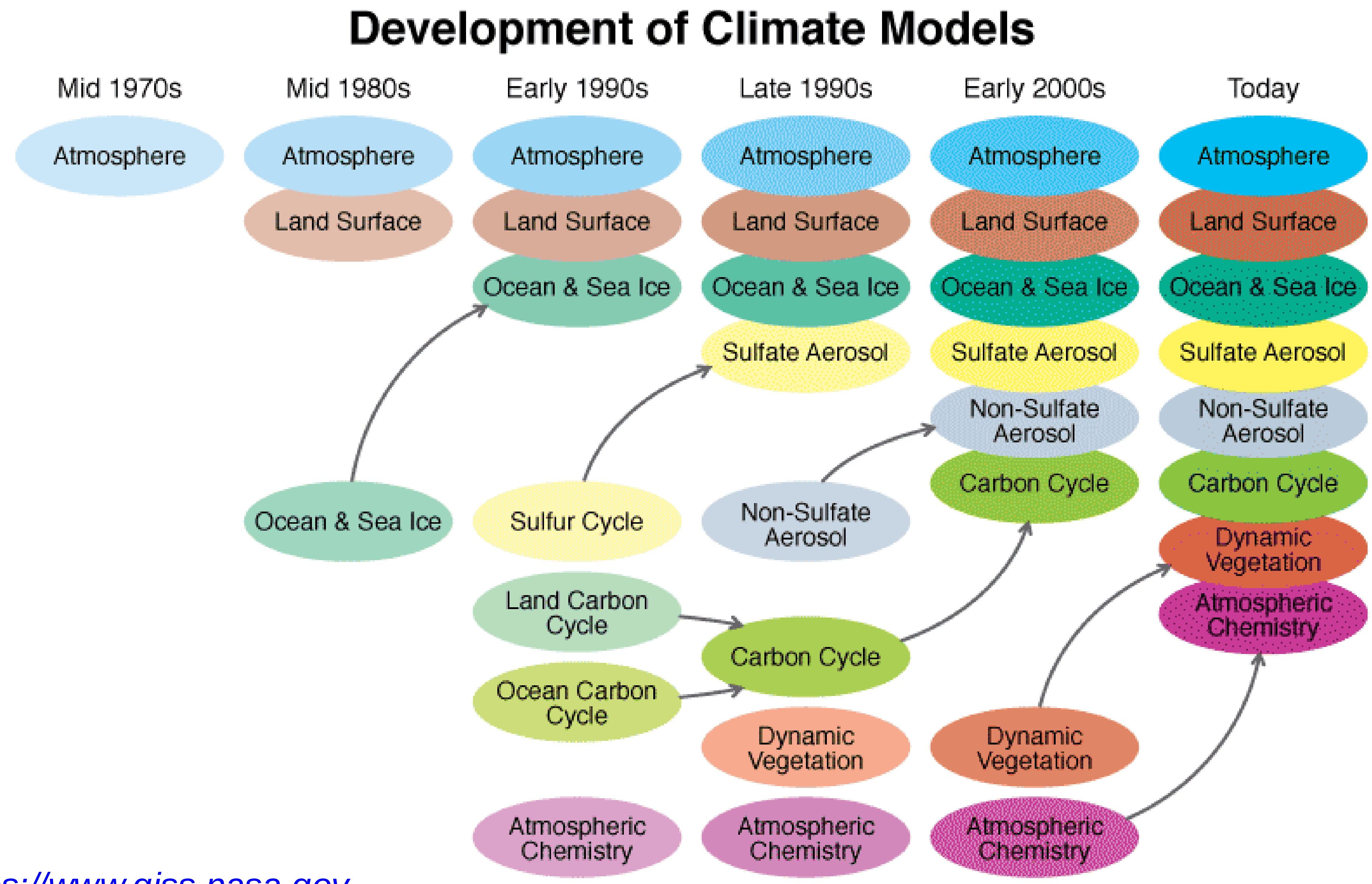
$$\frac{\partial f(x_0,t)}{\partial x} \approx \frac{f(x_0+\Delta x,t) - f(x_0,t)}{\Delta x},$$

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

Represent subgrid-scale

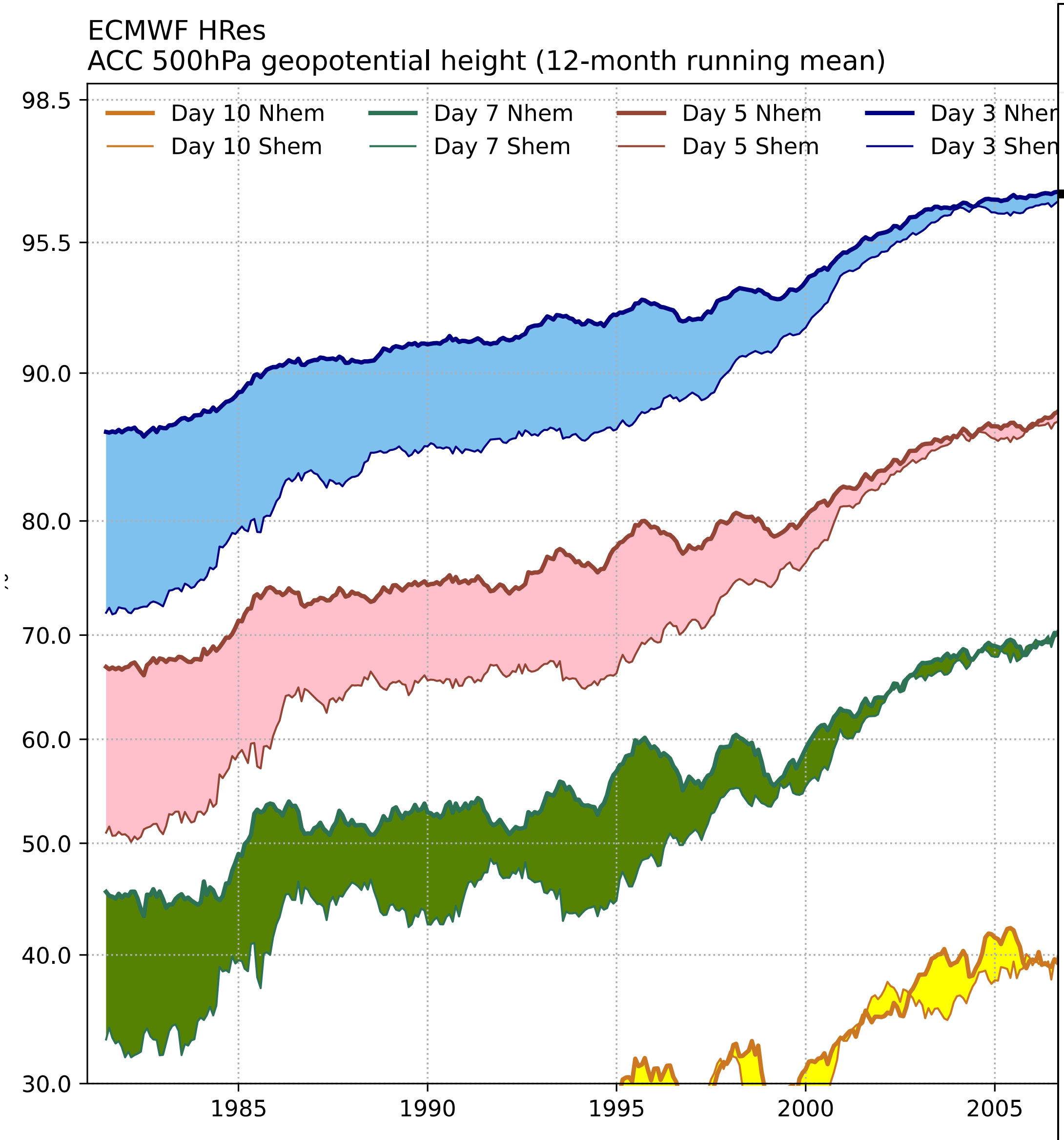
Constantly compare against observations

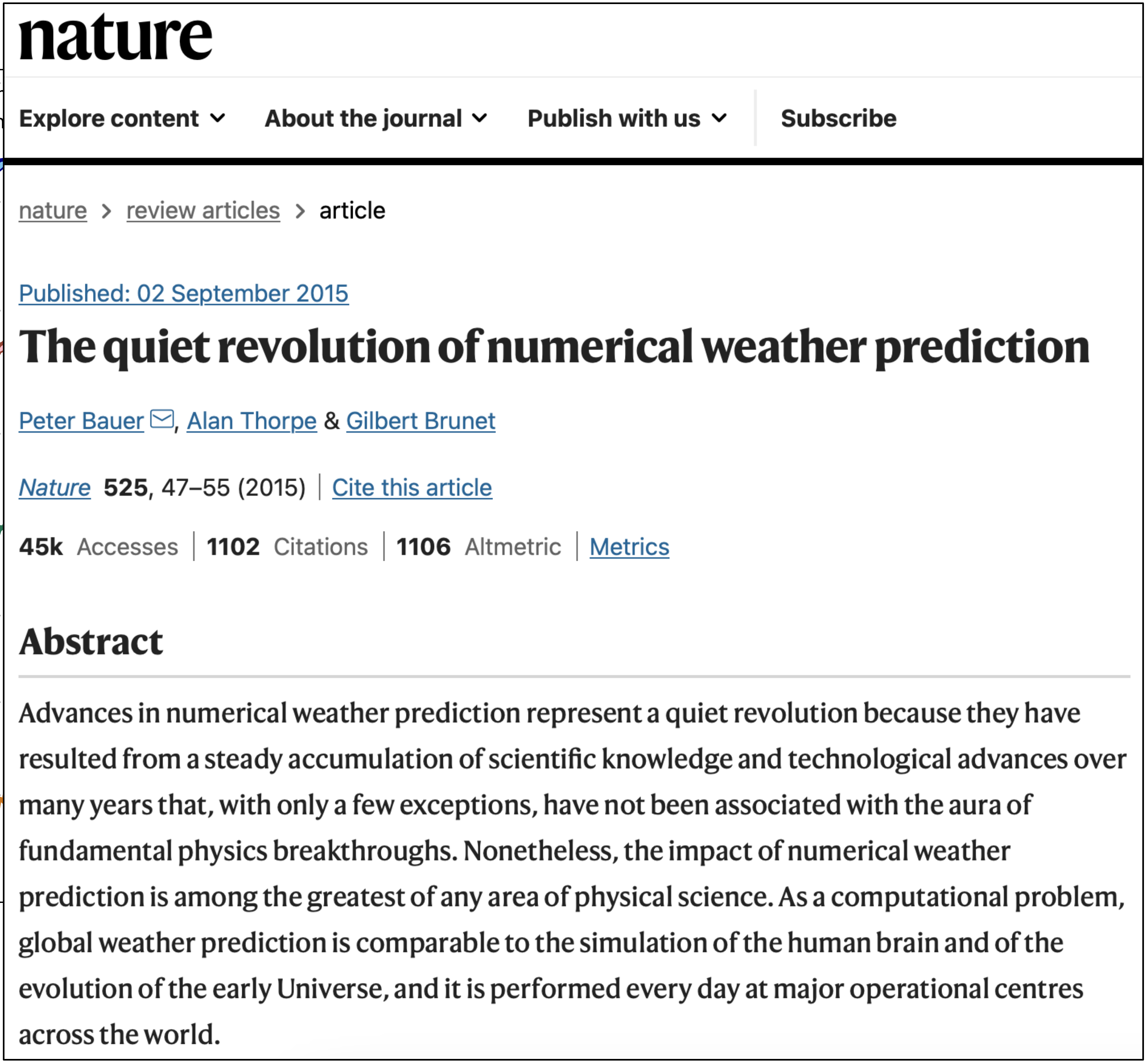
Earth System model complexity



Source: https://www.giss.nasa.gov

1980-2020: The quiet revolution of numerical weather prediction





nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄ Subscribe

nature > review articles > article

Published: 02 September 2015

The quiet revolution of numerical weather prediction

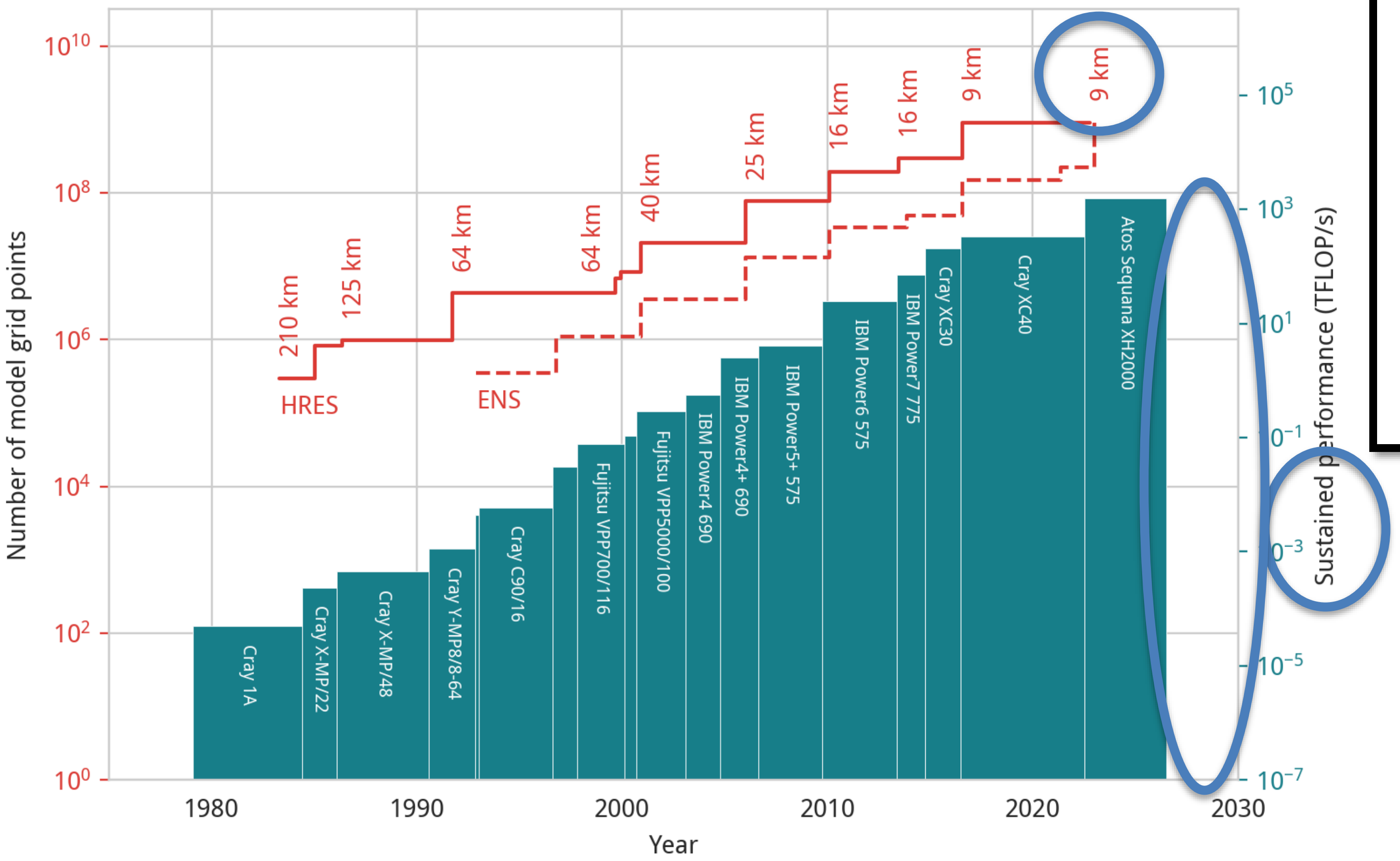
Peter Bauer, Alan Thorpe & Gilbert Brunet

