Machine learning, high-performance computing and numerical weather prediction

Peter Dueben
Royal Society University Research Fellow & ECMWF's Coordinator for Machine Learning and AI Activities

The strength of a common goal

This research used resources of the Oak Ridge Leadership Computing Facility (OLCF), which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.

The ESIWACE, MAELSTROM and AI4Copernicus projects have received funding from the European Union under grant agreement No 823988, 955513 and 101016798.

Machine learning in three communities

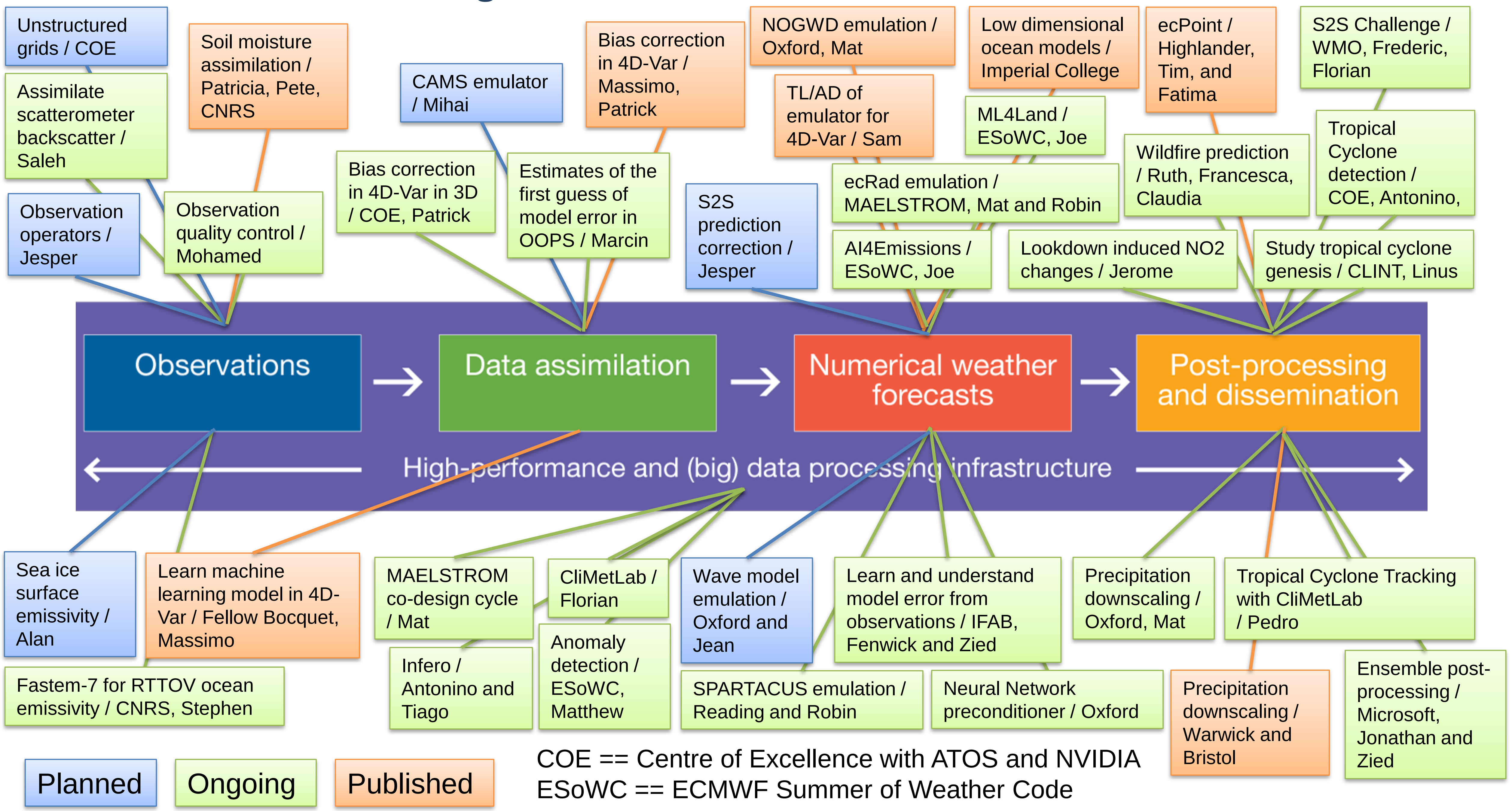
How did the view on machine learning change since the last HPC workshop in 2018?

The bold Machine Learning scientist:
“Machine learning will replace everything”
→ “Machine learning will replace everything, look here…”

The HPC hardware developer:
“Machine learning will dominate future HPC developments”
→ “Here is our new machine learning hardware, please use it”

The sceptical weather and climate domain scientist:
“Machine learning is just a wave going through…”
→ “Machine learning is just a method…”

Status of machine learning at ECMWF



Machine learning in three communities

How did the view on machine learning change since the last HPC workshop in 2018?

