Disruptive Changes in Weather and Climate Modelling

Peter Dueben

Head of the Earth System Modelling Section

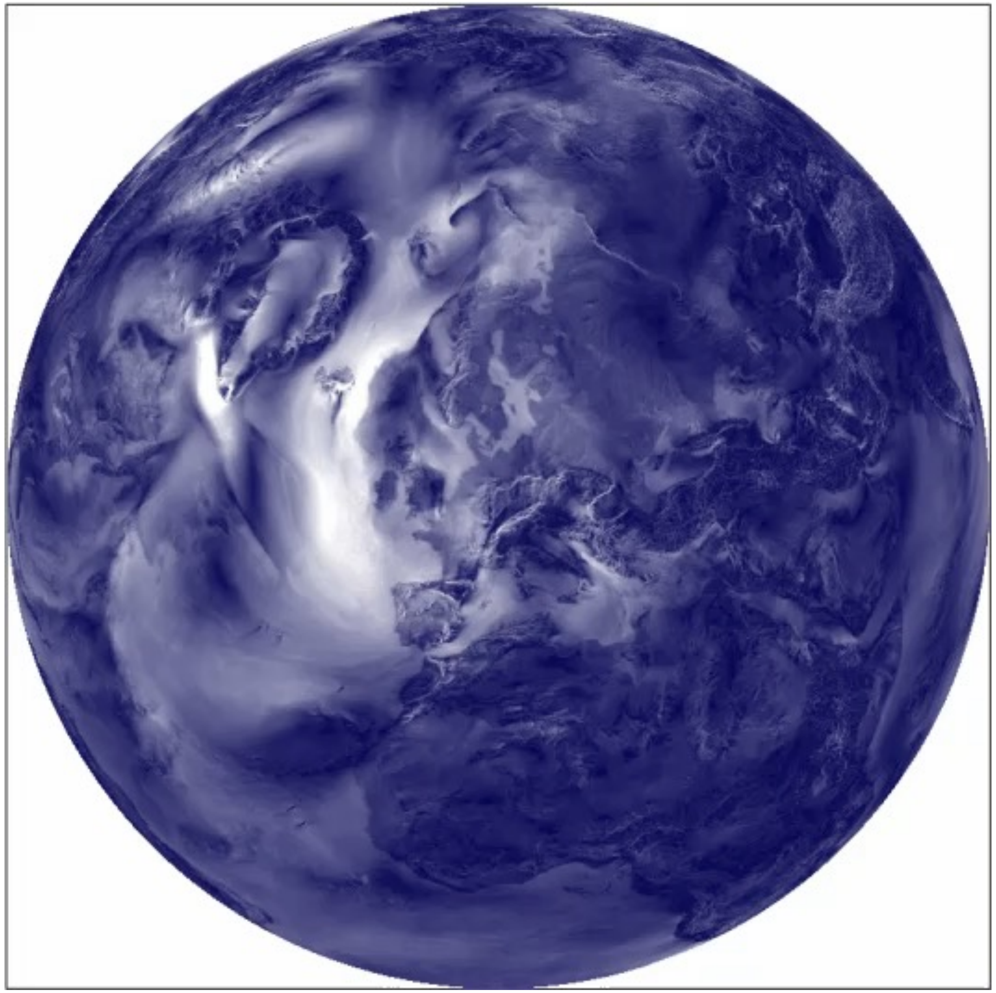
ECMWF | EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

The strength of a common goal

Funded by the European Union

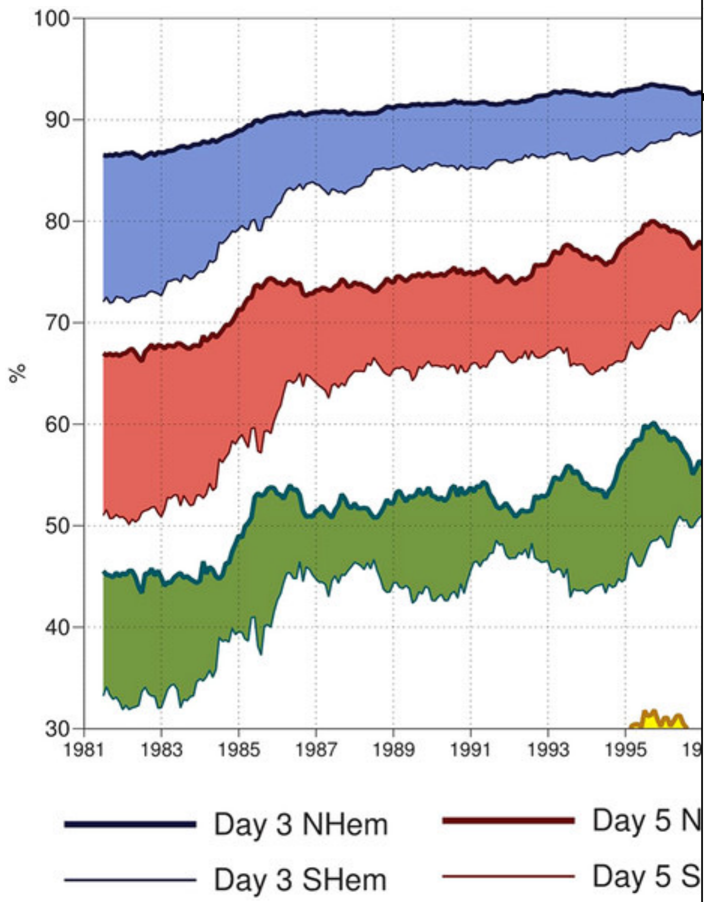
The MAELSTROM and ESiWACE projects have received funding from the EuroHPC-Joint Undertaking under grant agreement No 955513 and 101093054.

Numerical weather predications



Special thanks to Simon Lang

Half a century of success stories in numerical weather predictions



ECMWF

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.

And there is more to come with km-scale simulations

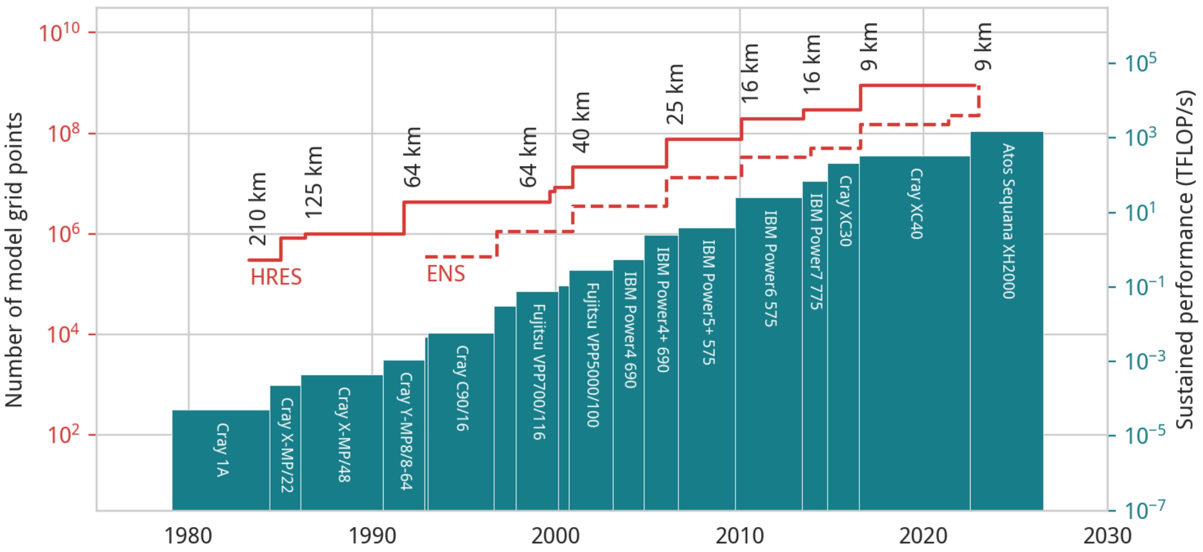
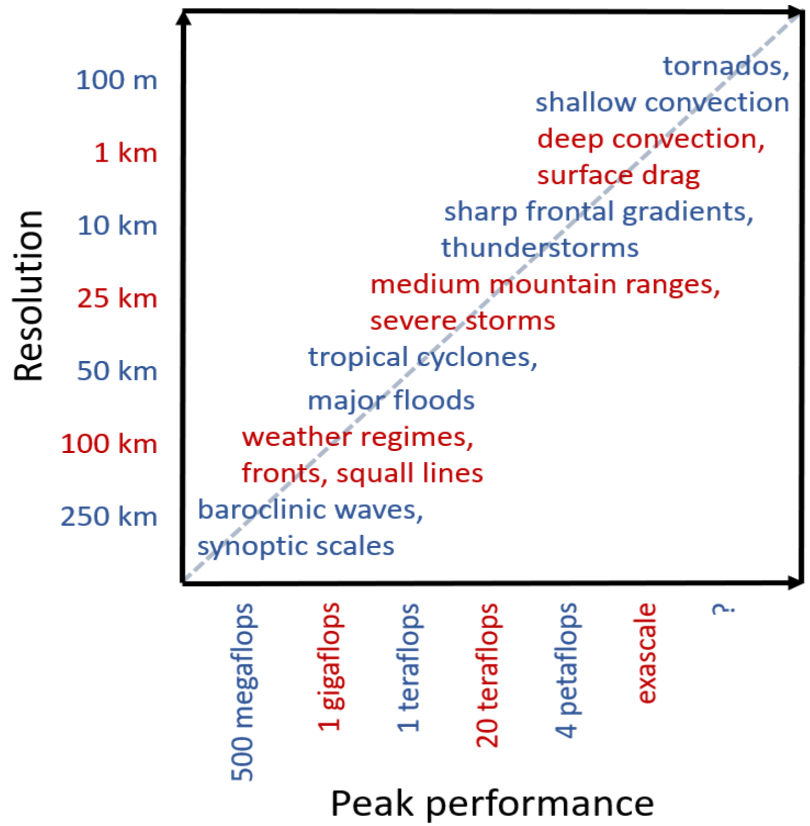
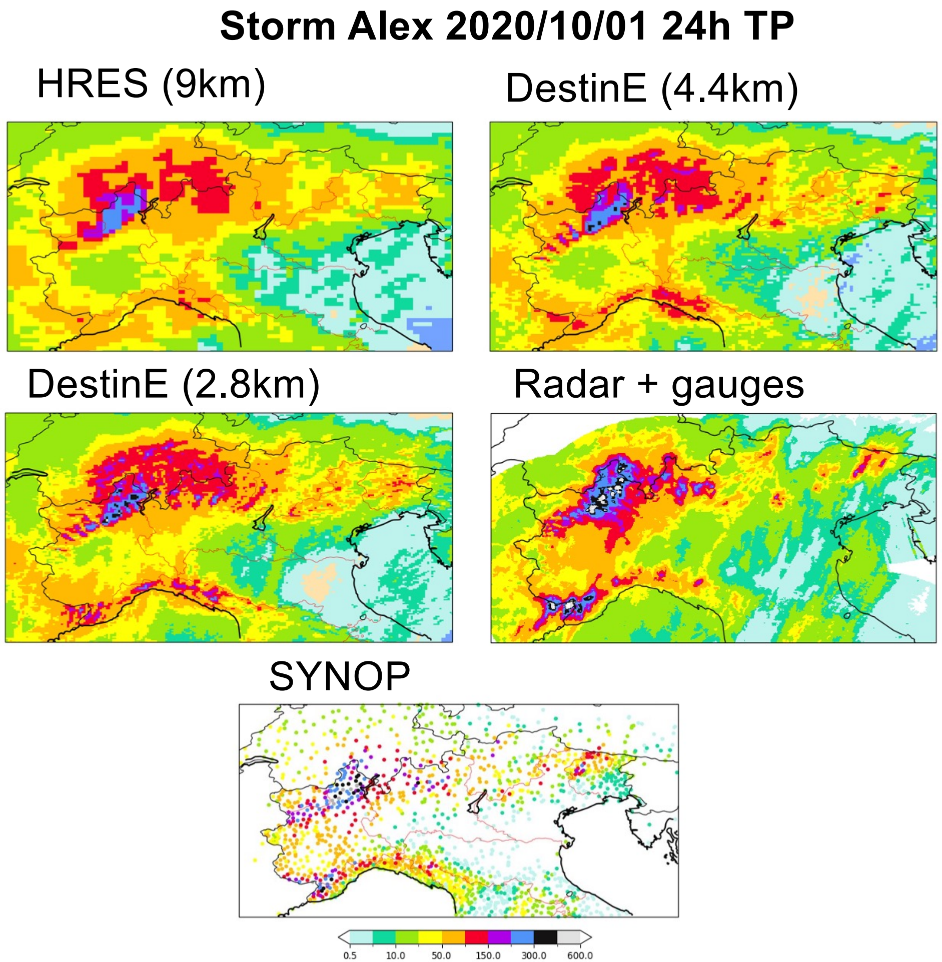