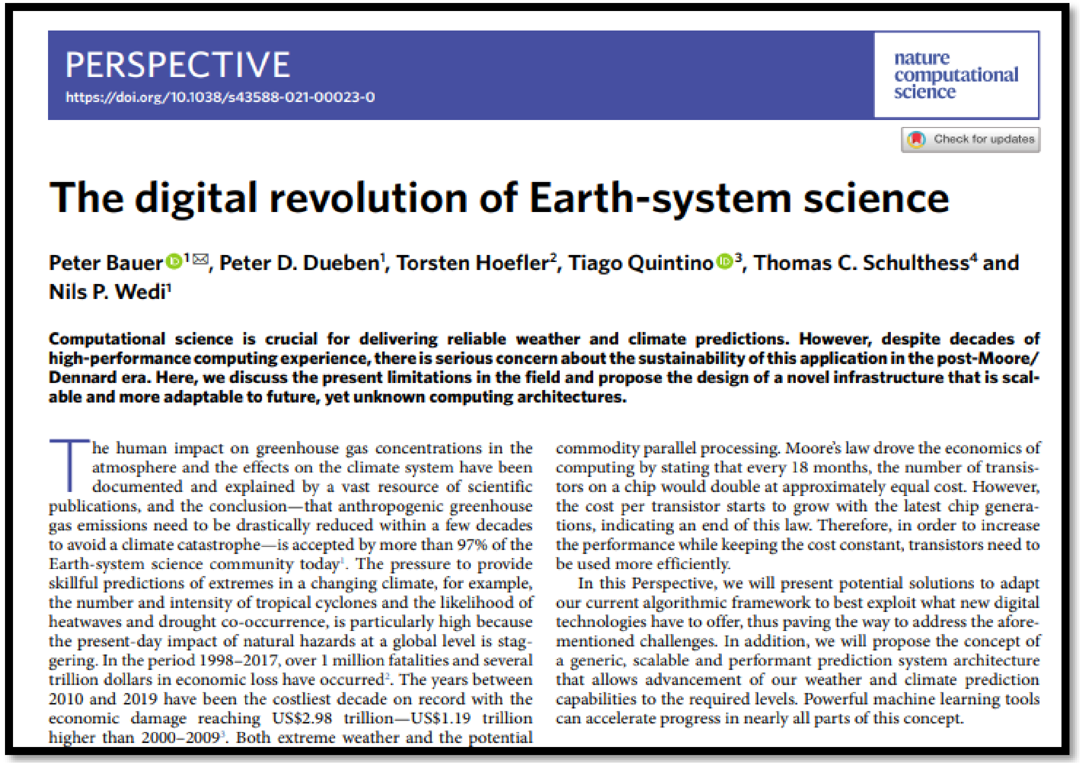
Nature **525**, 47–55 (2015) | Cite this article

Abstract

Advances in numerical weather prediction represent a quiet revolution because they have resulted from a steady accumulation of scientific knowledge and technological advances over many years that, with only a few exceptions, have not been associated with the aura of fundamental physics breakthroughs. Nonetheless, the impact of numerical weather prediction is among the greatest of any area of physical science. As a computational problem, global weather prediction is comparable to the simulation of the human brain and of the evolution of the early Universe, and it is performed every day at major operational centres across the world.

2015-today: The digital revolution





PERSPECTIVE
https://doi.org/10.1038/s43588-021-00023-0
nature computational science
Check for updates

The digital revolution of Earth-system science

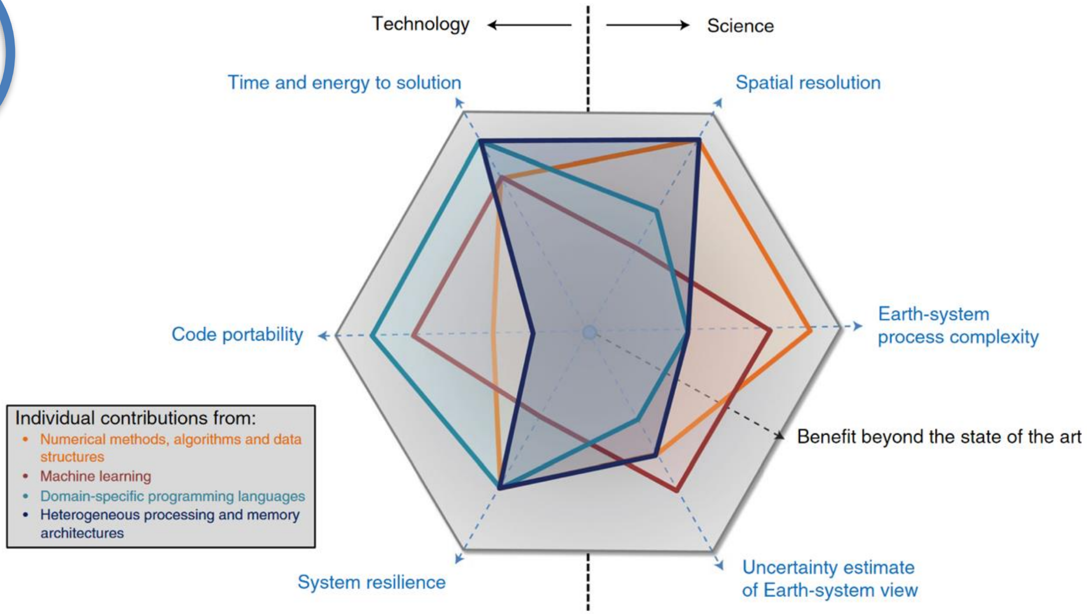
Peter Bauer[1]✉, Peter D. Dueben[1], Torsten Hoefler[2], Tiago Quintino[3], Thomas C. Schulthess[4] and Nils P. Wedi[1]

Computational science is crucial for delivering reliable weather and climate predictions. However, despite decades of high-performance computing experience, there is serious concern about the sustainability of this application in the post-Moore/Dennard era. Here, we discuss the present limitations in the field and propose the design of a novel infrastructure that is scalable and more adaptable to future, yet unknown computing architectures.

The human impact on greenhouse gas concentrations in the atmosphere and the effects on the climate system have been documented and explained by a vast resource of scientific publications, and the conclusion—that anthropogenic greenhouse gas emissions need to be drastically reduced within a few decades to avoid a climate catastrophe—is accepted by more than 97% of the Earth-system science community today[1]. The pressure to provide skillful predictions of extremes in a changing climate, for example, the number and intensity of tropical cyclones and the likelihood of heatwaves and drought co-occurrence, is particularly high because the present-day impact of natural hazards at a global level is staggering. In the period 1998–2017, over 1 million fatalities and several trillion dollars in economic loss have occurred[2]. The years between 2010 and 2019 have been the costliest decade on record with the economic damage reaching US\$2.98 trillion—US\$1.19 trillion higher than 2000–2009[3]. Both extreme weather and the potential

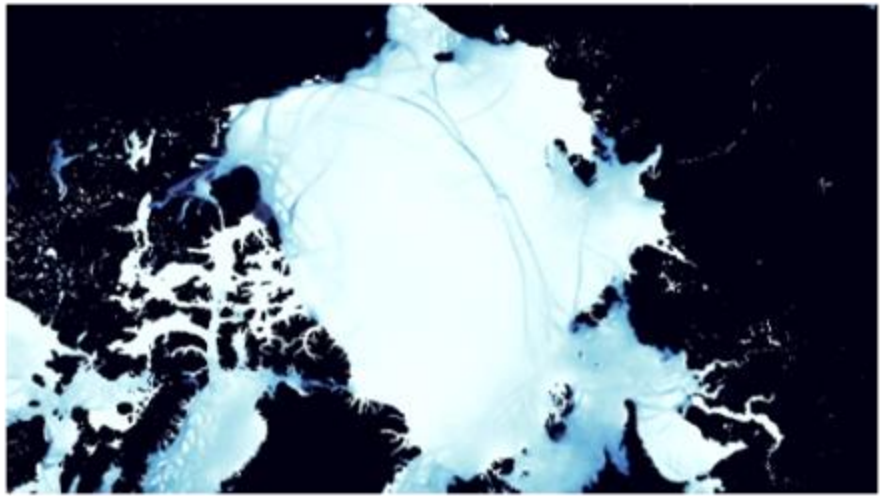
commodity parallel processing. Moore's law drove the economics of computing by stating that every 18 months, the number of transistors on a chip would double at approximately equal cost. However, the cost per transistor starts to grow with the latest chip generations, indicating an end of this law. Therefore, in order to increase the performance while keeping the cost constant, transistors need to be used more efficiently.

In this Perspective, we will present potential solutions to adapt our current algorithmic framework to best exploit what new digital technologies have to offer, thus paving the way to address the aforementioned challenges. In addition, we will propose the concept of a generic, scalable and performant prediction system architecture that allows advancement of our weather and climate prediction capabilities to the required levels. Powerful machine learning tools can accelerate progress in nearly all parts of this concept.

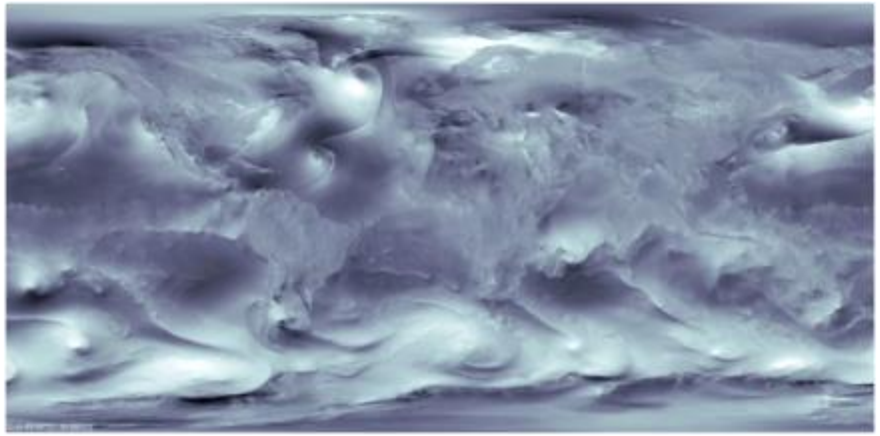


Destination Earth allowed to realise the Digital Revolution in Europe

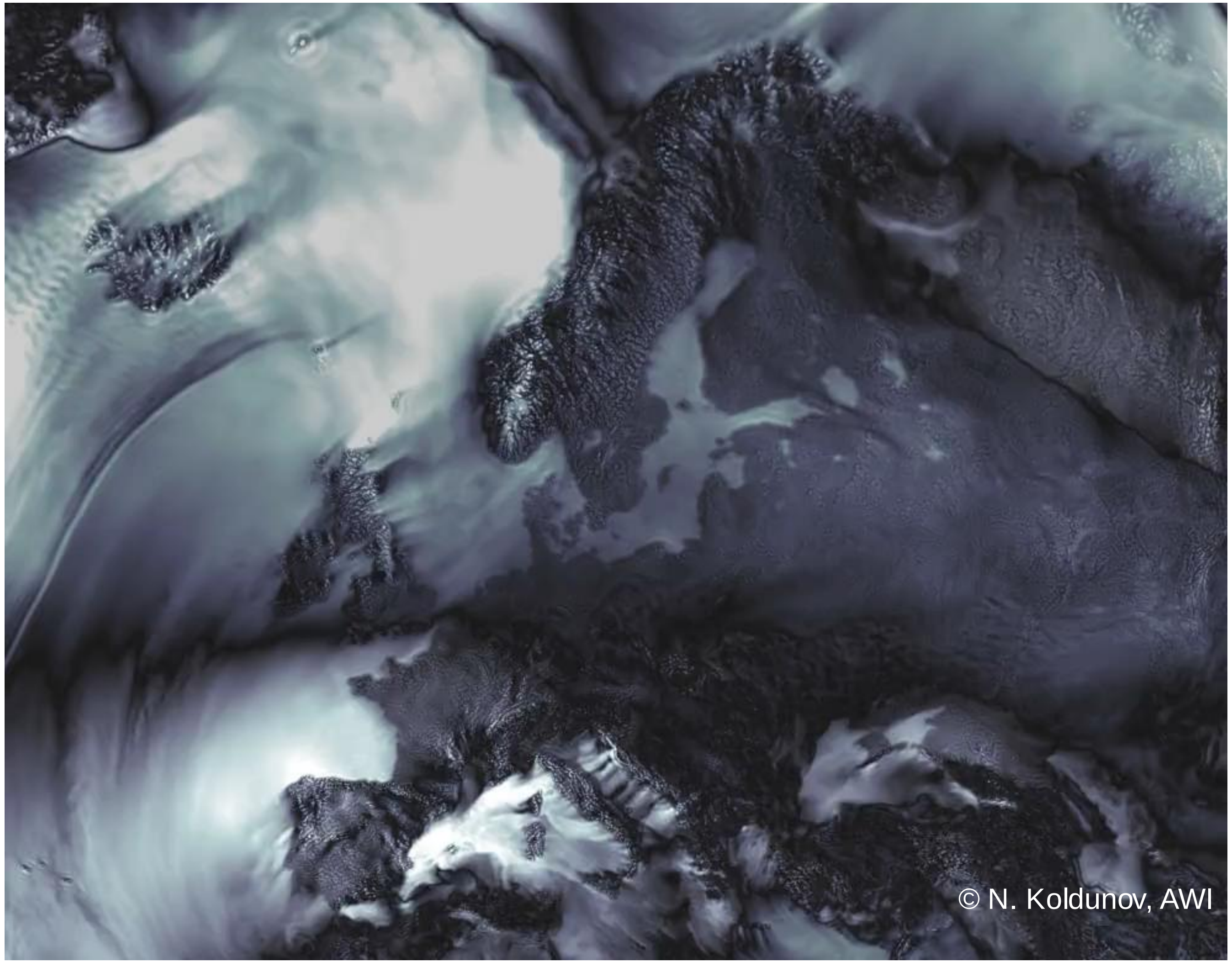
2015-today: The digital revolution to allow for km-scale models



More realistic at local scale



More realistic at global scale



Better results via a coupled model system

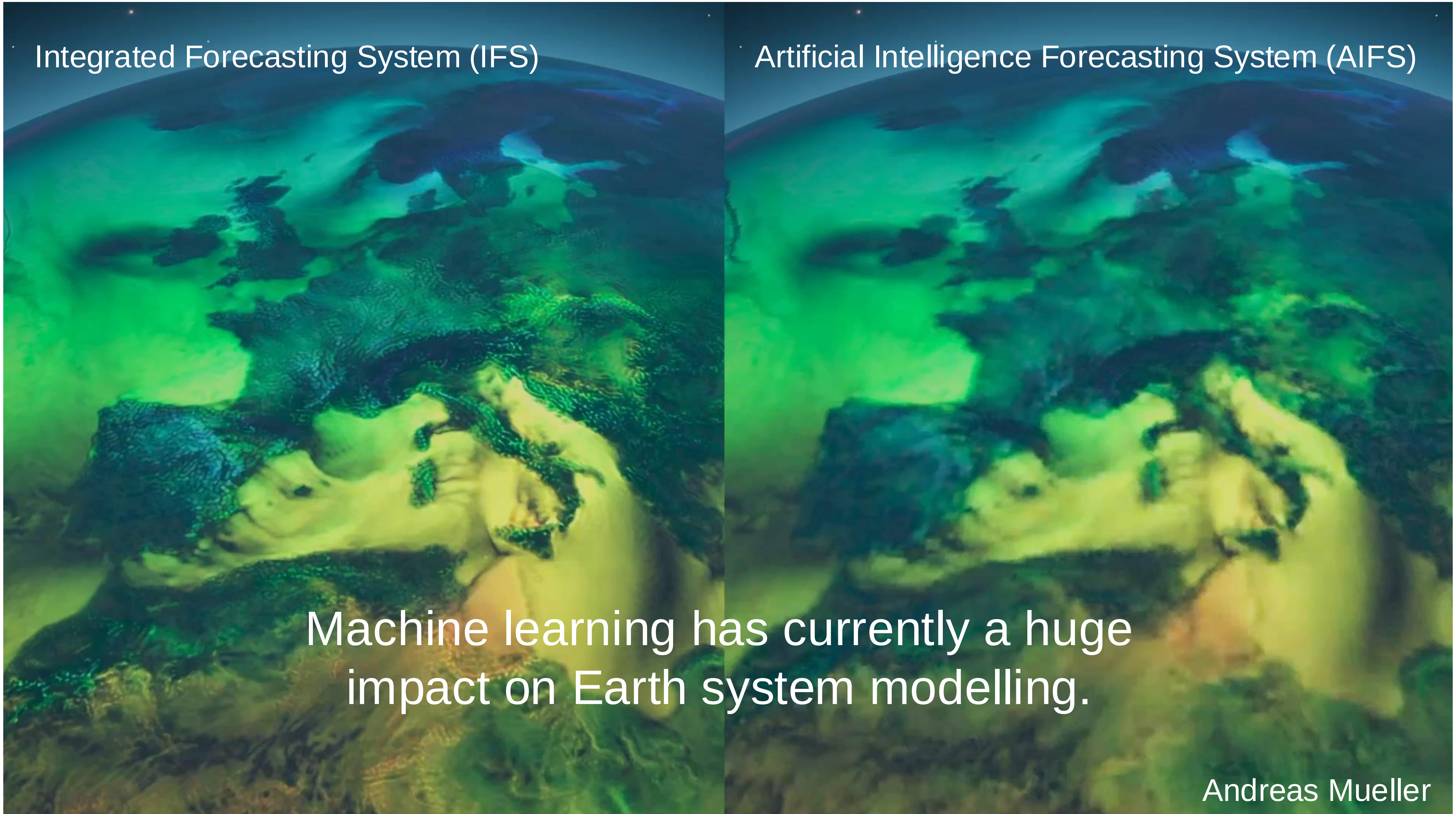
ECMWF

Global km-scale models improve realism of simulations significantly and are now becoming available.

Destination Earth | EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

next GEMS

Km-scale models improve model realism significantly, but this is not the end of the story…



2022-today: The machine learning revolution

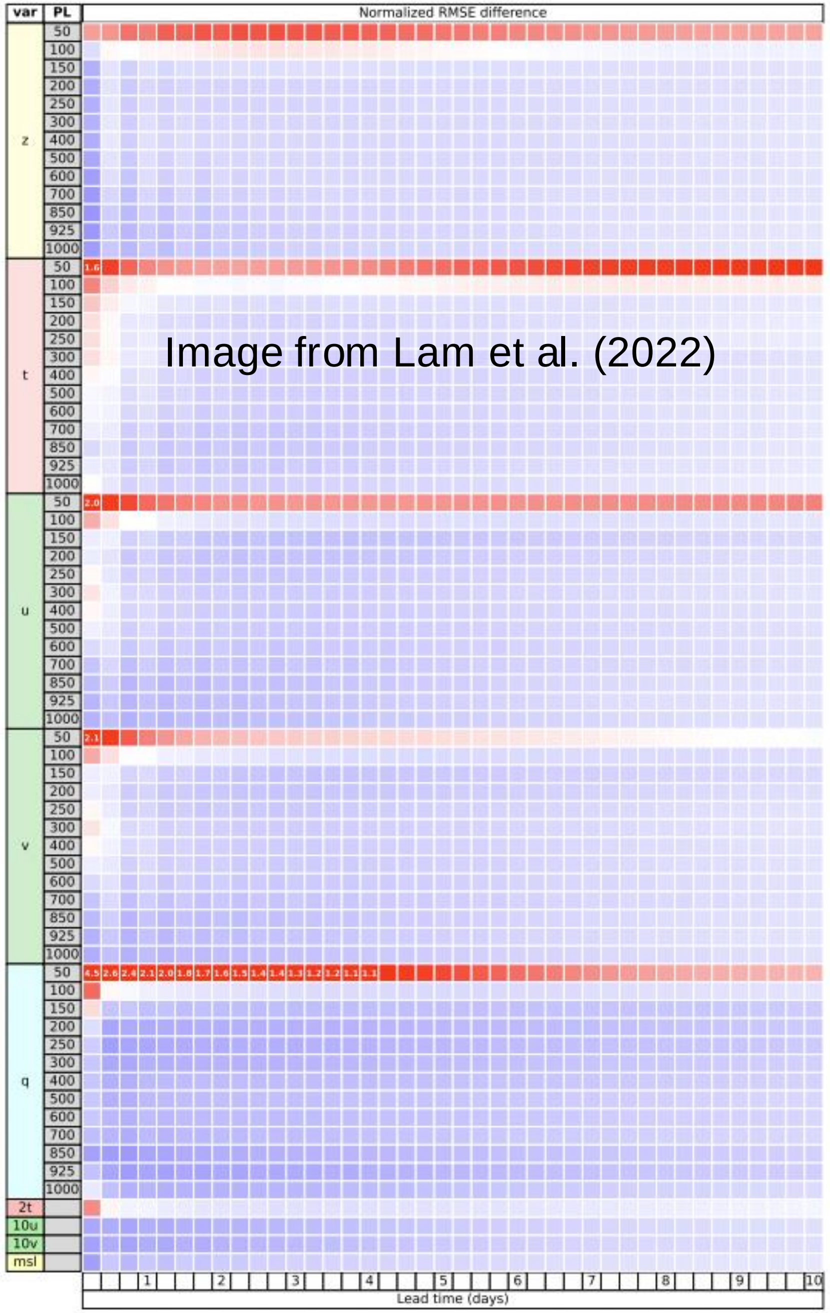
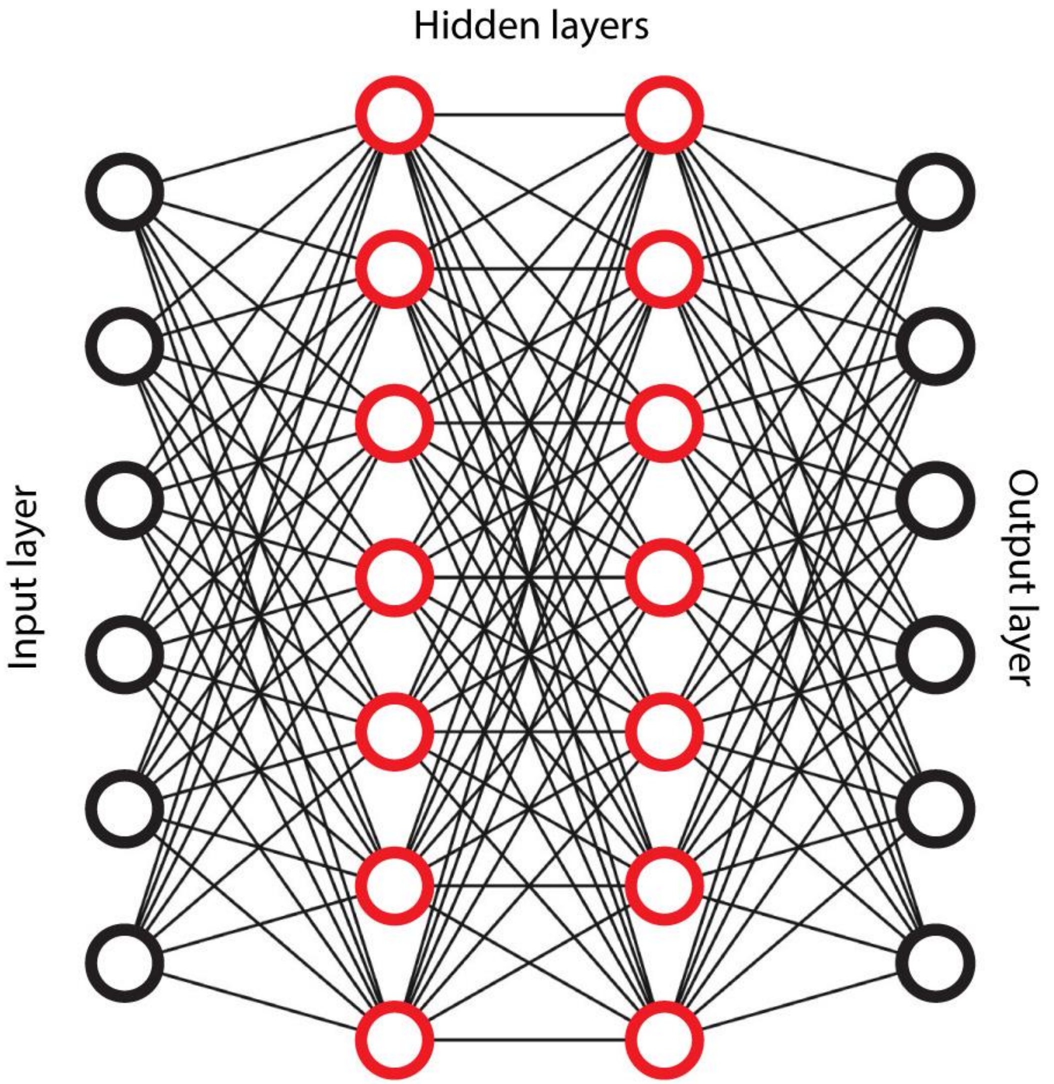


Image from Lam et al. (2022)

In 2022 machine learned forecast models from Google, NVIDIA and Huawei are beating conventional weather forecast models in deterministic scores and are orders of magnitude faster during inference.

2022-today: The machine learning revolution

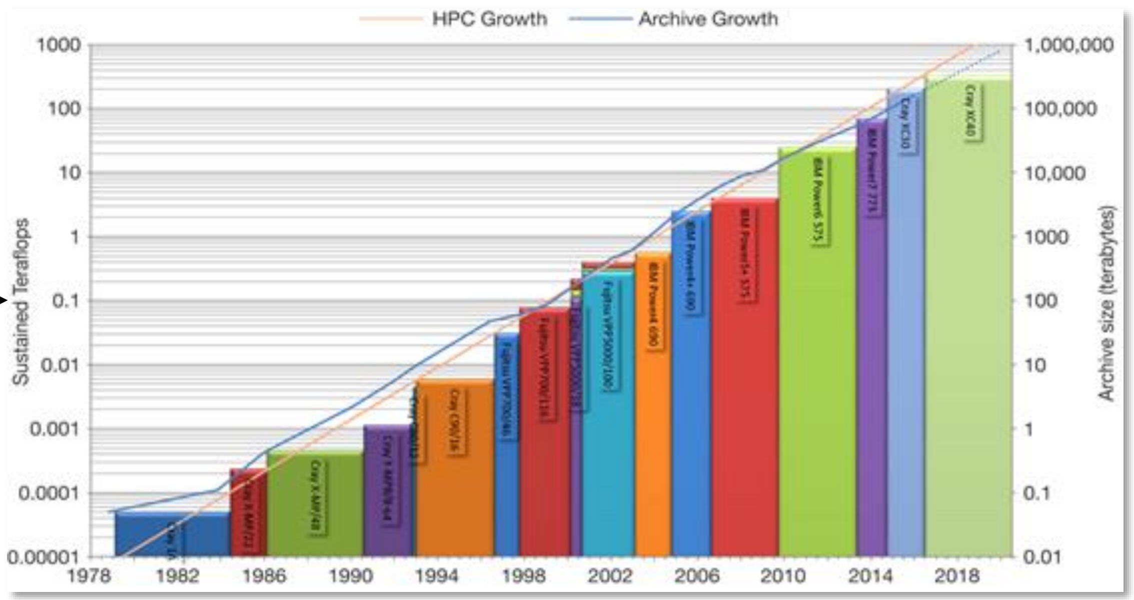


The concept:
Take input and output samples from a large data set
Learn to predict outputs from inputs
Predict the output for unseen inputs

The key:
Neural networks can learn a complex task as a “black box”
No previous knowledge about the system is required
More data will allow for better networks

Machine Learning – Why now?

Increase in data volume



New computing hardware

New machine learning software

TensorFlow

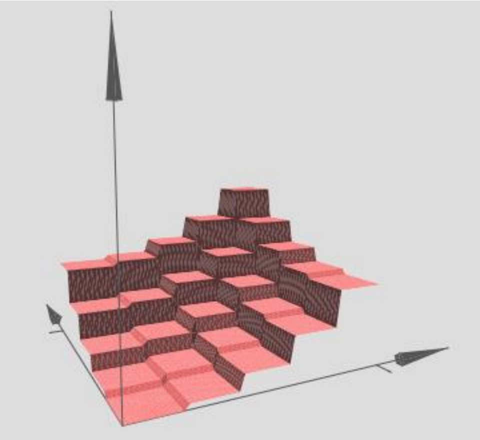
Keras

PyTorch

Increase in knowledge

Neural networks – the good, the bad, and the ugly

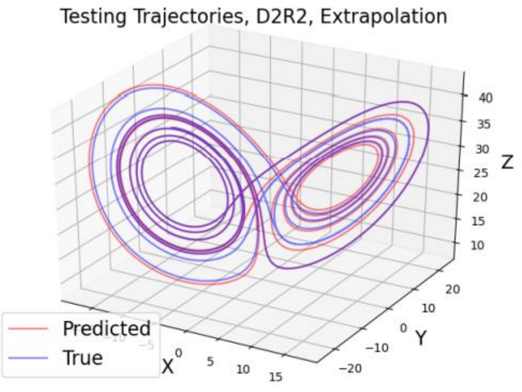
The good news: Neural networks are universal. No matter what function we want to learn, there is a neural network that can do the job if enough data is available and if the neural network is complex enough. See the “universal approximation theorem”.



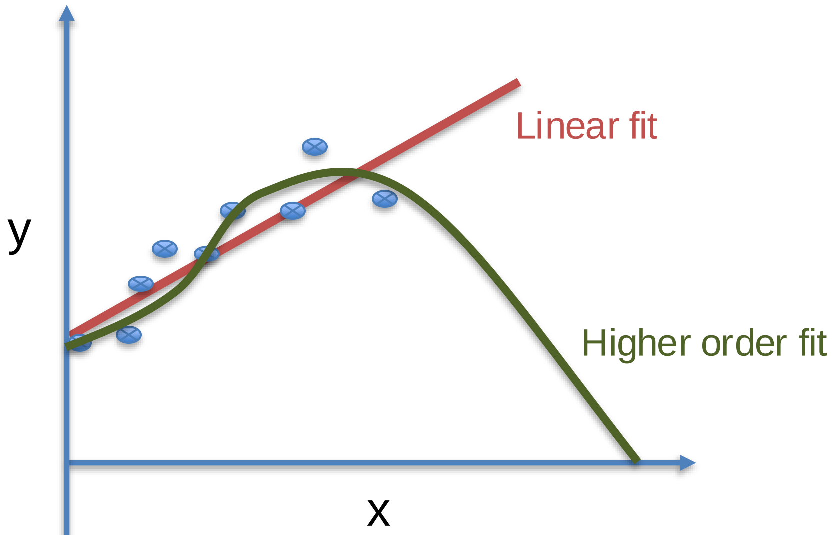
Source:http://neuralnetworksanddeeplearning.com/chap4.html

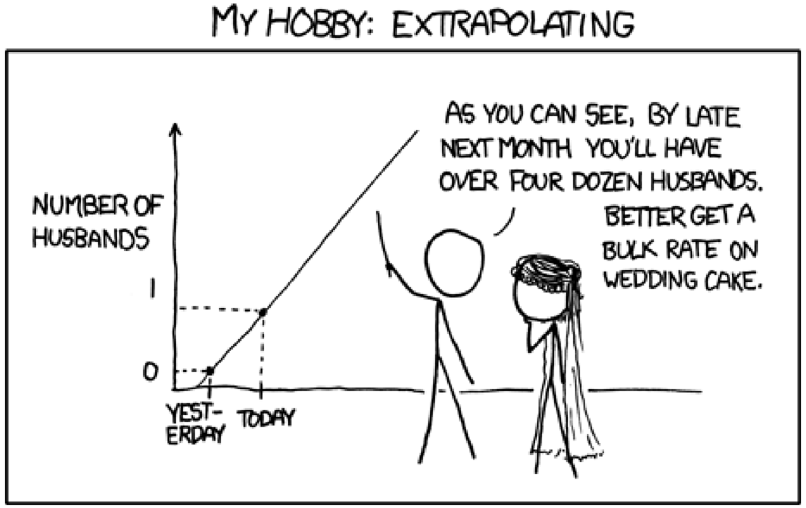
The bad news:
However, neural networks are often not the best tool for the job.
(Take the example to learn an equation.)
They are also very difficult to interpret.