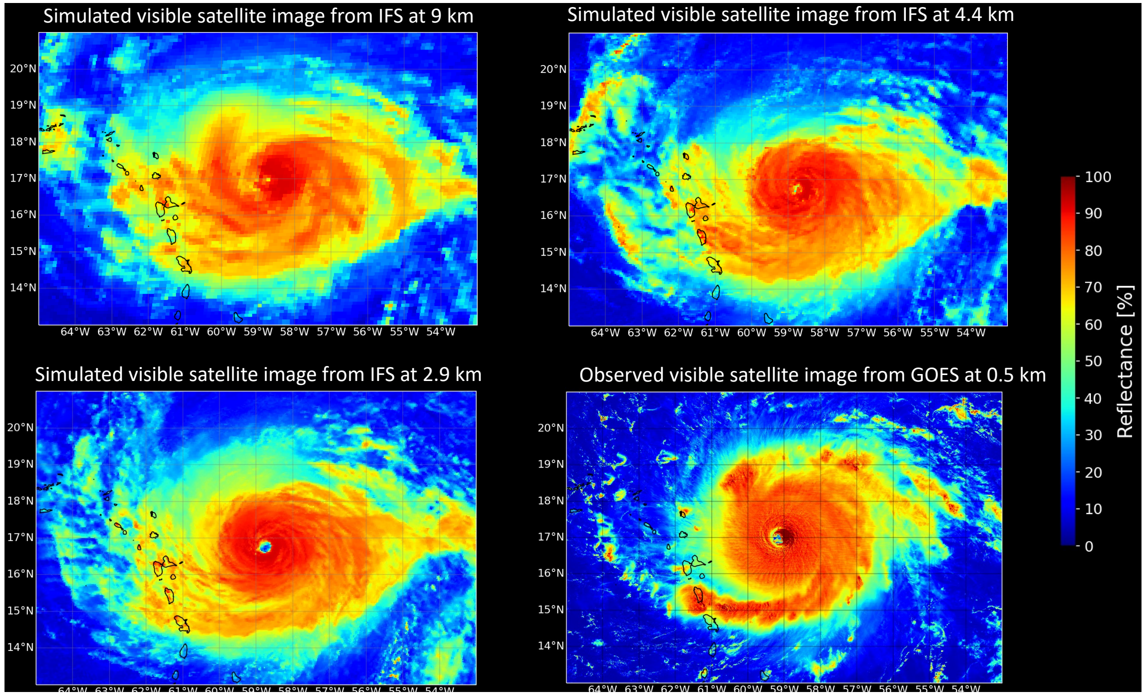
Figure courtesy of Sam Hatfield



Adapted from Neumann, Dueben et al. Phil Trans A 2018

And there is more to come with km-scale simulations





ECMWF

Destination Earth | EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

But progress in km-scale modelling is tough...

Compute power?
9 km → 1 km → Factor 9^3 = 729 compute power

Waiting for Moore's law.
→ 2^9 = 512 → Let's wait for 18 years?

Data and storage?
9km: 6,599,680 points x 137 levels x 10 variables
→ 9 billion points → > 0.5 TB

1.5km: 256,800,000 points x 137 levels x 10 variables
→ 352 billion points → > 20 TB

Uff…

ECMWF

TOP500 LIST - JUNE 2023

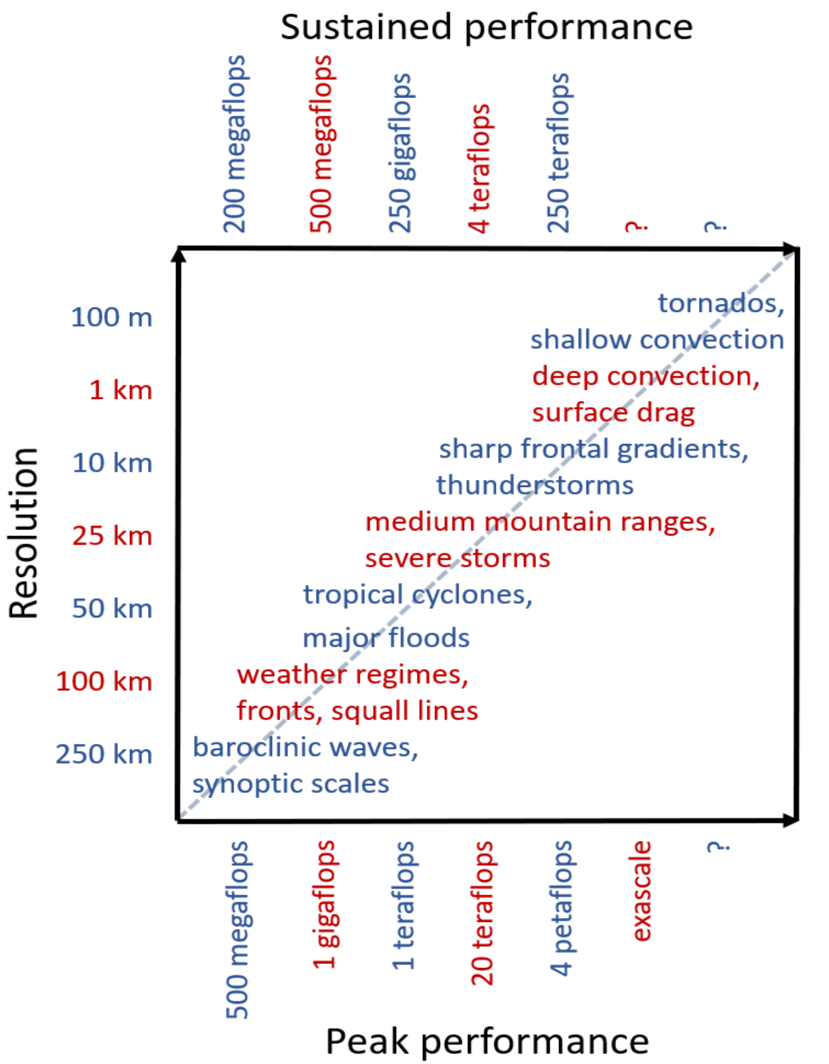
R_{max} and R_{peak} values are in PFlop/s. For more details about other fields, check the TOP500 description.

R_{peak} values are calculated using the advertised clock rate of the CPU. For the efficiency of the systems you should take into account the Turbo CPU clock rate where it applies.