$$\frac{dx}{dt} = \sigma(y - x)$$
$$\frac{dy}{dt} = x(\rho - z) - y$$
$$\frac{dz}{dt} = xy - \beta z$$



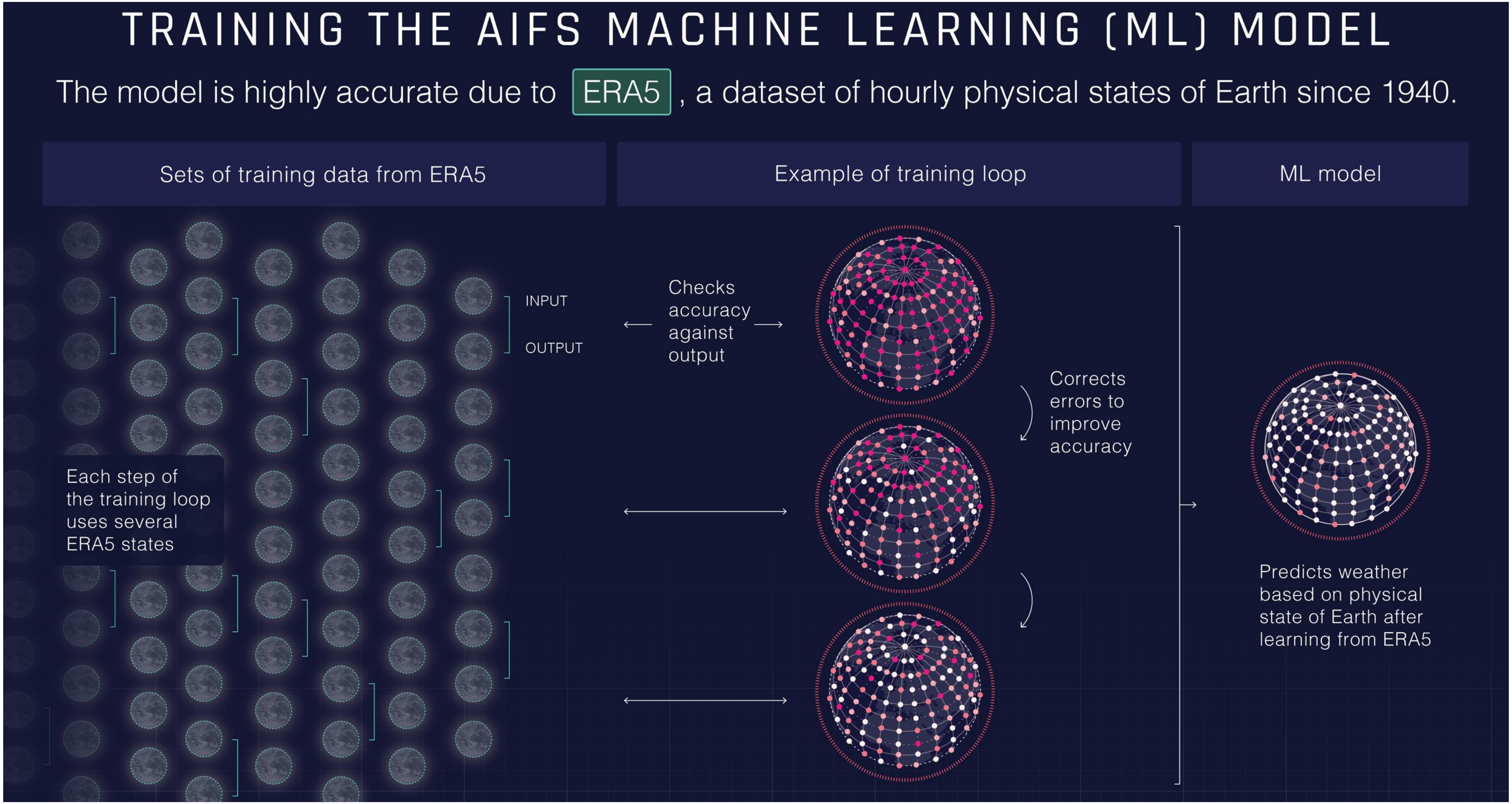
The ugly news: Neural networks are horrible when they are asked to extrapolate.





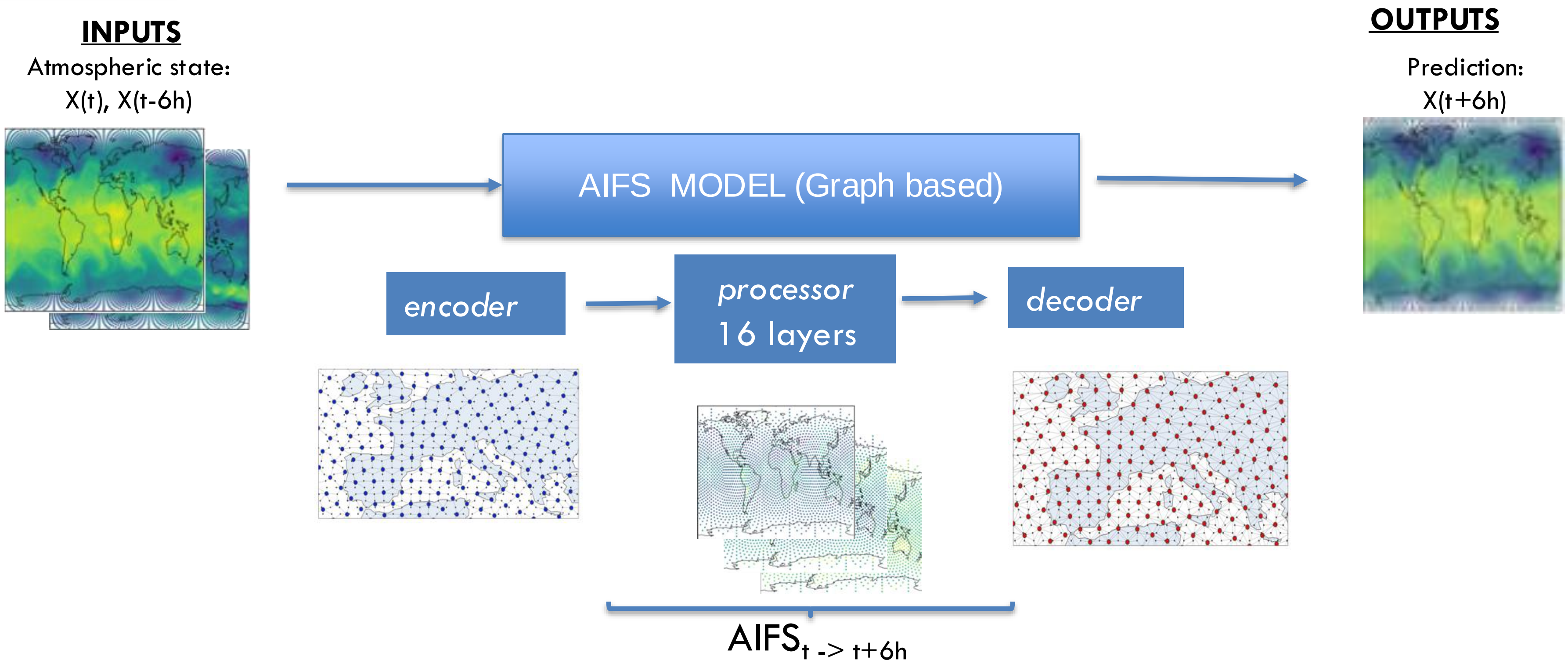
Source: https://xkcd.com/605/

How to build a machine learned model with knowledge about the physical system?



Artificial Intelligence Forecasting System

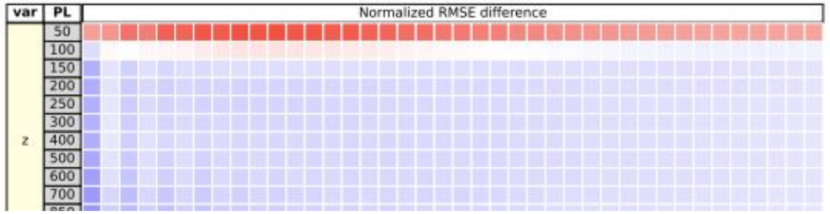
TRAINING

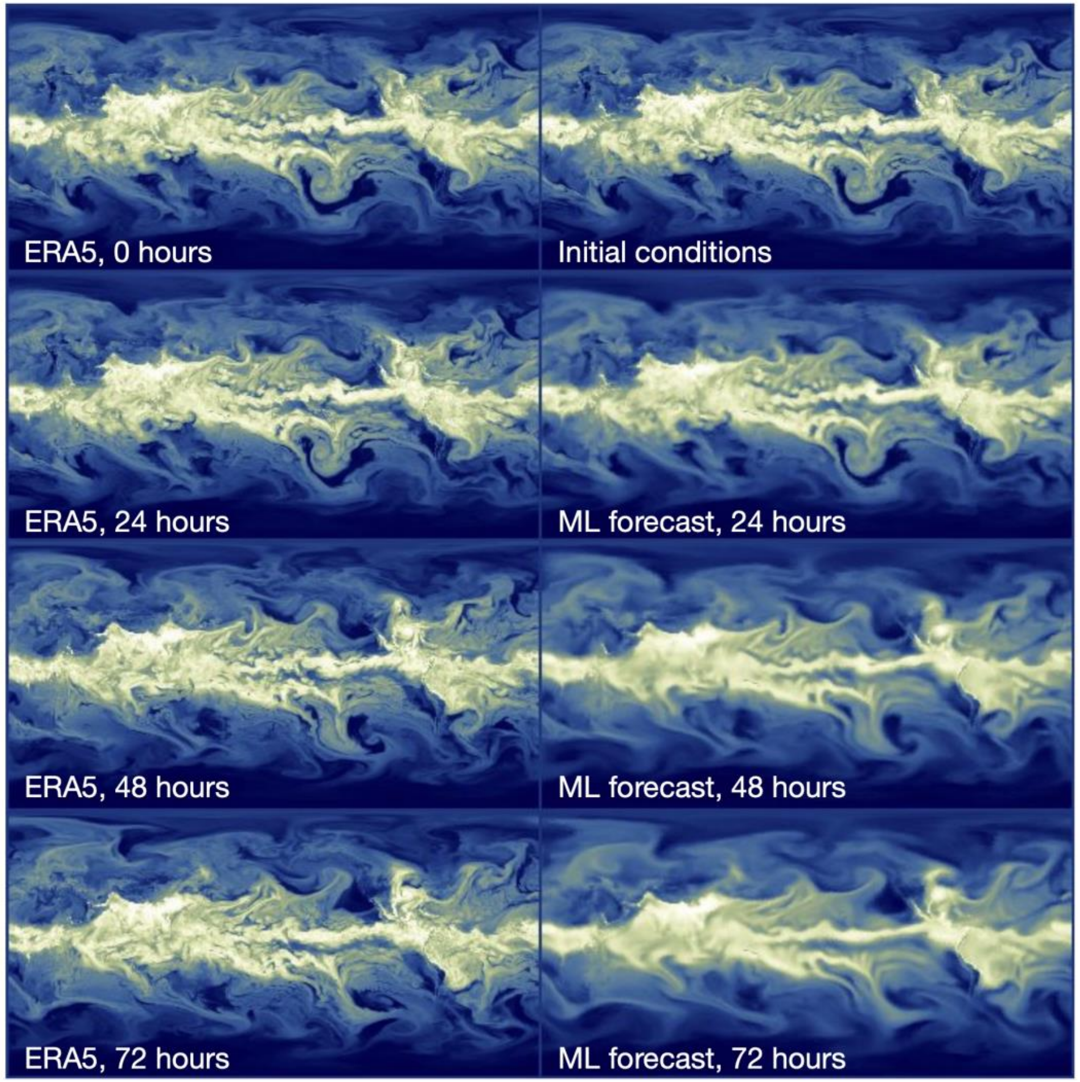


Lots of neural network architectures successful.
All share weights across space to some extent.

Slide from Mat Chantry

2022-today: The machine learning revolution





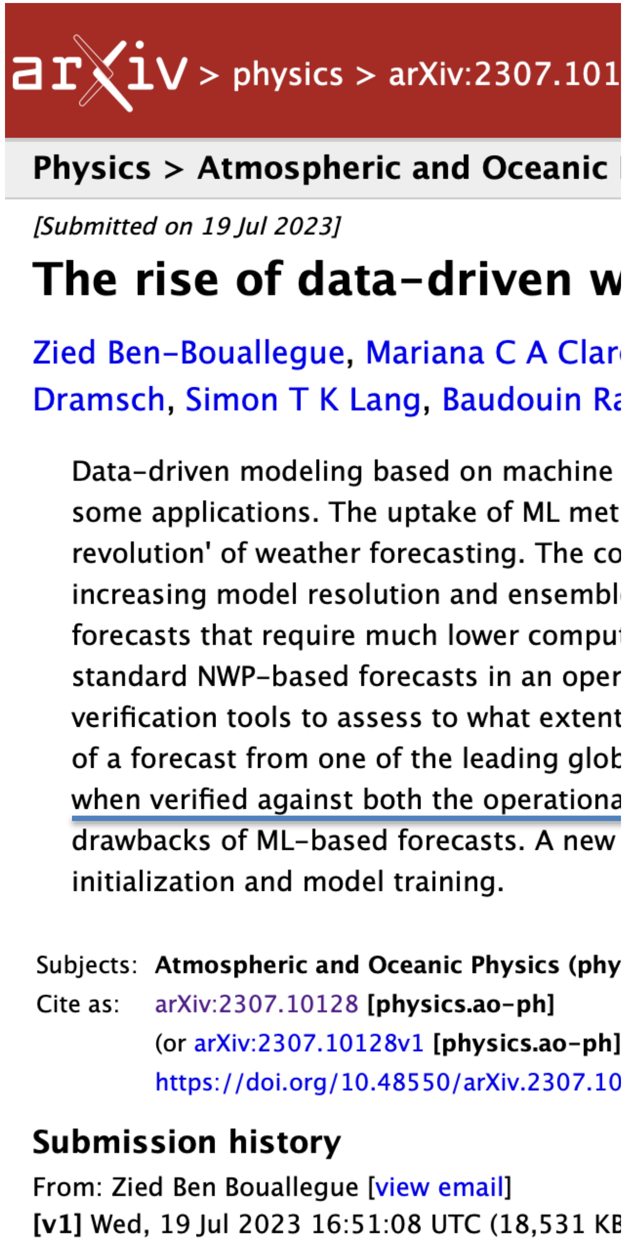
In 2022 machine learned forecast models from Google, NVIDIA and Huawei are beating conventional weather forecast models in deterministic scores and are orders of magnitude faster during inference.

But how do these models actually work?

In 2023 we still had many questions:
Can they avoid the smearing out for long predictions?
Can they learn uncertainty?
Can they extrapolate and faithfully represent extreme events?
Can they represent physically consistent forecasts?
Can they do data assimilation?

Images from Keisler (2022)

2022-today: The machine learning revolution



arXiv > physics > arXiv:2307.101

Physics > Atmospheric and Oceanic

[Submitted on 19 Jul 2023]

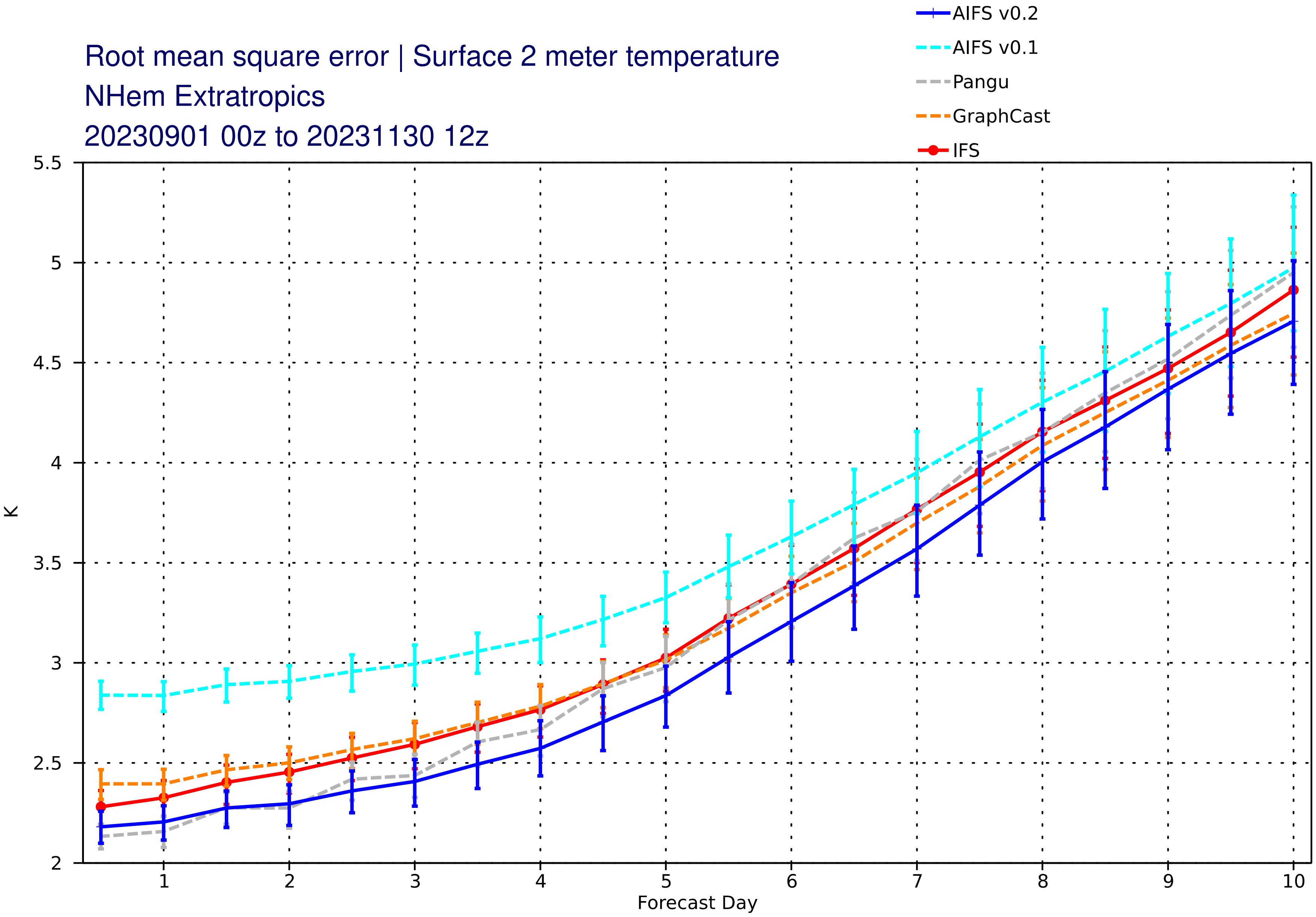
The rise of data-driven w

Zied Ben-Bouallegue, Mariana C A Clar
Dramsch, Simon T K Lang, Baudouin Ra

Data-driven modeling based on machine
some applications. The uptake of ML met
revolution' of weather forecasting. The co
increasing model resolution and ensembl
forecasts that require much lower comput
standard NWP-based forecasts in an oper
verification tools to assess to what extent
of a forecast from one of the leading glob
when verified against both the operationa
drawbacks of ML-based forecasts. A new
initialization and model training.

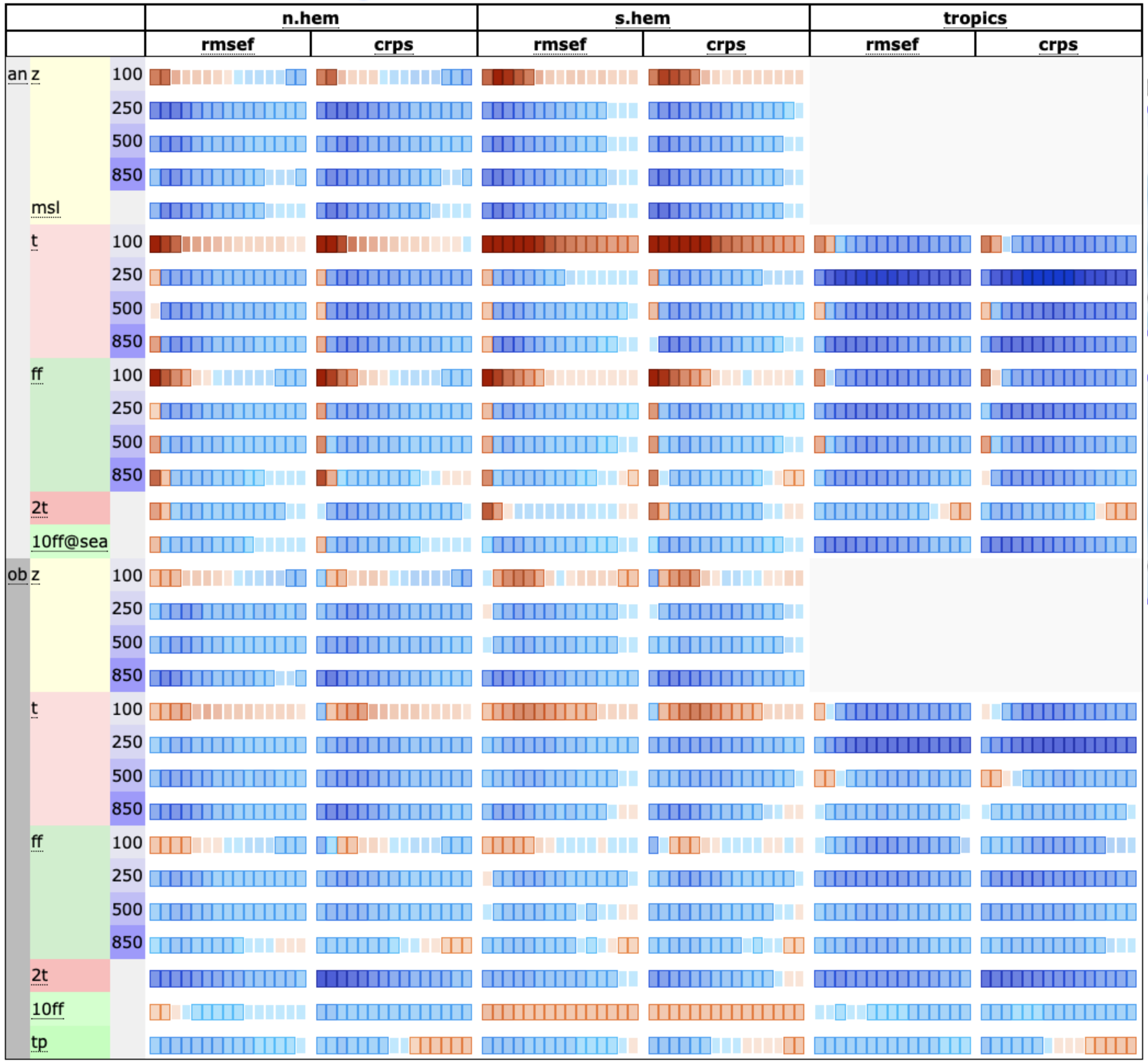
Subjects: **Atmospheric and Oceanic Physics (phy**
Cite as: arXiv:2307.10128 **[physics.ao-ph]**
(or arXiv:2307.10128v1 **[physics.ao-ph]**
https://doi.org/10.48550/arXiv.2307.10

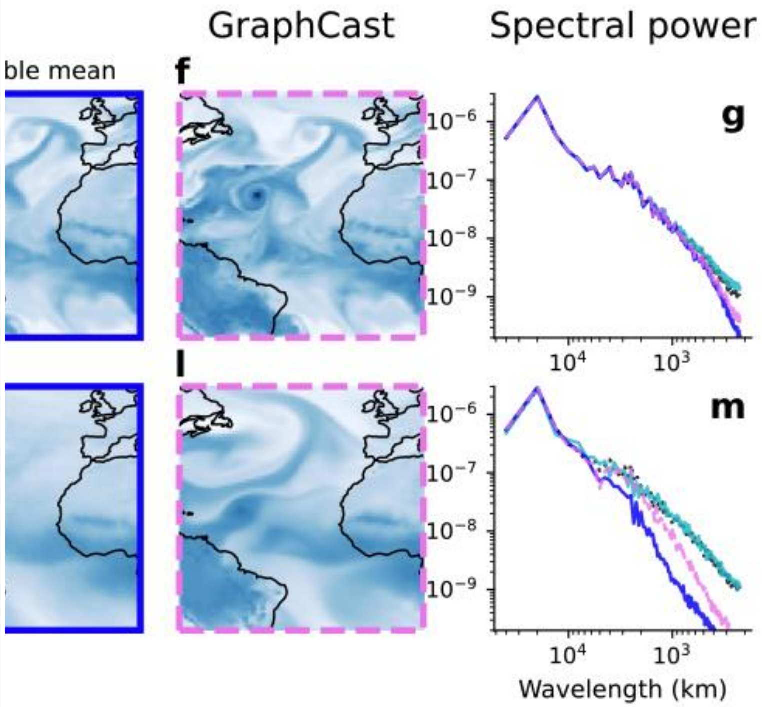
Submission history
From: Zied Ben Bouallegue [view email]
[v1] Wed, 19 Jul 2023 16:51:08 UTC (18,531 KE



Lang et al. ECMWF Newsletter 2024 and the AIFS team

Can machine learning models avoid the smearing out for long predictions?
Can machine learning models learn uncertainties?





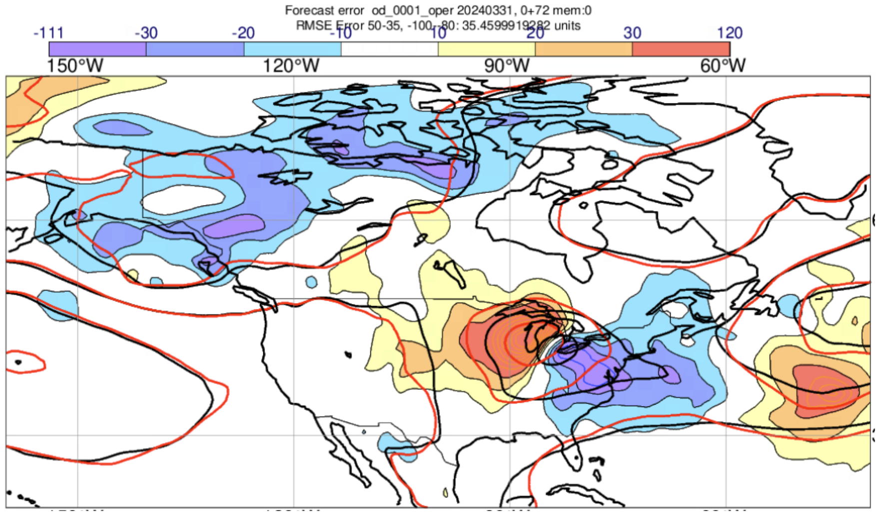
Yes.

We now also have ensemble AIFS. Lang et al. 2024 arxiv:2412.15832v1.

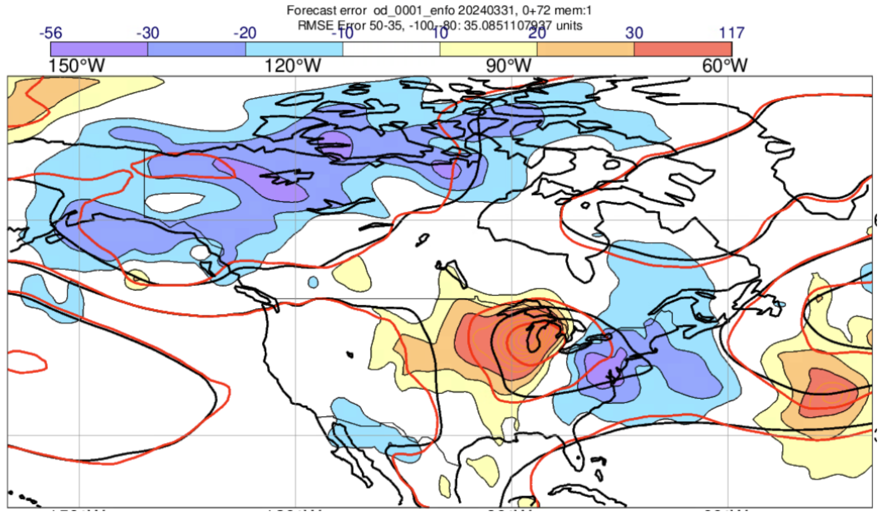
Can machine learning models represent extreme events?
Can machine learning models represent physical consistency?

Yes.

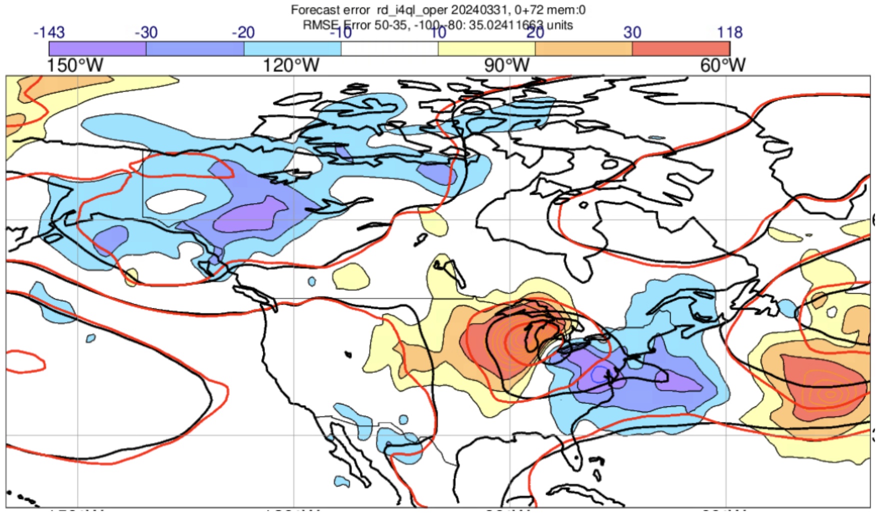
9km CTL



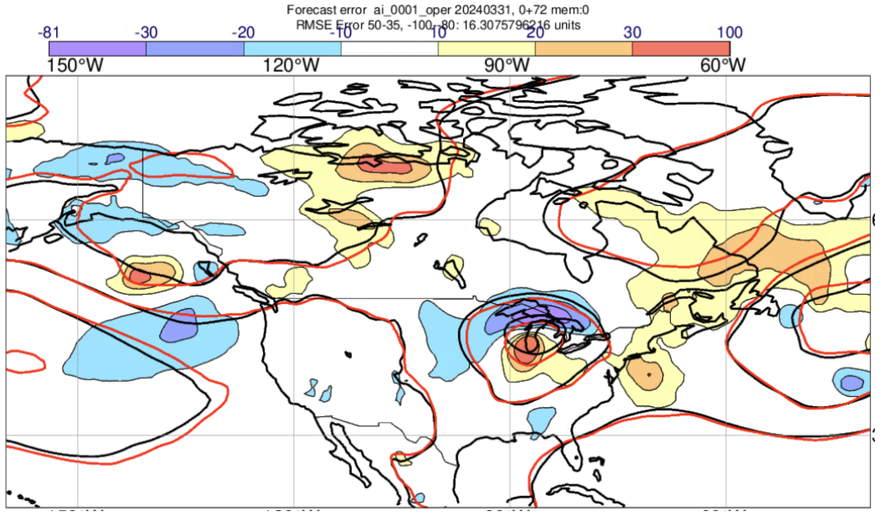
Ensemble mean



DestinE 4.4km



AIFS

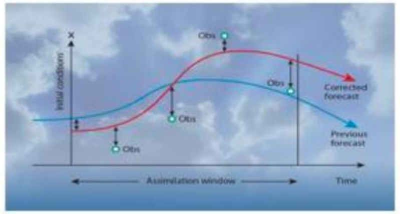


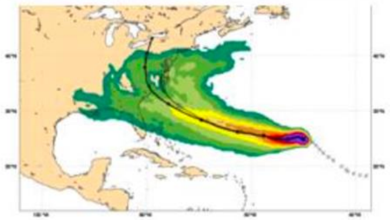
FC - black
AN – red
Error - shade

Z850 forecast error 2024-03-31 00UTC +72h

Can machine learning models do data assimilation?