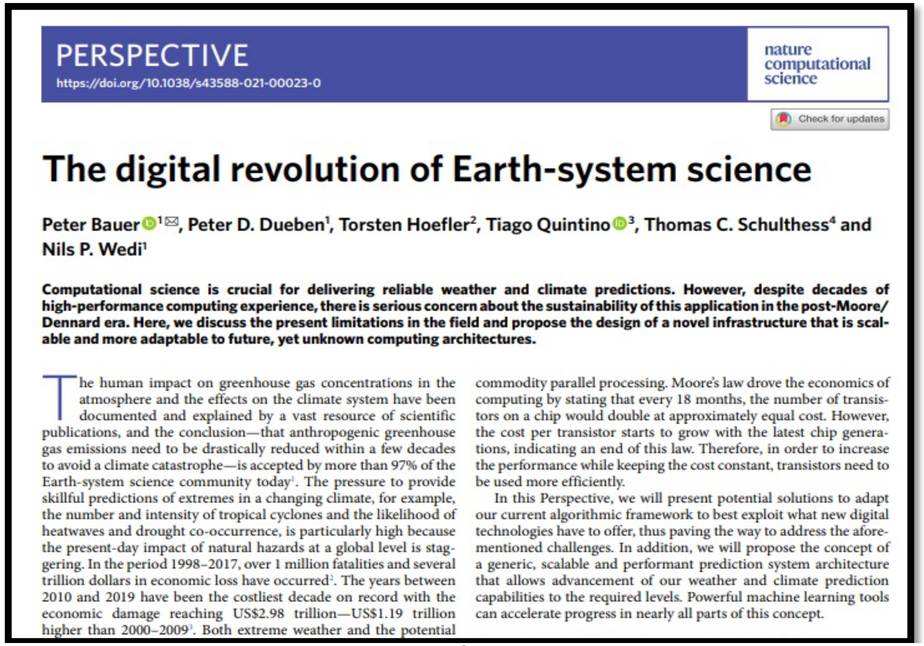
← 1-100 101-200 201-300 301-400 401-500 →

Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	**Frontier** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States	8,699,904	1,194.00	1,679.82	22,703
2	**Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442.01	537.21	29,899
3	**LUMI** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE EuroHPC/CSC Finland	2,220,288	309.10	428.70	6,016
4	**Leonardo** - BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64 GB, Quad-rail NVIDIA HDR100 Infiniband, Atos EuroHPC/CINECA Italy	1,824,768	238.70	304.47	7,404
5	**Summit** - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States	2,414,592	148.60	200.79	10,096

But progress in km-scale modelling is tough…



Adapted from Neumann, Dueben et al. Phil Trans A 2018



PERSPECTIVE

https://doi.org/10.1038/s43588-021-00023-0

nature computational science

Check for updates

The digital revolution of Earth-system science

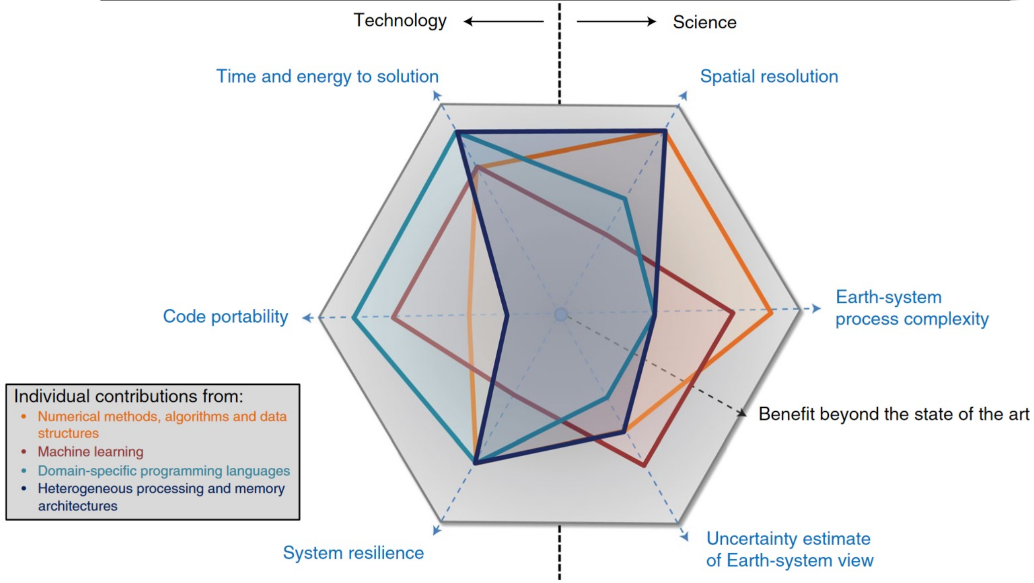
Peter Bauer[1], Peter D. Dueben[1], Torsten Hoefler[2], Tiago Quintino[3], Thomas C. Schulthess[4] and Nils P. Wedi[1]

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

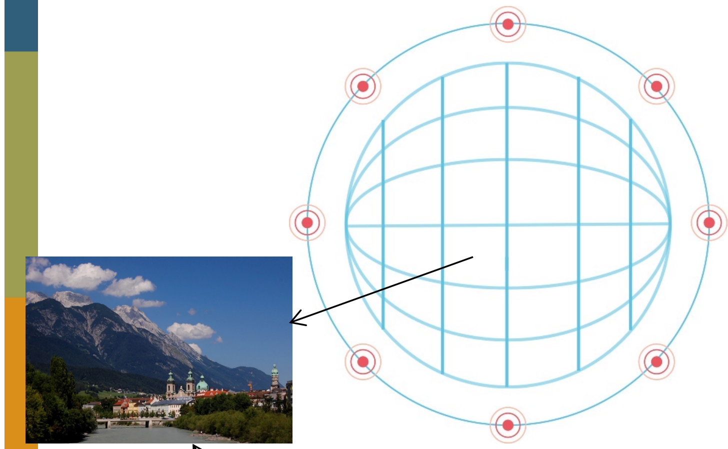
commodity parallel processing. Moore's law drove the economics of computing by stating that every 18 months, the number of transistors on a chip would double at approximately equal cost. However, the cost per transistor starts to grow with the latest chip generations, indicating an end of this law. Therefore, 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.



Current Systems

Earth System
models & observations

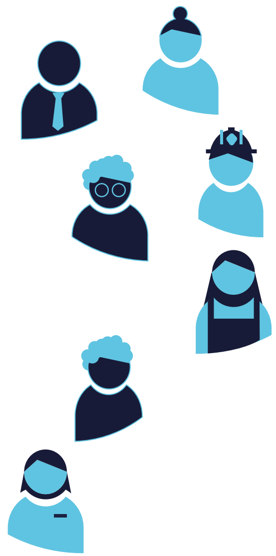


- Limited resolutions
- Small-scall processes not represented
- Separation of earth system & impact sector models

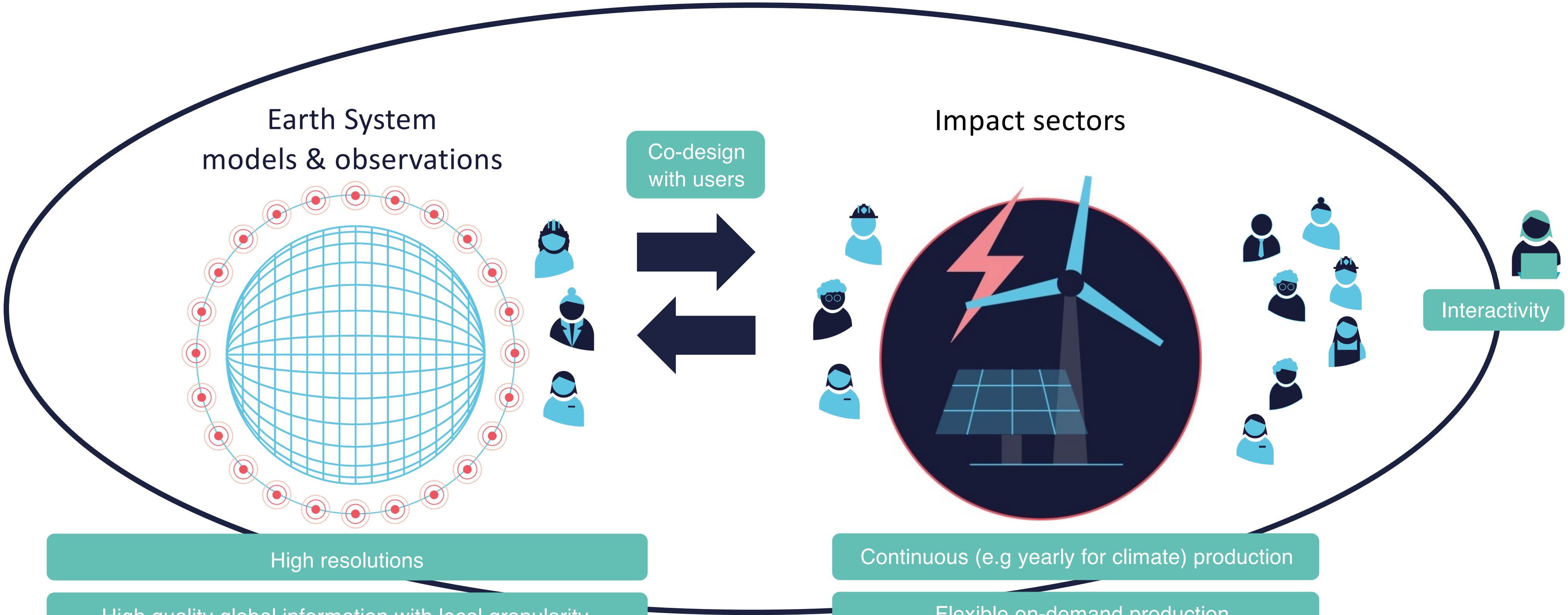
Impact sectors

- Limited ways to change experiment design or output
- Limited versatility to access and interact with the data
- Pre-defined configurations & simulation run times (every 7 years for climate!)

Users



DestinE builds Digital Twins of the Earth



High resolutions

High quality global information with local granularity

Bringing impact sector models WITHIN the same workflow

Continuous (e.g yearly for climate) production

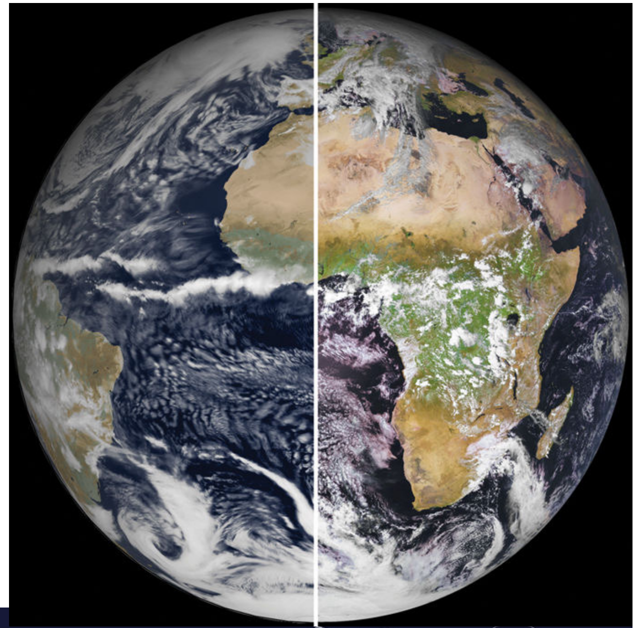
Flexible on-demand production

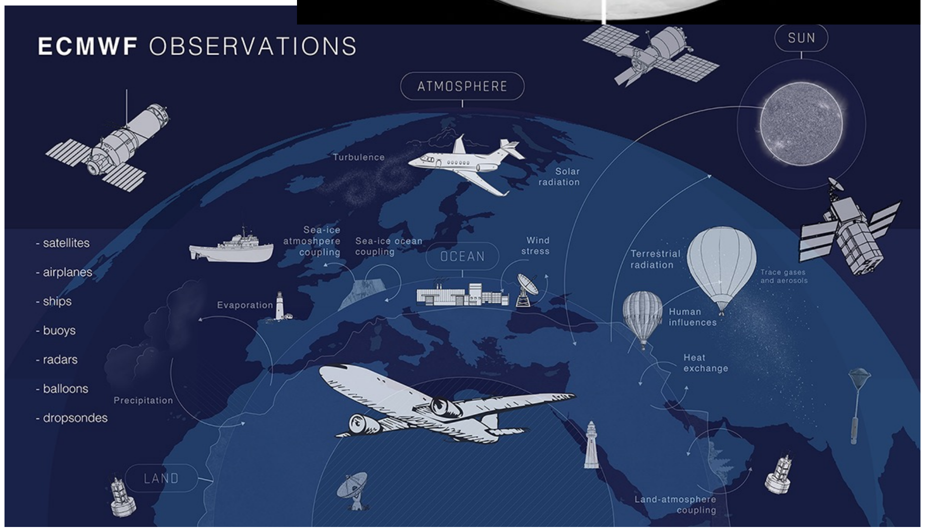
Data streaming and tailored data access

What about the machine learning revolution?

Machine Learning – Why in Earth System modelling

Inputs —— machine learning ——> Outputs





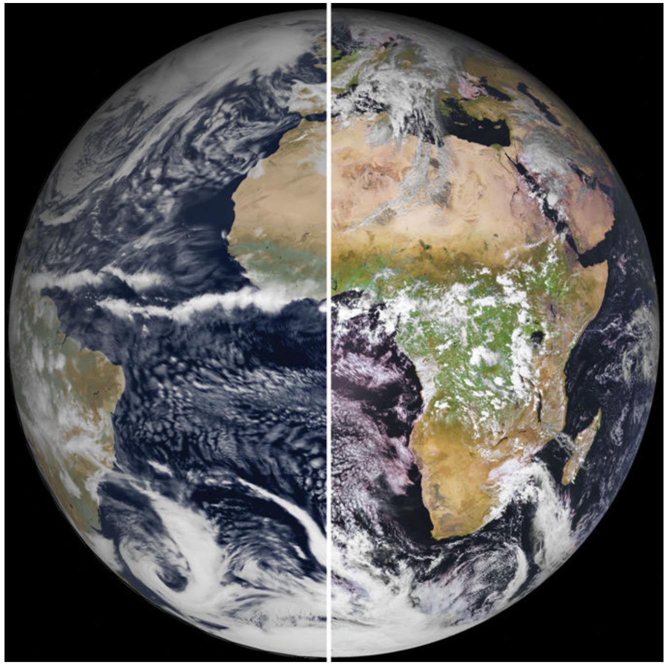
Earth system science is difficult as the Earth system is huge, complex and chaotic, and as the resolution of our models is limited

However, we have a huge amount of observations and Earth system data

- ➢ **There are many application areas for machine learning in Earth system science**

Explore the space of machine learning for weather and climate modelling

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



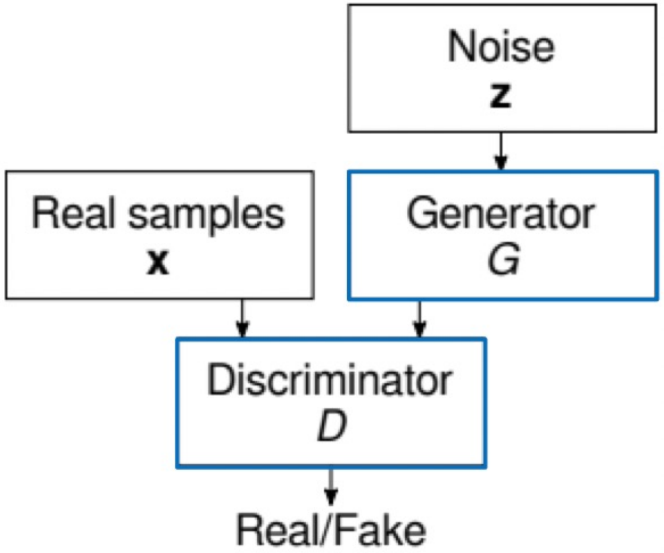
Speed up simulations
- Emulate model components
- Port emulators to heterogeneous hardware
- Use reduced numerical precision and sparse machine learning
- Optimise HPC and data workflow
- Data compression
- …

Improve models
- Learn components from observations
- Correct biases
- Quality control of observations and observation operators
- Feature detection
- …

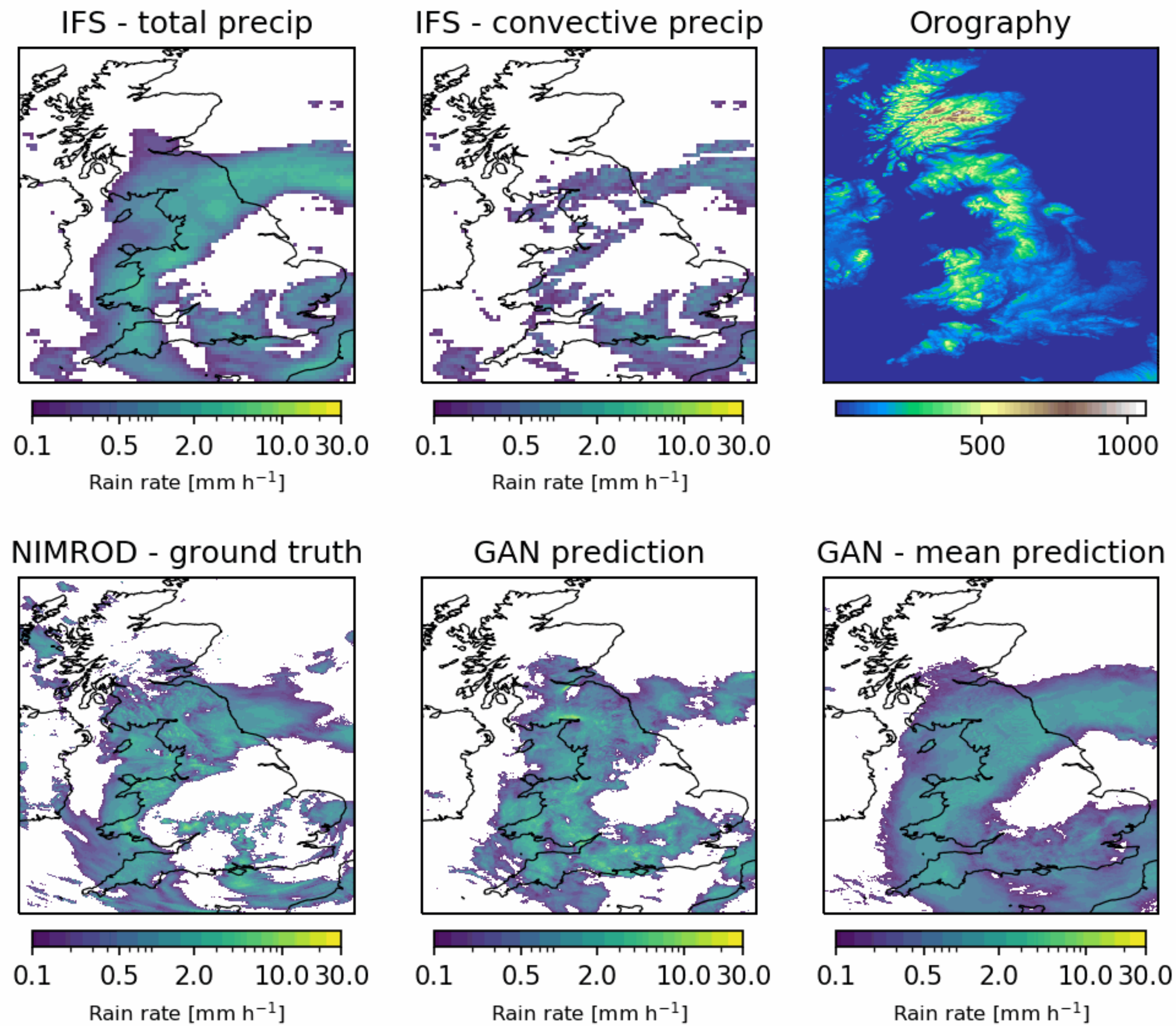
Link communities
- Health – e.g. for predictions of risks
- Energy – e.g. for local downscaling
- Transport – e.g. to combine weather and IoT data
- Pollution – e.g. to detect sources
- Extremes – e.g. to predict wild fires
- …

Weather and climate modelling centres mostly explore the "hybrid space" coupling machine learning to conventional models.

Science example 1: Downscaling with Generative Adversarial Networks

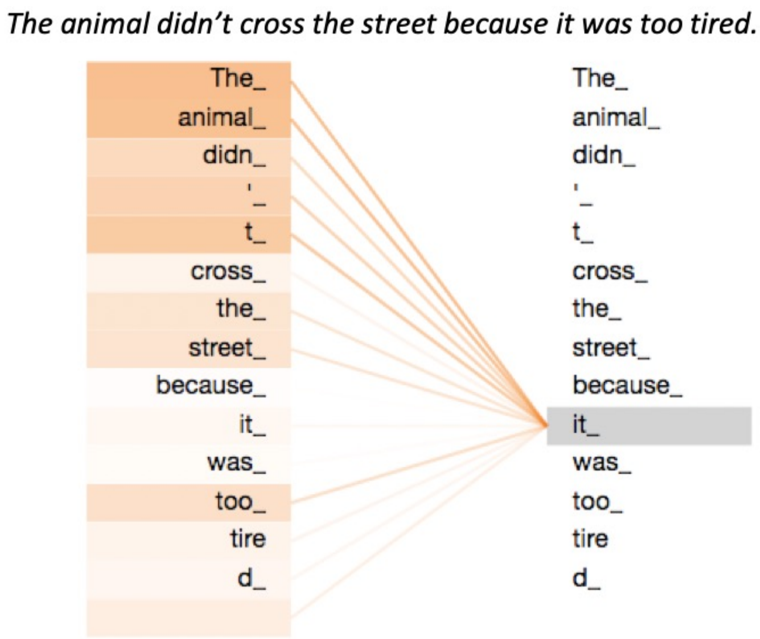


Input: IFS Model Simulation fields on coarse (9 km) grid
Output: Precipitation observation on fine (1 km) grid