Observations $O(10^8)$ data points → Data-assimilation $O(10^7)$ data points → Numerical weather forecasts $O(10^{10})$ data points → Post-processing of model output $O(10^{10})$ data points → Product generation $O(10^5)$ data points





Data assimilation is the process to blend information from observations and model simulations to find the optimal initial conditions.

Why would it not work?

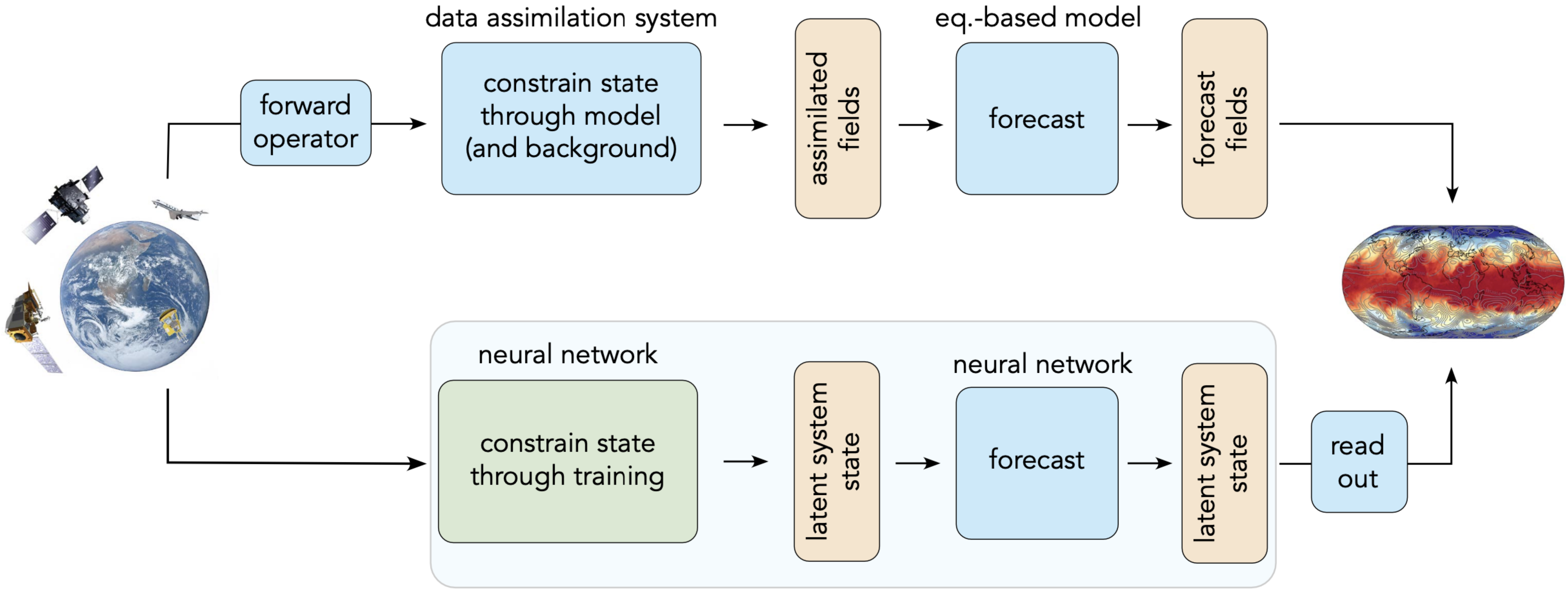
There are no easy training datasets for observations comparable to ERA5 or WeatherBench.
The further you go back in time, the less observations you have, and the more information needs to be filled.
There will be a huge null-space.
What do you do if satellites appear or disappear?

Why would it work?

All individual steps of conventional data assimilation can be replaced by machine learning.
Generative methods can fill in gaps, for example in down-scaling or ensemble simulations.
Machine learning is amazing.

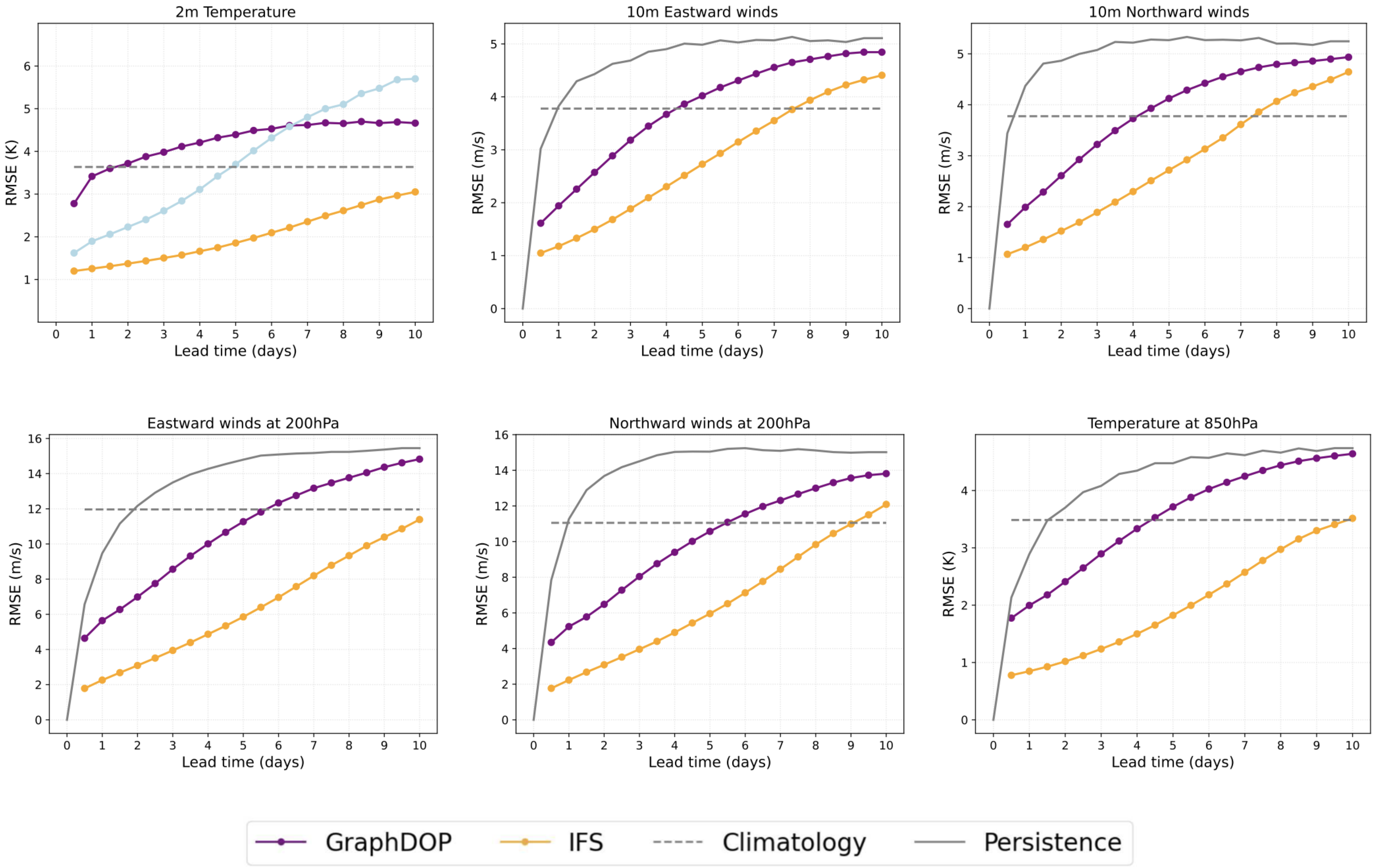
Can machine learning models do data assimilation?

Quote from Christian Lessig (ECMWF):
*We should not only try to **replace** data assimilation as we **may not need** data assimilation in the future.*



Christian Lessig

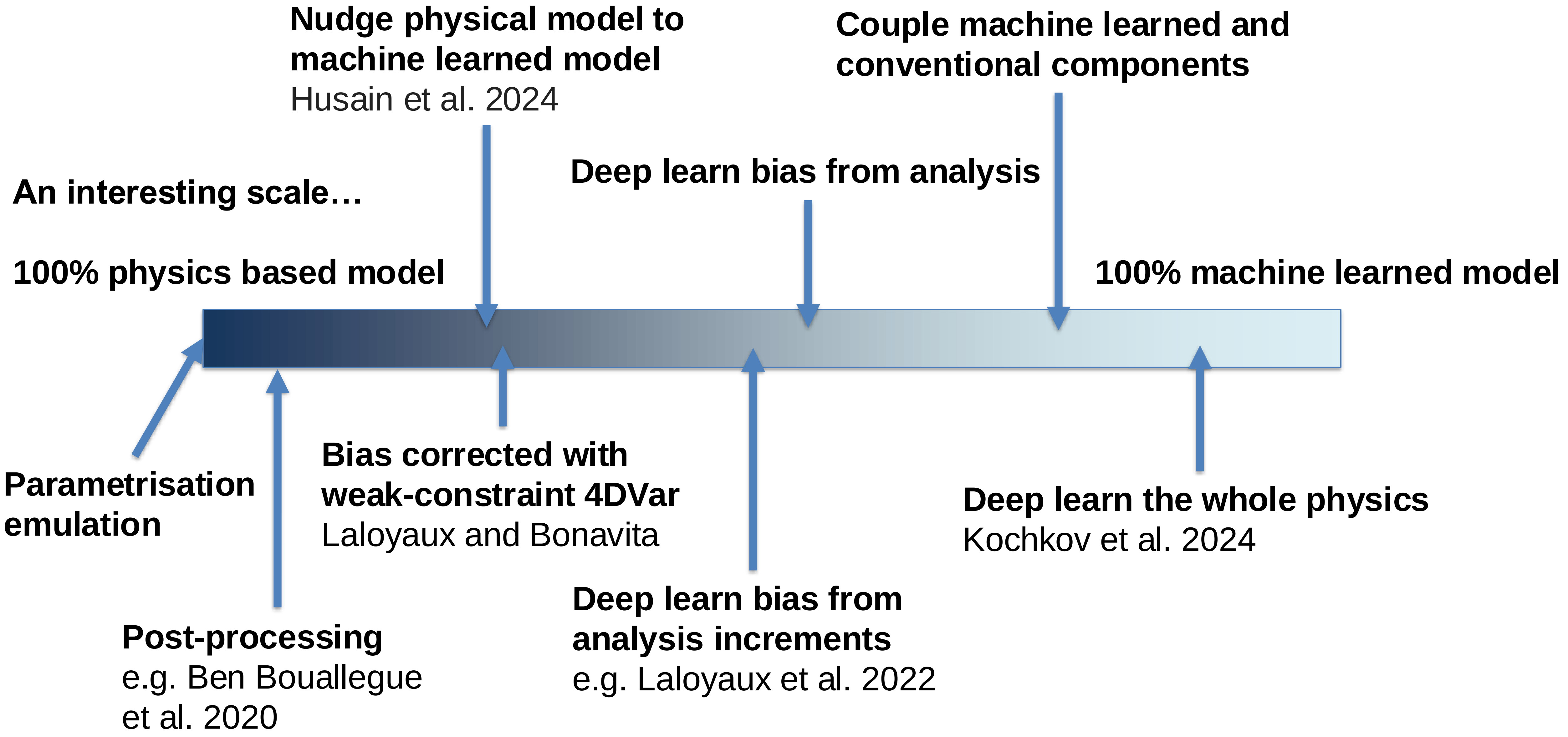
Can machine learning models do data assimilation?



Alexe et al., *arXiv:2412.15687*, 2024.

<u>Almost.</u>

2022-today: The machine learning revolution – What about hybrid?

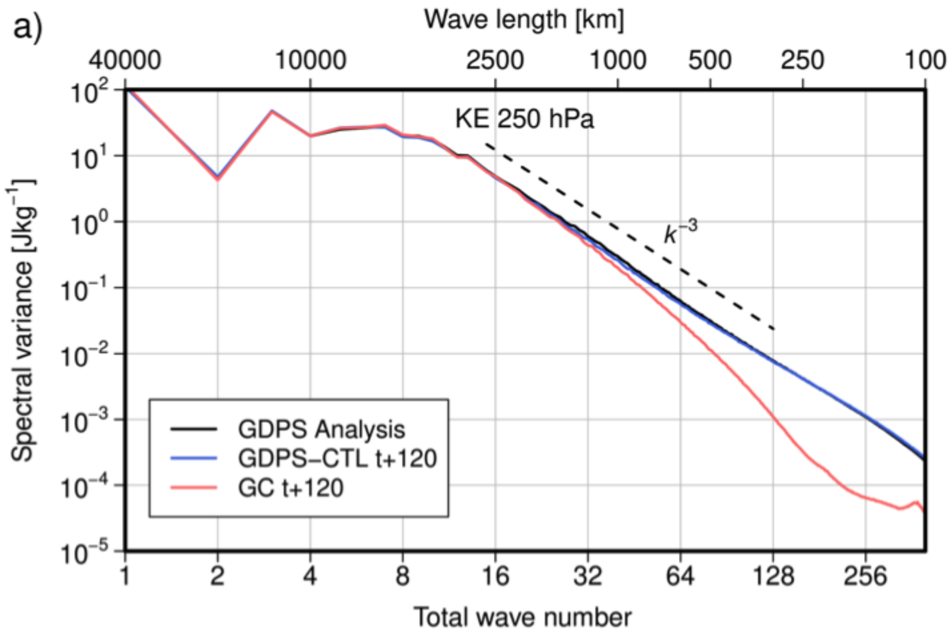


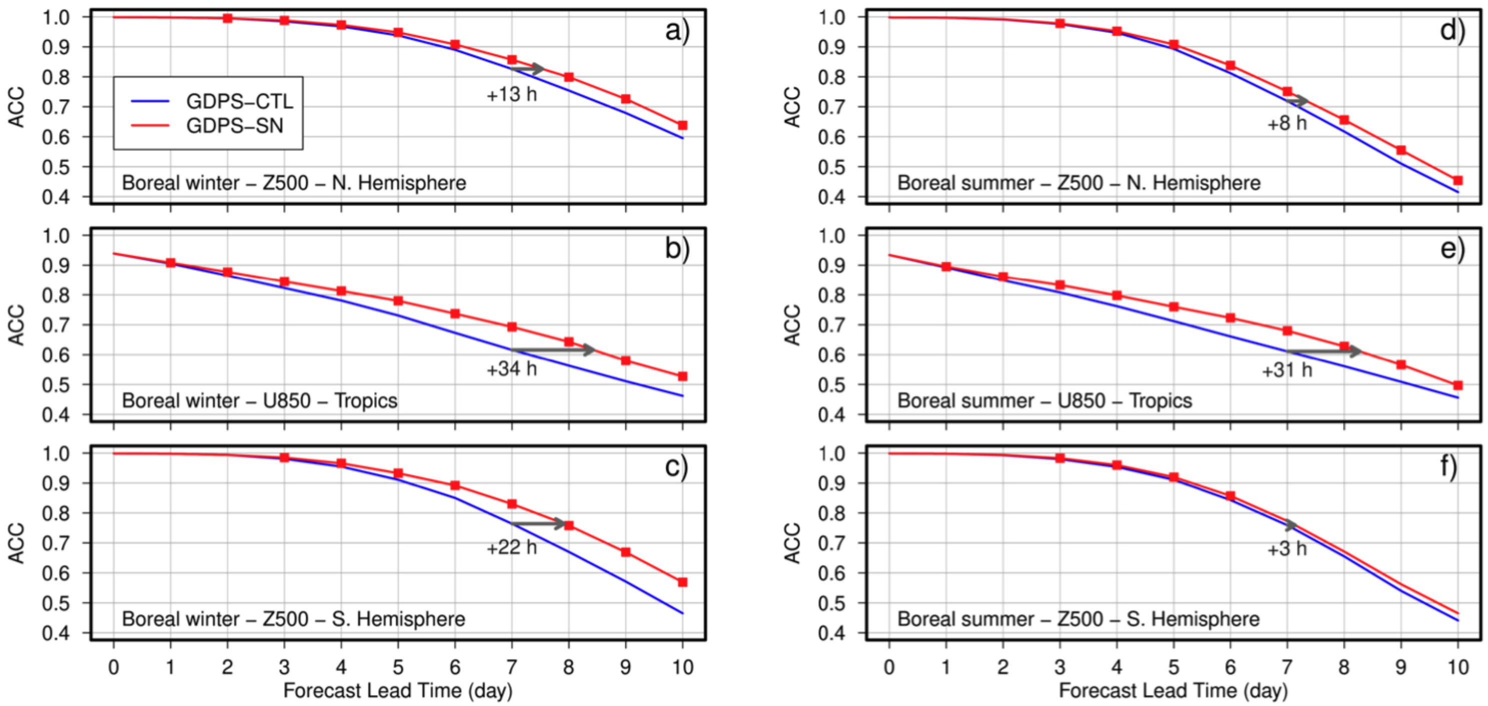
What is the best way to combine machine learning and physical models?