Harris, McRae, Chantry, Dueben, Palmer JAMES 2022

Science example 2: Transformer networks for ensemble post-processing



https://jalammar.github.io/illustrated-transformer/

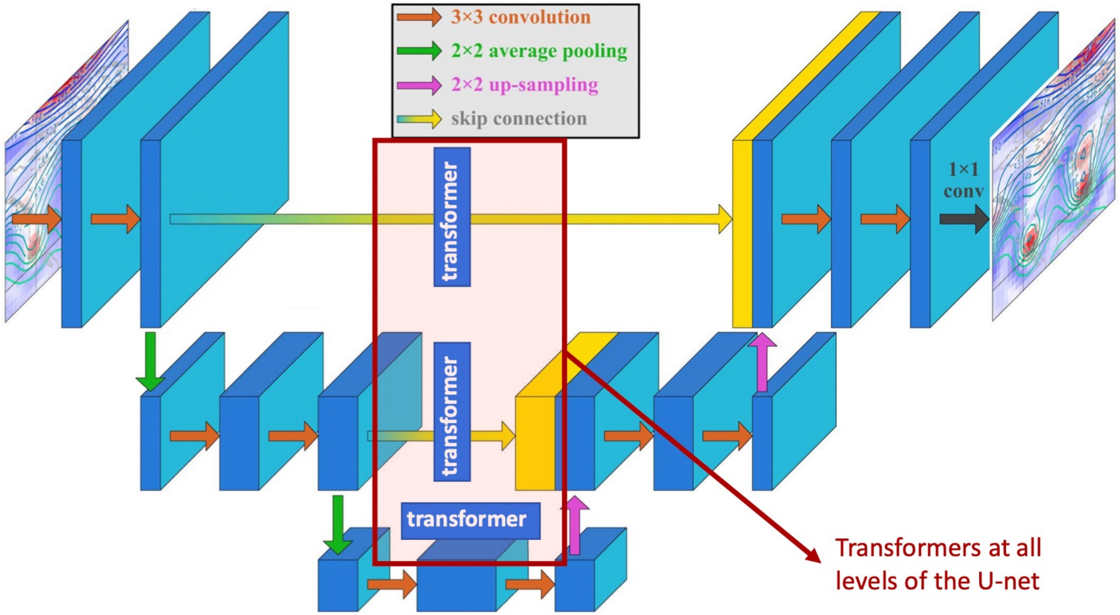
Let's use transformers in the ensemble dimension following the work of Tobias Finn

"Self-Attentive Ensemble Transformer: Representing Ensemble Interactions in Neural Networks for Earth System Models." *arXiv preprint arXiv:2106.13924*

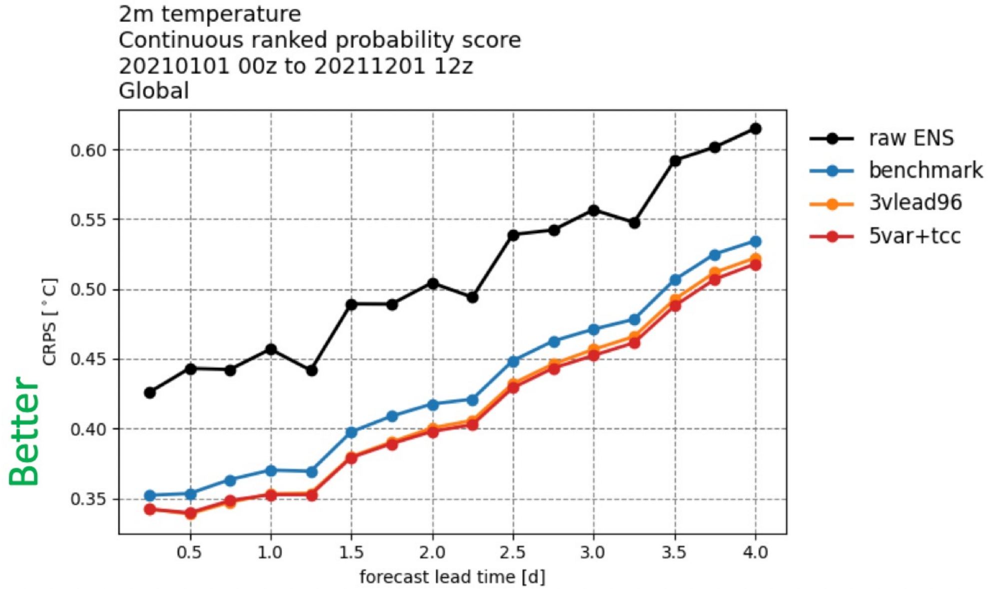
Let's test this for hindcast ensembles in a collaboration between Microsoft and ECMWF

Ben Bouallegue, Weyn, Clare, Dramsch, Dueben, Chantry arXiv 2023

Science example 2: Transformer networks for ensemble post-processing



CRPS



In comparison to the ENS-10 benchmarks from https://arxiv.org/abs/2206.14786

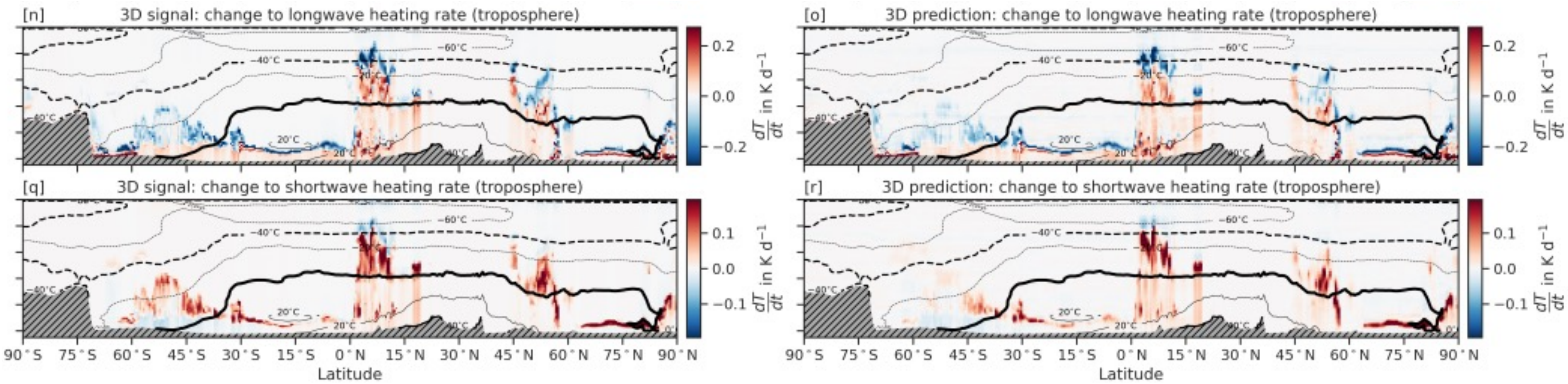
Metric	Model	**Z500** $[\mathrm{m}^2\,\mathrm{s}^{-2}]$ 5-ENS	**Z500** 10-ENS	**T850** [K] 5-ENS	**T850** 10-ENS	**T2m** [K] 5-ENS	**T2m** 10-ENS
CRPS	Raw	81.03	78.24	0.748	0.719	0.758	0.733
	EMOS	$79.08^{\pm 0.739}$	$81.74^{\pm 6.131}$	$0.725^{\pm 0.002}$	$0.756^{\pm 0.052}$	$0.718^{\pm 0.003}$	$0.749^{\pm 0.054}$
	MLP	$75.84^{\pm 0.016}$	$74.63^{\pm 0.029}$	$0.701^{\pm 2e-4}$	$0.684^{\pm 4e-4}$	$0.684^{\pm 6e-4}$	$0.672^{\pm 5e-4}$
	LeNet	$\mathbf{75.56}^{\pm 0.101}$	$\mathbf{74.41}^{\pm 0.109}$	$0.689^{\pm 2e-4}$	$0.674^{\pm 2e-4}$	$0.669^{\pm 7e-4}$	$0.659^{\pm 4e-4}$
	U-Net	$76.66^{\pm 0.470}$	$76.25^{\pm 0.106}$	$0.687^{\pm 0.003}$	$0.669^{\pm 0.009}$	$0.659^{\pm 0.005}$	$0.644^{\pm 0.006}$
	Transformer	$77.30^{\pm 0.061}$	$74.79^{\pm 0.118}$	$\mathbf{0.686}^{\pm 0.002}$	$\mathbf{0.665}^{\pm 0.002}$	$\mathbf{0.649}^{\pm 0.004}$	$\mathbf{0.626}^{\pm 0.004}$
PoET U-net Transformer			**73.97**		**0.650**		

Ben Bouallegue, Weyn, Clare, Dramsch, Dueben, Chantry arXiv 2023

Science example 3: Improve results via emulation

To represent 3D cloud effects for radiation (SPARTACUS) within simulations of the Integrated Forecast Model is four time slower than the standard radiation scheme (Tripleclouds)

Can we emulate the difference between Tripleclouds and SPARTACUS using neural networks?



Rel. Cost	Tripleclouds	SPARTACUS	Neural Network	Tripleclouds+Neural Network
	1.0	4.4	0.003	1.003

Meyer, Hogan, Dueben, Mason JAMES 2022

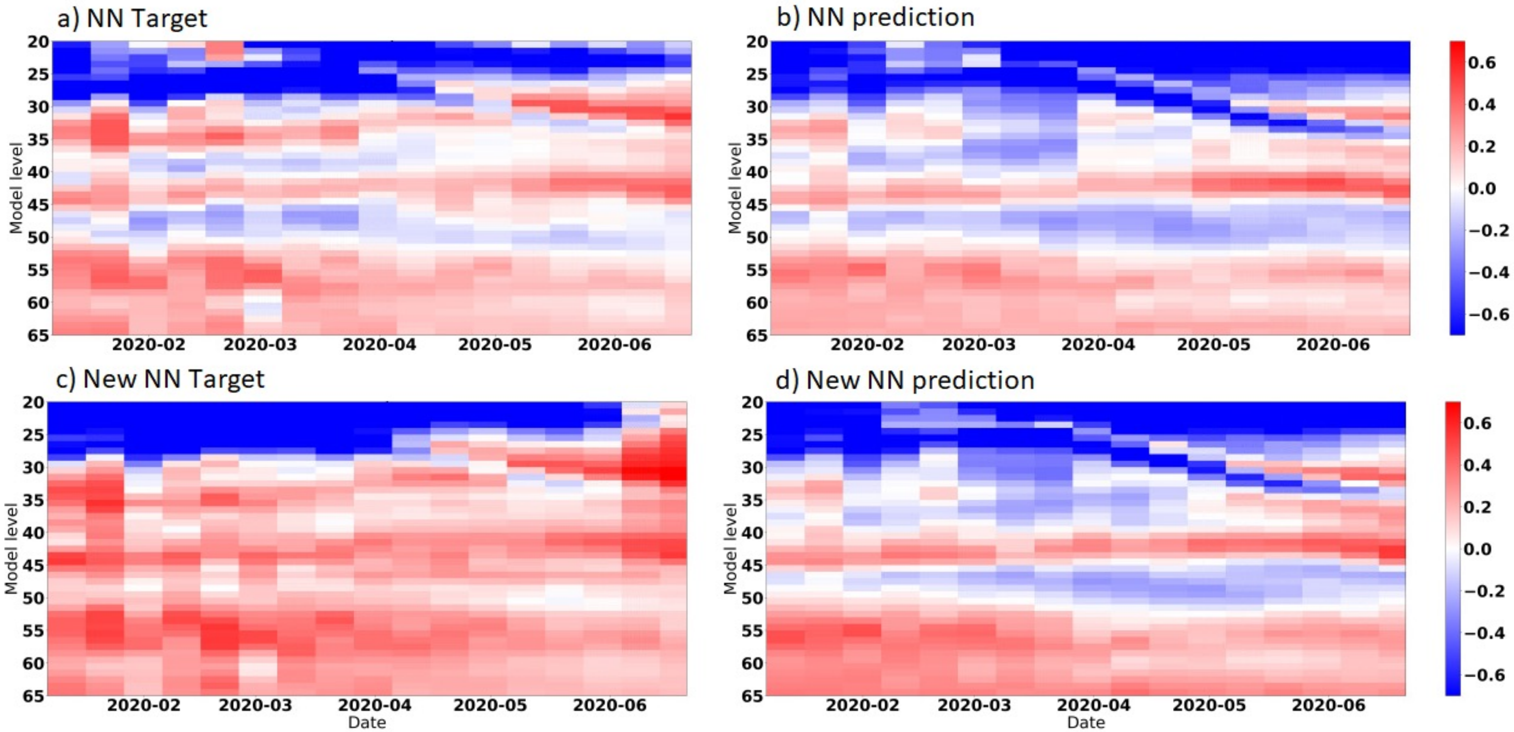
Learn how to combine operational models and machine learning

- During data-assimilation the model trajectory is “synchronised” with observations
- It is possible to learn model error when comparing the model with (trustworthy) observations

Approach: Learn model error from a direct comparison of the model trajectory and observations

Benefit: Correct for model error and understand model deficiencies

Question: What happens when the model is upgraded and the error pattern change?



Laloyaux, Kurth, Dueben, Hall JAMES 2022

Dude, when do you finally wake up and do machine learning for the full model?

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

THE SHIFT

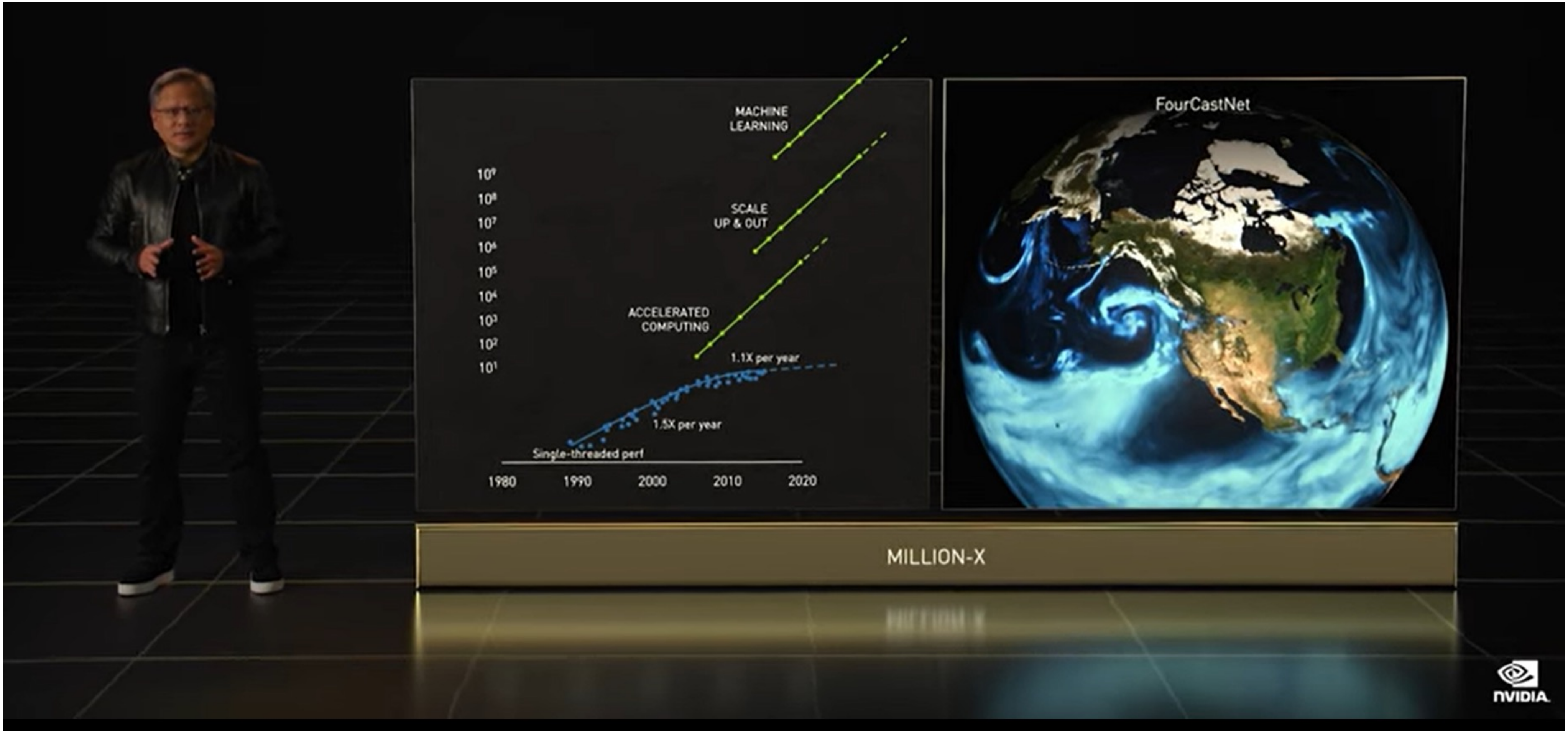
An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy.

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

Give this article 1.5K

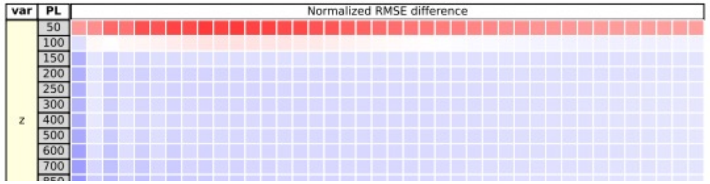
Jason Allen's A.I.-generated work, "Théâtre D'opéra Spatial," took first place in the digital category at the Colorado State Fair. via Jason Allen

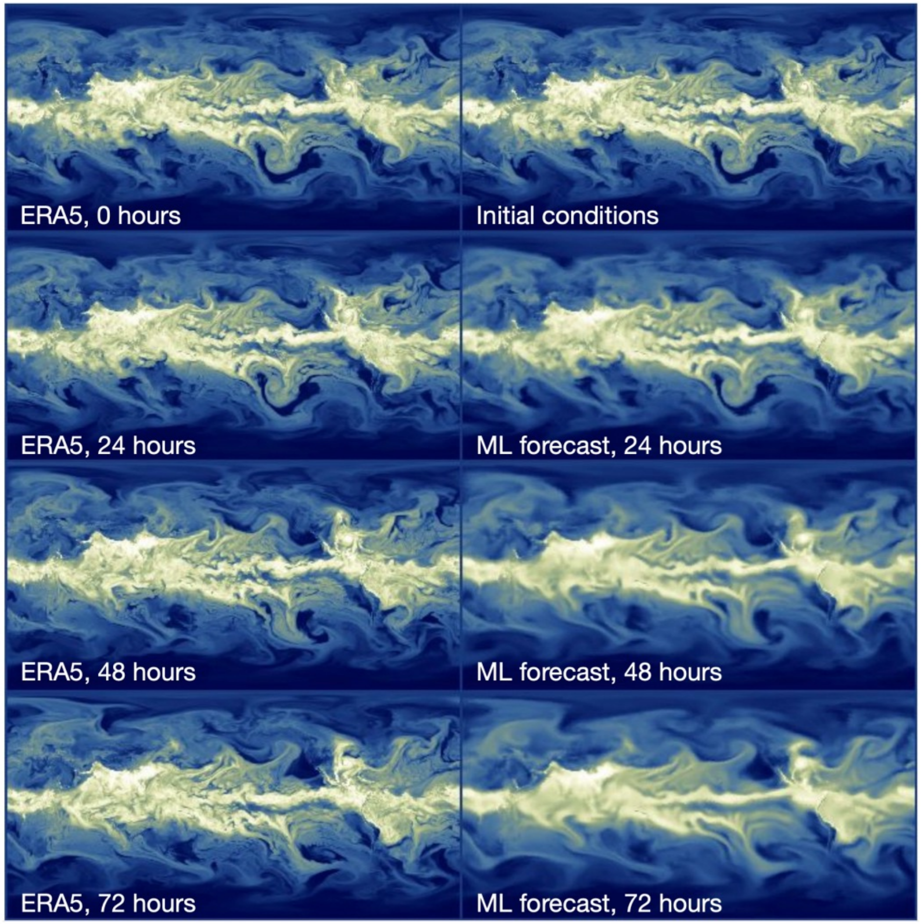
Can we replace conventional Earth System models by deep learning?



NIVIDA’s Earth-2 is coming with FourCastNet

Can we replace conventional Earth System models by deep learning?