One of the general assumptions of the quiet revolution and physical modelling:
The large scales of the model simulations are well resolved and therefore correct.
The small scales of the model simulations are not well resolved and therefore incorrect.
→ Higher resolution leads to better predictions

However… Machine learned models are coarse, fail to represent small scales, and are still competitive.

→ Get best of both worlds by nudging large scales of machine learned models to the physical models.



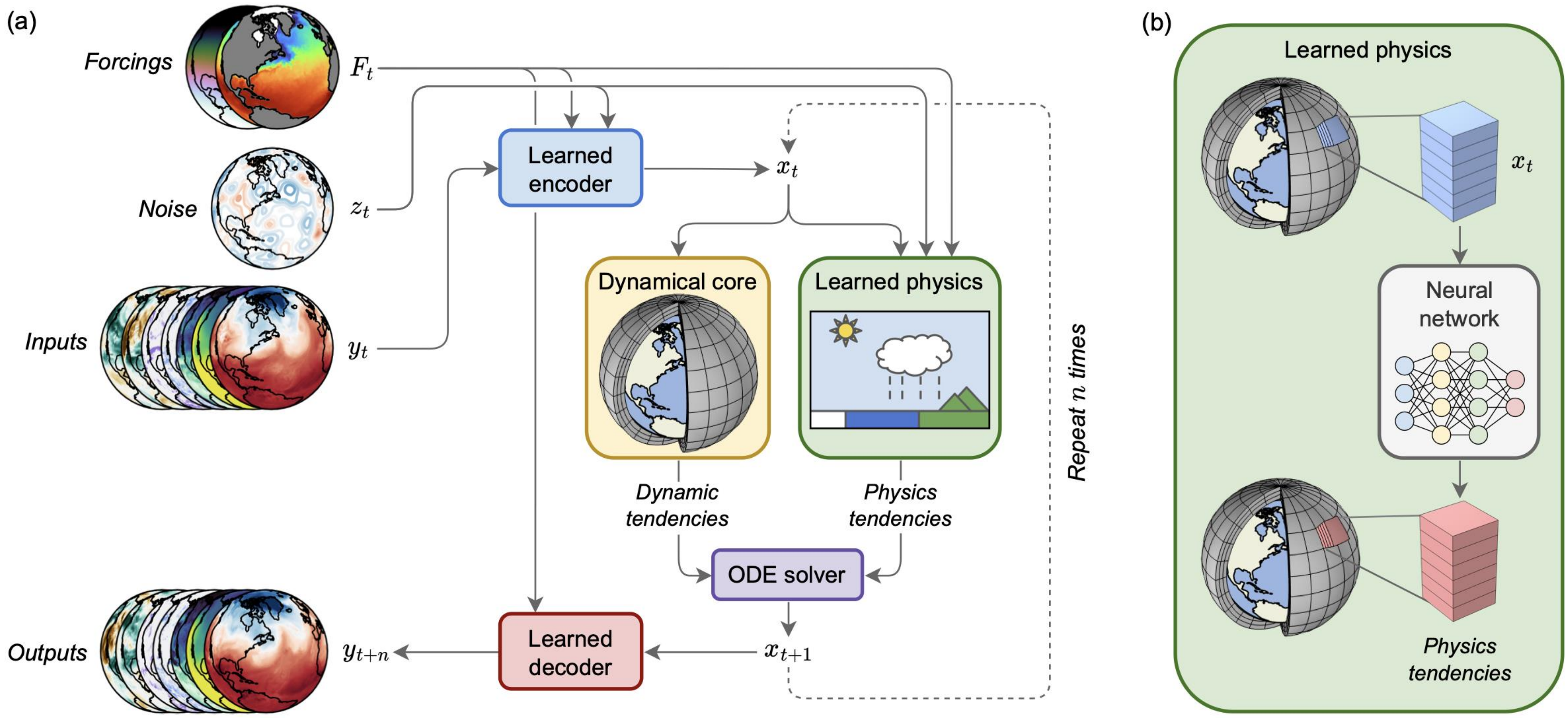


Husain et al. *arXiv:2407.06100* (2024).

<u>We checked, it is also working with IFS and AIFS.</u>

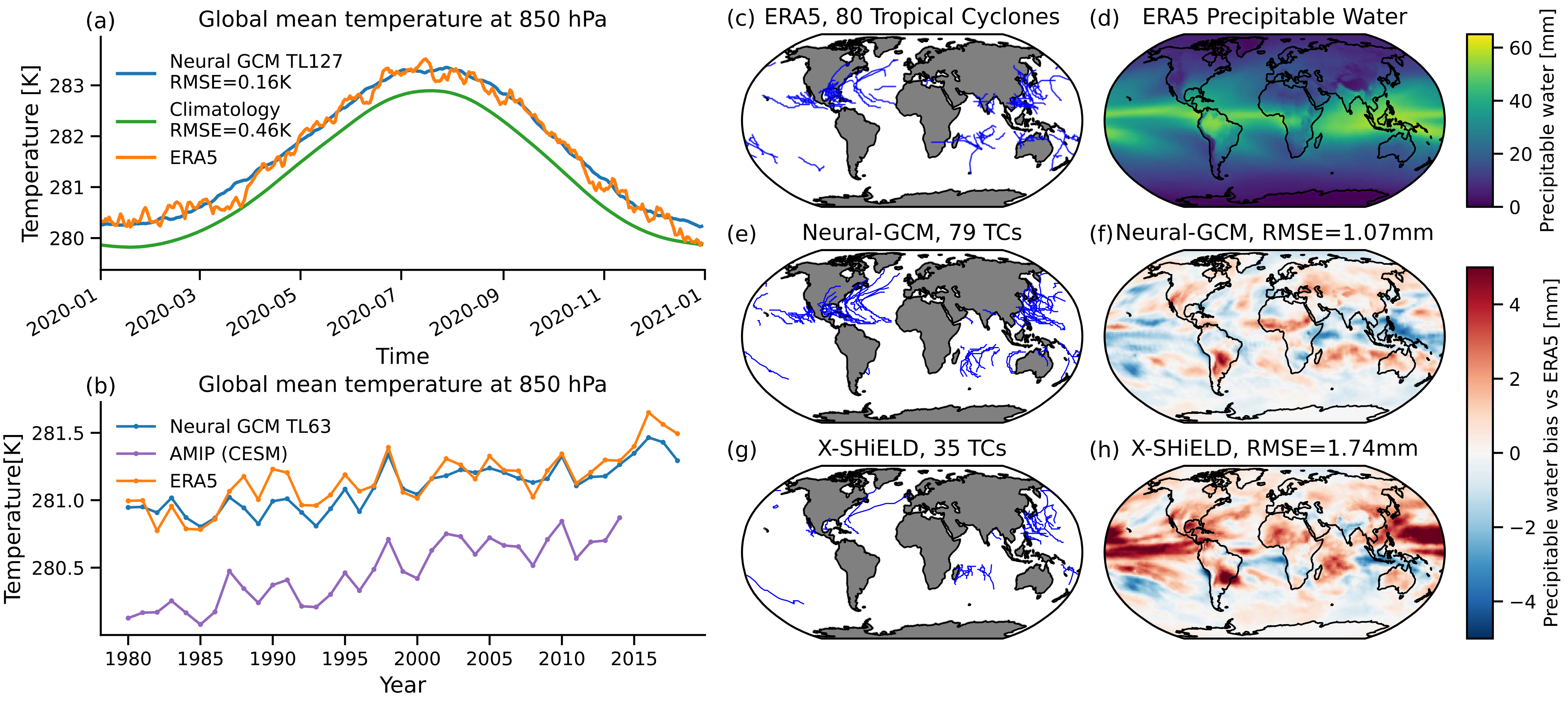
What is the best way to combine machine learning and physical models?

Or learn the small scales – see NeuralGCM from Google



Kochkov et al., *Nature* **632**, 1060–1066 (2024)

What is the best way to combine machine learning and physical models?



Machine learned models can now also do AMIP simulations.

Kochkov et al., *Nature* **632**, 1060–1066 (2024)

Next step: Machine learned climate simulations

Why would it not work?

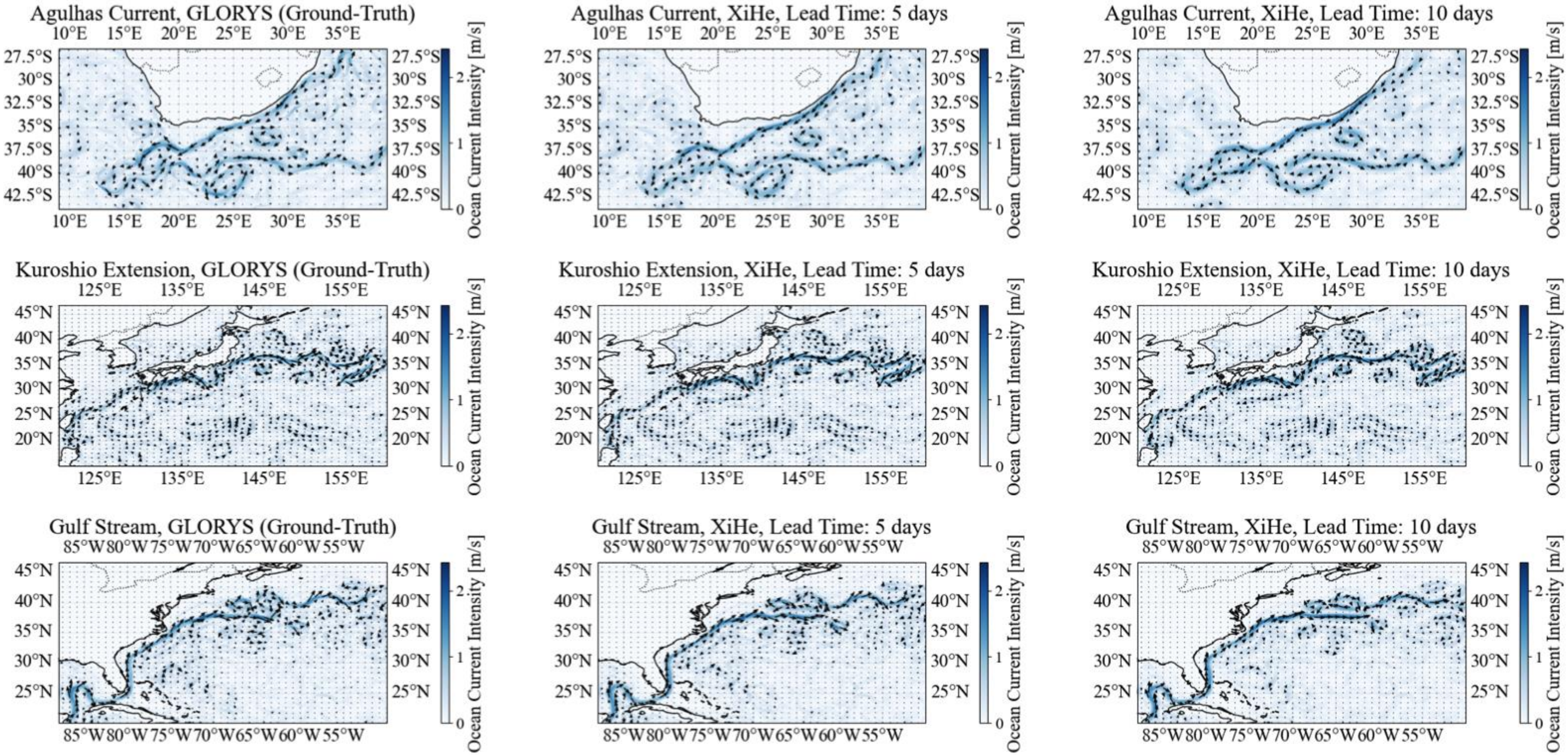
- Extrapolation.
- How do you represent CO_2?
- Need for more Earth system components including ocean, sea-ice, land, aerosols…
- We will not be able to represent the deep ocean.
- How could we trust the models for predictions that forecast the year 2100?

Why would it work?

- We can do AMIP simulations already.
- We can learn all model components needed.
- Who can represent the deep ocean correctly?
- Existing ML models are remarkably robust regarding unseen weather situations.
- The machine learned climate models will be bias free and beat conventional climate models in almost all comparable diagnostics for today's climate.