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.

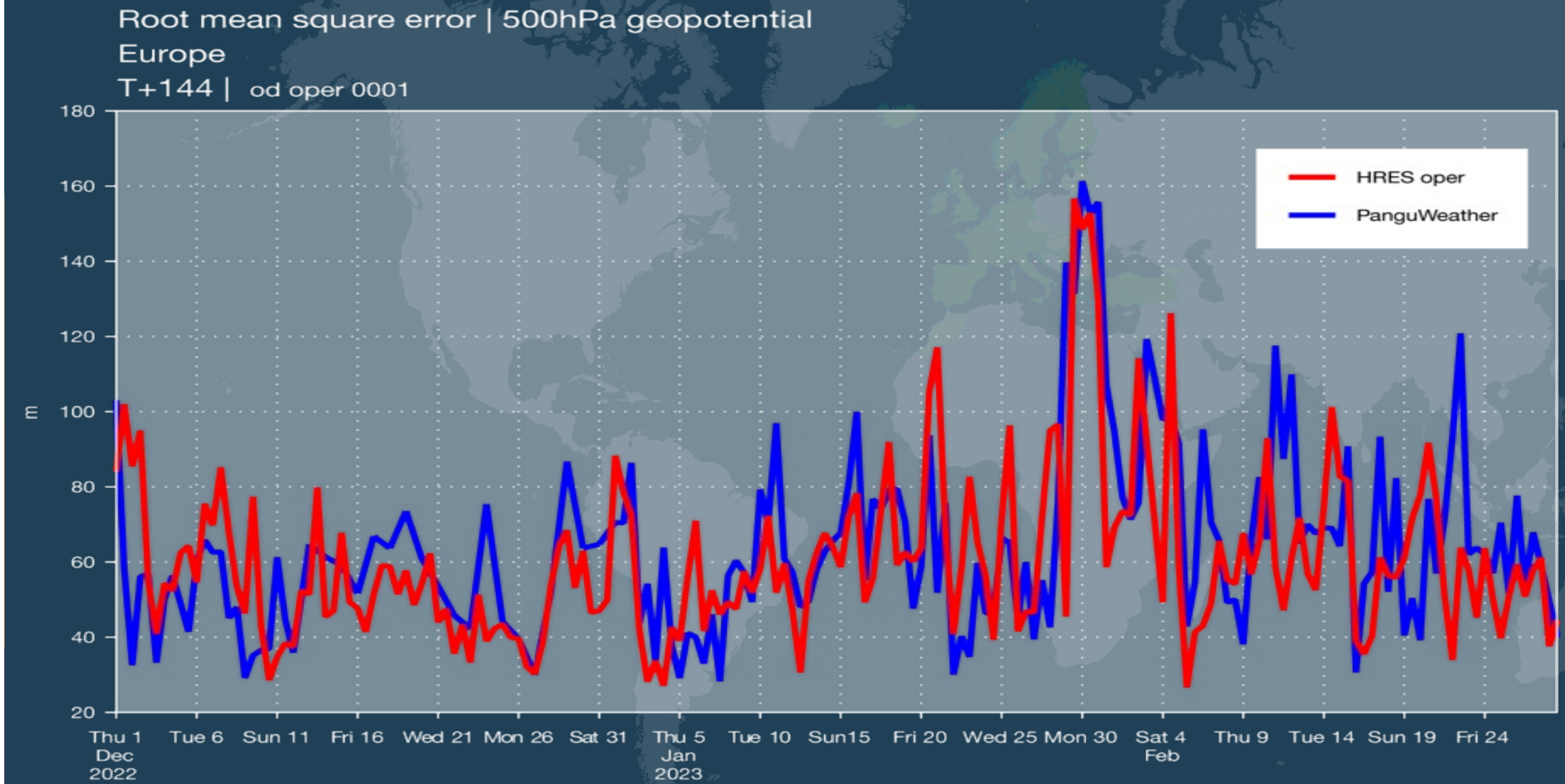
They are trained for a small Root Mean Square Error.
→ They smear out for large lead times.

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

Images from Keisler (2022)

What results are showing: **Time-series of day 6, RMSE over Europe**

Same starting point....Similar results

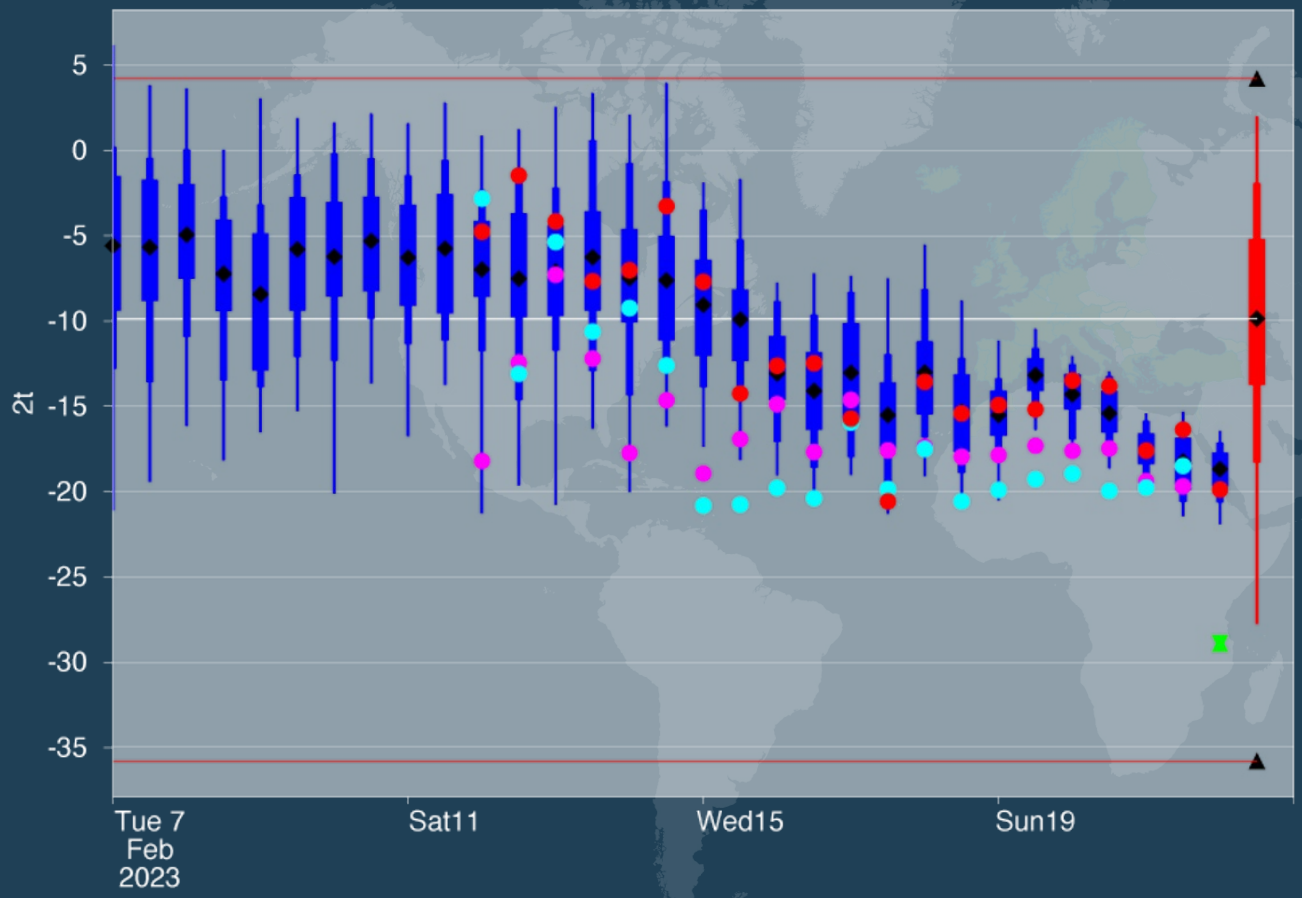


PanguWeather vs ECMWF

- Results extremely close...
- ‘forecast busts’ at similar time
- More ‘physical’ than one might think!

Figure from
Zied Ben-Bouallegue

What the forecasts are showing: **Severe Cold** / Sodankylä, Finland, 22 Feb 00UTC