Next step: Machine learned Earth system models

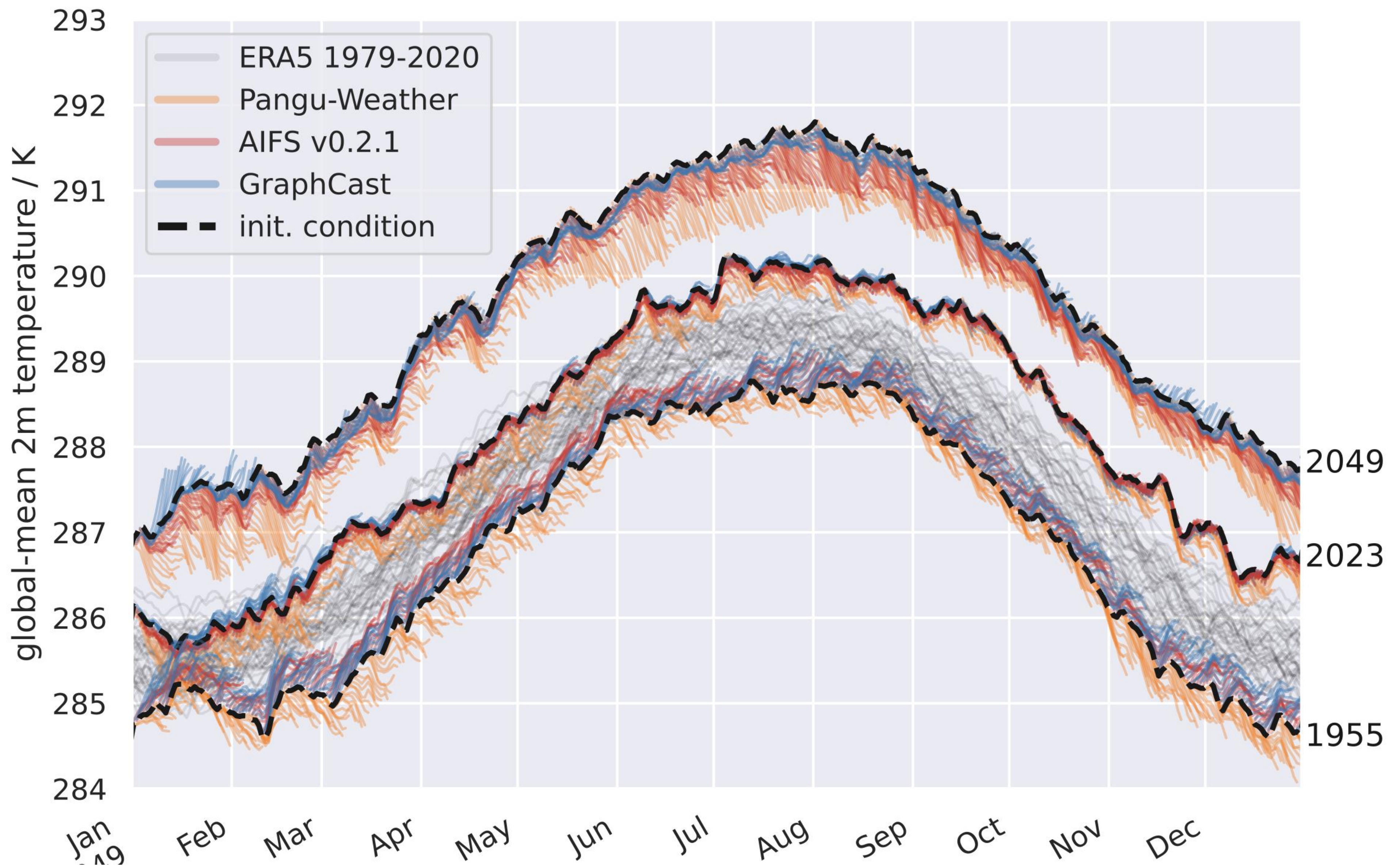
Land models should be easy. Ice models should be easy. The first ocean models exist already.



This is XiHe.
Wang et al. 2024.

How will we couple the models?

Next step: Machine learned climate simulations



Rackow et al.: arXiv:2409.18529 (2024)

How will physics-based and machine learning modelling develop?

Physics-based:
- Efforts in development will reduce
- They will be replaced by machine learning models in many applications
- They will still be needed – e.g. for grey-swan weather events, tipping points or training data generation
- Development will focus less on forecast scores and more on realism – e.g. km-scale
- Models need to be easy to use, portable and efficient or it will be impossible to find good staff

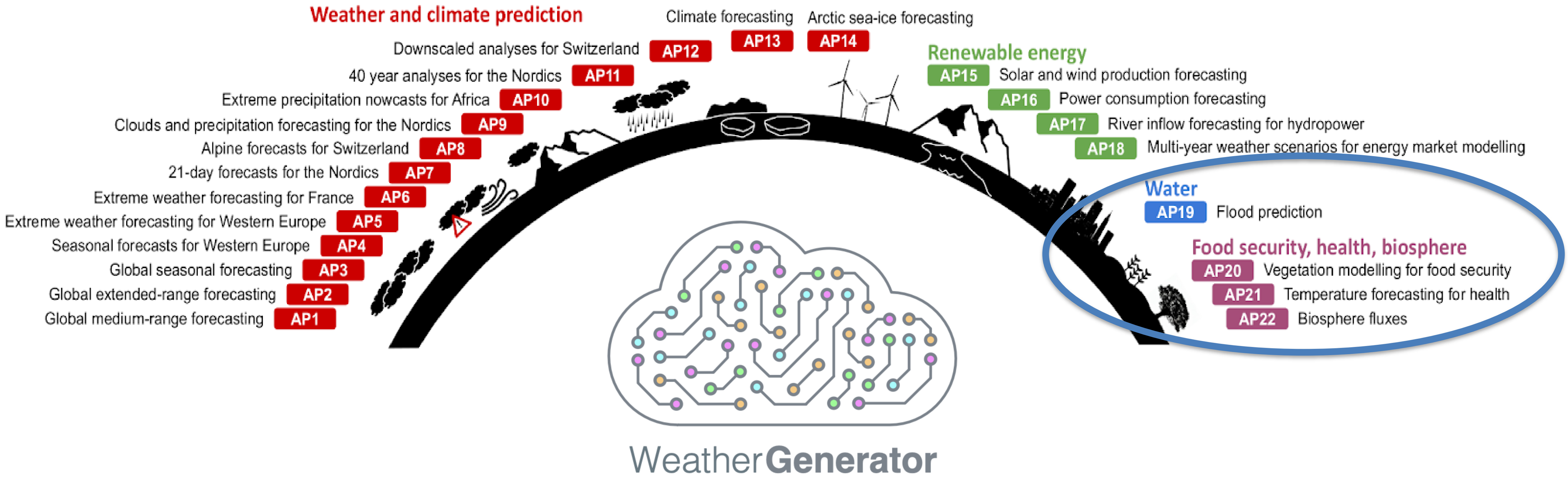
Machine-learned models:
- Efforts in development will increase
- There will be many different application areas that interact (but how many models eventually?)
- Datasets and models will increase in size (but what will be the limit?)

Modelling centres:
- Revise physics-based modelling – e.g. via tools such as GT4Py
- Embrace machine learning
- Embrace federated computing and federated data
- Invest in good visualisation to discover data efficiently
- Consider difference between operational HPC and scalable HPC (see Bauer 2024)

Europe is prepared for these steps, in particular due to initiatives such as DestinE.

Destination Earth | EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

Next steps: Foundation models for weather and climate



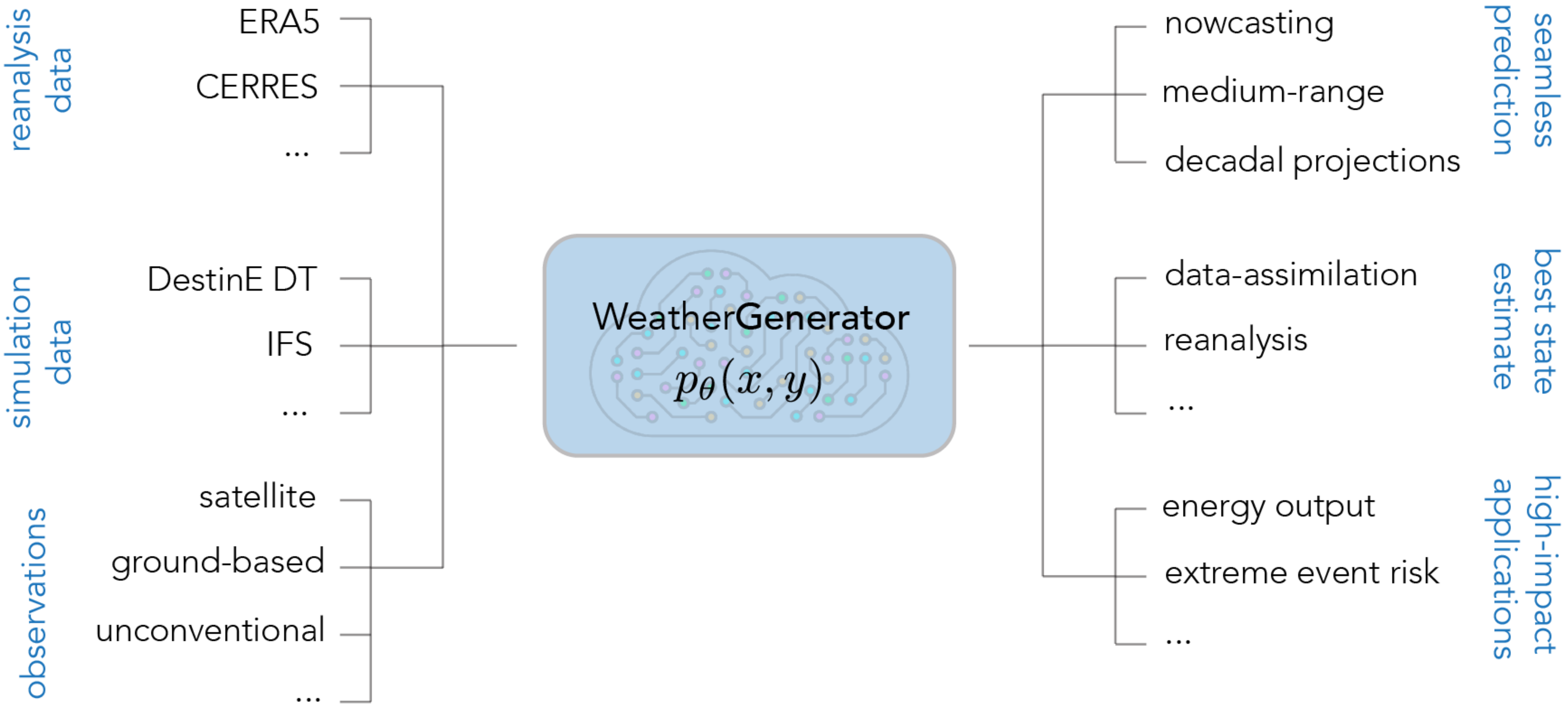
What about a unification of the machine learning applications via a Foundation Model for Earth system science?

There is already AtmoRep (Lessig et al. 2023) and other models such as ClimaX, Aurora and Orbit.

Aim: This project will build the machine-learned WeatherGenerator – the world's best generative Foundation Model of the Earth system – that will serve as a Digital Twin in Destination Earth (DestinE).

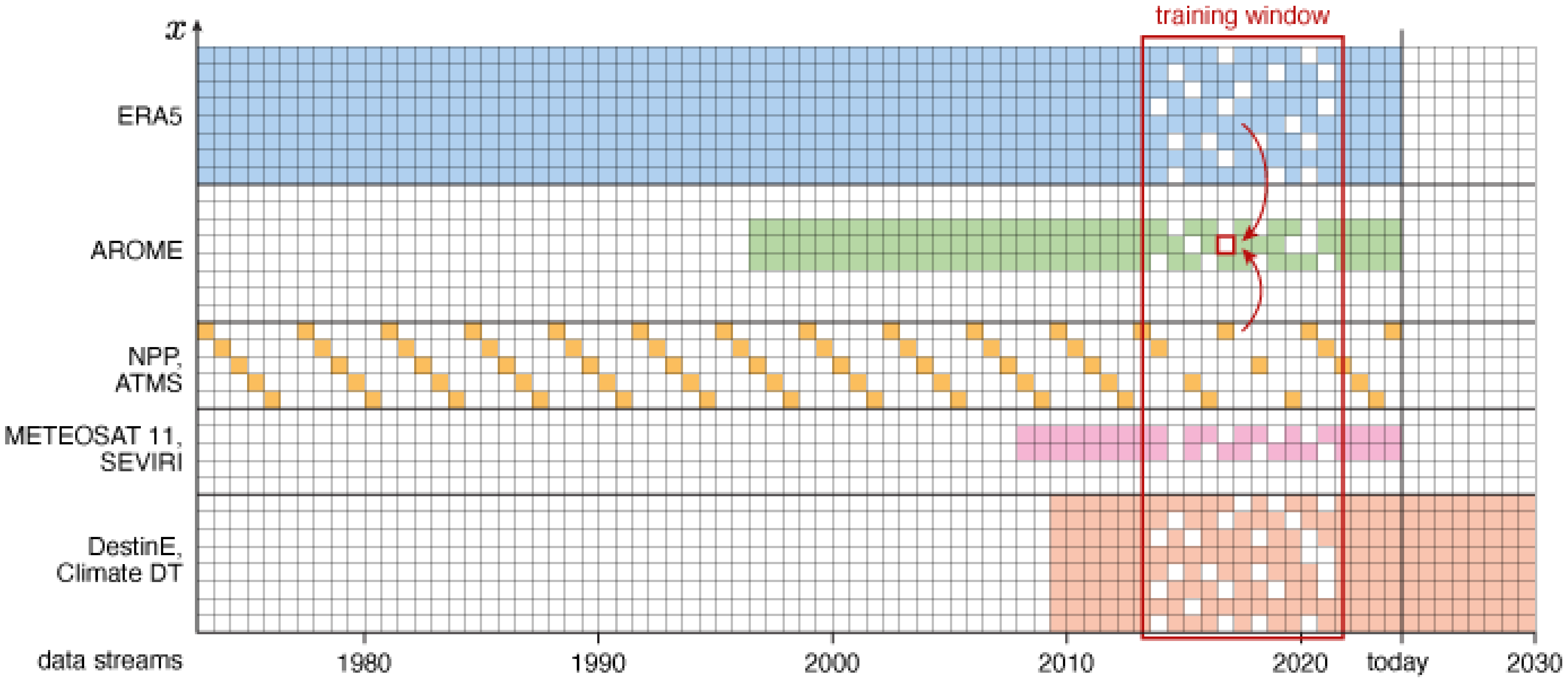
https://www.ecmwf.int/en/about/media-centre/news/2024/weathergenerator-project-aims-recast-machine-learning-earth-system

WeatherGenerator – A foundation model for weather and climate



https://www.ecmwf.int/en/about/media-centre/news/2024/weathergenerator-project-aims-recast-machine-learning-earth-system

WeatherGenerator – A foundation model for weather and climate



https://www.ecmwf.int/en/about/media-centre/news/2024/weathergenerator-project-aims-recast-machine-learning-earth-system

What is special about machine learning for land modelling?

Can we build machine learned land models?

- Yes, machine learning and parameter optimisation is nothing new for land modelling (e.g. Raoult et al. 2024).
- Yes, to learn a land model is comparably easy – small problem, zero dimensional.
- But it will be difficult to learn models that enable to represent long term memory and trends.

What are the limits for machine-learned land models?

- The pure model emulation is not too helpful as land models are comparably cheap.
- Learning is only possible if we have the data, and data is not very well distributed.
    - For which tasks is the data we have enough?
    - How can we apply learned models in data poor regions (in space and in future times)?
    - How do we know where and if we can generalise?
    - What are the surface fields that we actually need and can describe best?

What are the interesting science questions?

- How can we optimise land models to work in a coupled Earth system model?
- How do we know when we have reached limits of our machine-learned land models?
- How can we trust machine learned land models in climate simulations?
- When will the assumption of zero-dimensionality break down?

What have we learned?

Many thanks!

Peter.Dueben@ecmwf.int

The quiet revolution (1980-2020):
- Steady investment into Earth system modelling and Earth system observations made a difference.

The digital revolution (2015-today):
- Conventional models need to be made future proof via new software and hardware standards.
- Large projects such as DestinE make km-scale models possible today and will make a difference.

The machine learning revolution (2022-today):
- Models such as AIFS can beat physics-based models for deterministic and ensemble predictions.
- There is loads of interesting science to explore regarding hybrid models and predictability.
- We may soon see machine learning models that can do data assimilation and climate modelling.

The next step: Models will be better, tools will be easier, and data/HPC will be federated
- ECMWF will build a machine-learned Earth system model
- WeatherGenerator will build a machine-learned foundation model for Earth system science.
- To achieve this needs programmes such as Destination Earth, Earth-2 and EVE.

WeatherGenerator

MAELSTROM

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

Funded by the European Union

The WeatherGenerator project is funded by the European Union under grant agreement No 101187947. The MAELSTROM, ESiWACE and Hanami projects have received funding from the EuroHPC-Joint Undertaking under grant agreement No 955513, 101093054 and 101136269.