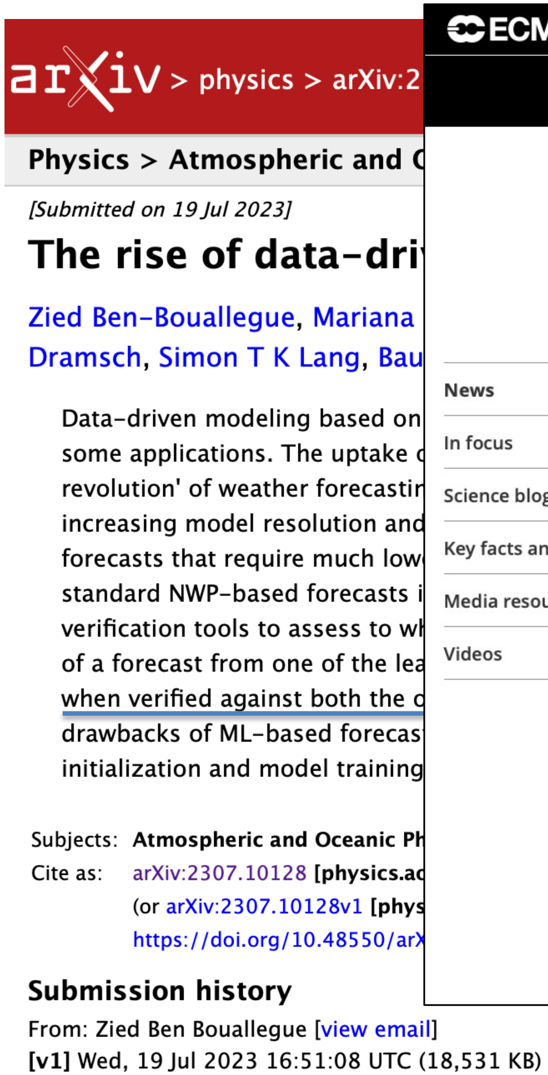


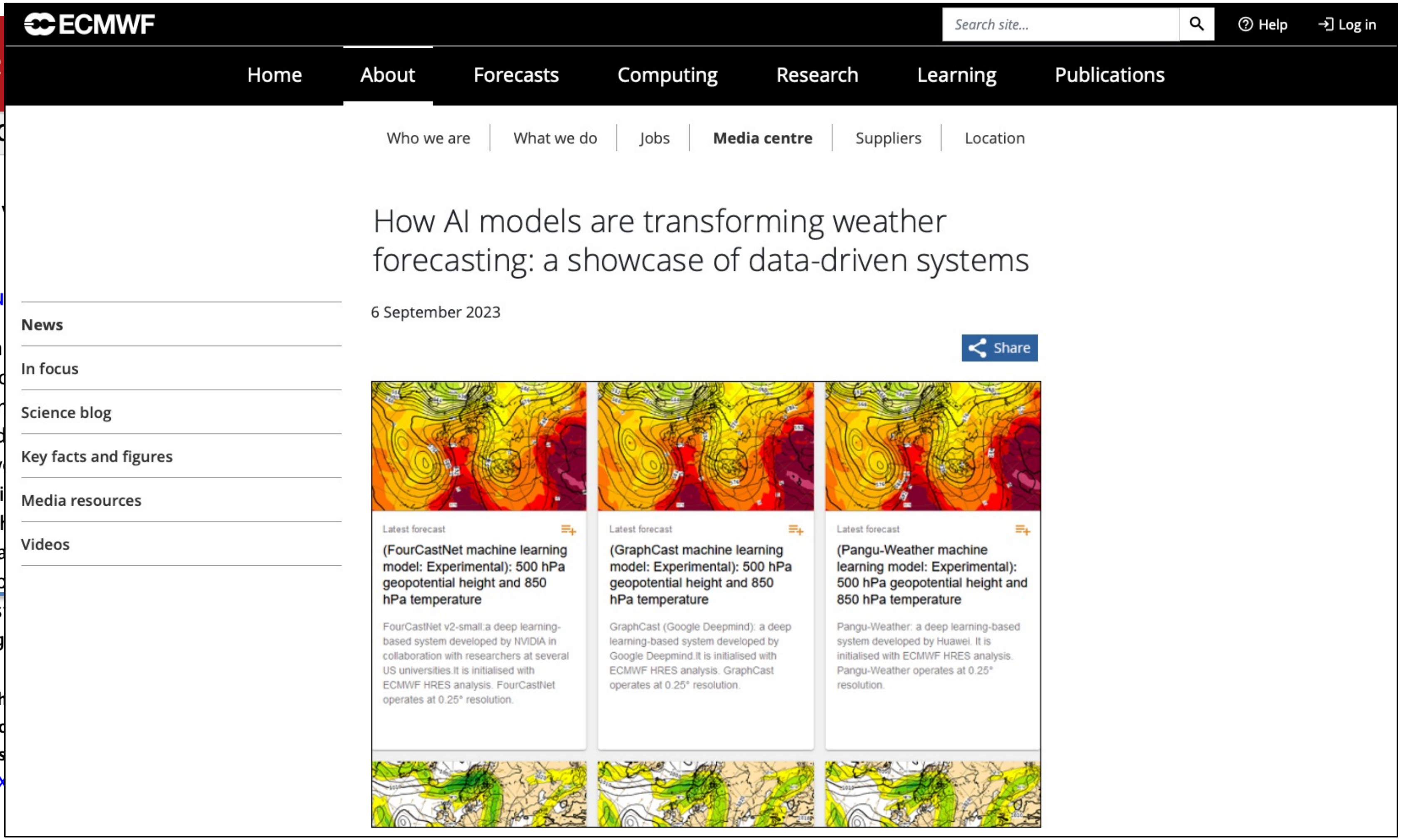
To explore the ability of data-driven models to capture extreme events we examine a case study from Finland from earlier this year, when -29C was observed.
We find that Pangu and FourCastNet recognised the severity of this event earlier, however all models underestimated the temperature significantly, to a similar degree.

Observation – green hourglass
IFS HRES – red dot
IFS ENS - blue
Pangu – cyan dot
FourCastNet – magenta dot
Climatology – red box plot

Figure from
Zied Ben-Bouallegue

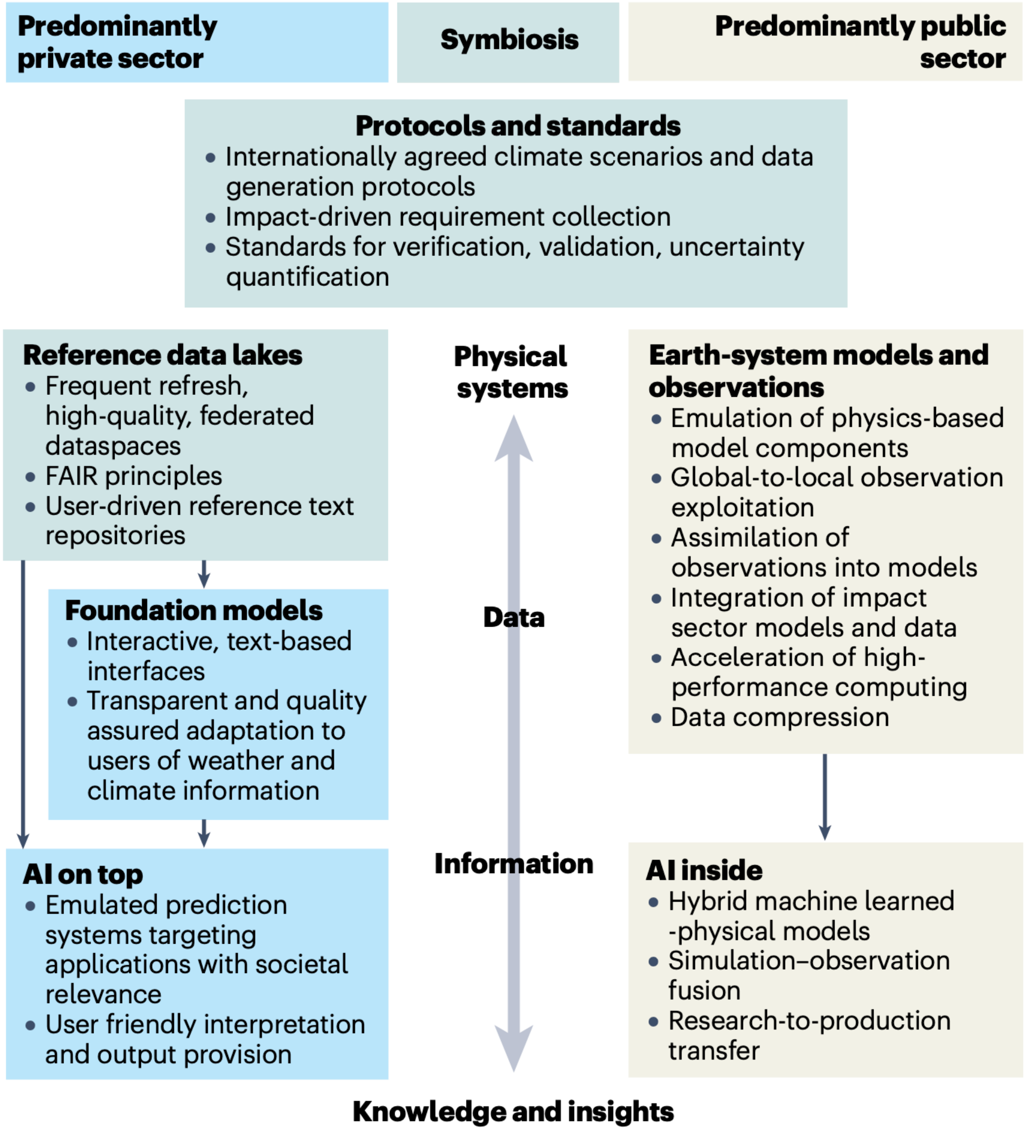
Can we replace conventional Earth System models by deep learning?





https://www.ecmwf.int/en/about/media-centre/news/2023/how-ai-models-are-transforming-weather-forecasting-showcase-data

How will ML for weather and climate evolve in a public/private partnership?




nature reviews earth & environment

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

nature > nature reviews earth & environment > comment > article

Comment | Published: 01 August 2023

Deep learning and a changing economy in weather and climate prediction

Peter Bauer, Peter Dueben, Matthew Chantry, Francisco Doblas-Reyes, Torsten Hoefler, Amy McGovern & Bjorn Stevens

Nature Reviews Earth & Environment **4**, 507–509 (2023) | Cite this article

638 Accesses | **34** Altmetric | Metrics

The rapid emergence of deep learning is attracting growing private interest in the traditionally public enterprise of numerical weather and climate prediction. A public–private partnership would be a pioneering step to bridge between physics- and data-based methods, and necessary to effectively address future societal challenges.

You want to learn more? – Have a look at our MOOC material

ECMWF Massive Open Online Course (MOOC) on Machine Learning in Weather & Climate:
https://lms.ecmwf.int/course/index.php?categoryid=1

40h of content, >9000 registered participants, 159 countries, 60 experts, 47 videos

What have we learned?

The quiet revolution (1980-2015):
- Investment into Earth system modelling and Earth system observations can make a huge difference.

The digital revolution (2015-today):
- Conventional models need to be made future proof via the use of new coding standards.
- Km-scale models are possible today and are starting to make a difference.

The machine learning revolution (2022-today):
- *A PhD student can write a machine learning tool of 2,000 lines of Python code that can beat the best weather prediction model in the world based on hundreds of person years of work and 1,000,000 lines of Fortran code.*
- Data needs to be open and easy to use to make progress.

What machine learned models can and cannot do

- Conventional models will not be replaced by machine learning models entirely.
- Machine learned models can predict weather extremes.
- Within the next couple of years most weather predictions will come from machine learning models.
- The availability and quality of data (observations, reanalysis and models) limits the quality of predictions.
- Km-scale models will make a difference for the generation of training datasets.
- Machine learning models will be able to predict the climate despite the current extrapolation problem.
- Not many meteorologists will be replaced by machine learning models.

What will we gain, what will we need?

What will we gain?

Better predictions for local and global weather and climate.

Models will become easier to use and easier to trigger.

It will be easy to build a specific machine learning model for a specific application.

What will we need?

Federated access to weather and climate data.

Projects such as DestinE and EVE to provide the infrastructure.

Many thanks! **Peter.Dueben@ecmwf.int** **@PDueben**

ECMWF | EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

The strength of a common goal