The bold Machine Learning scientist:
“Machine learning will replace everything”
→ “Machine learning will replace everything, look here…”

The HPC hardware developer:
“Machine learning will dominate future HPC developments”
→ “Here is our new machine learning hardware, please use it”

The sceptical weather and climate domain scientist:
“Machine learning is just a wave going through…”
→ “Machine learning is just a method…”

But there is still more that can be done with customised machine learning tools that are easy to use at scale.

Challenges for machine learning in weather and climate modelling

Different sets of tools for domain (Fortran on CPUs) and machine learning scientists (Python on GPUs)

Off-the-shelf machine learning tools are often not sufficient for weather and climate applications

Training datasets are often not good enough while the data size is huge

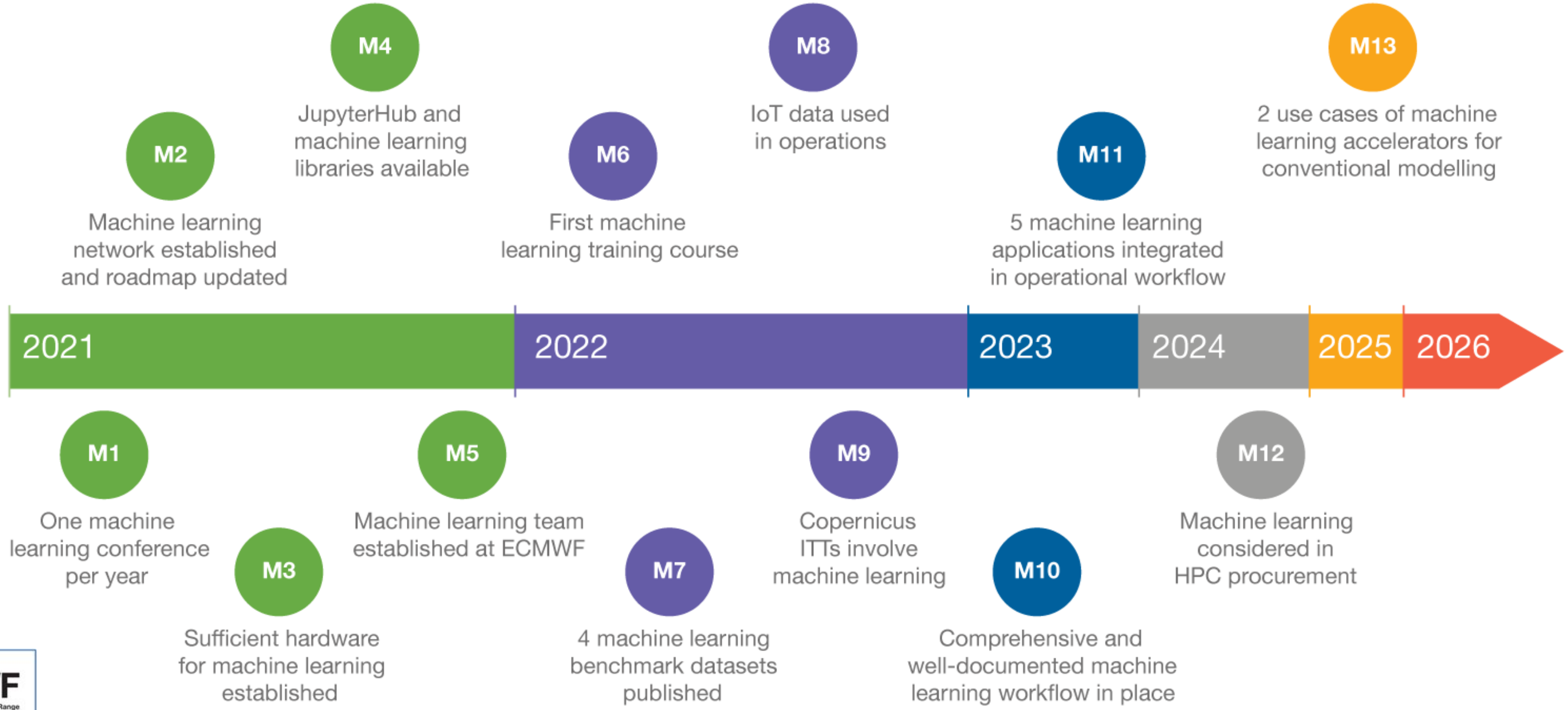
We still need to learn how to scale up to petascale supercomputers to make the most of machine learning

Integration of machine learning tools into the conventional numerical weather prediction workflow is difficult

Machine learning tools need to be updated in model cycles

Machine learning tools need to be reliable (extrapolating?) for use in operational predictions

Step 1: ECMWF's machine learning roadmap



Vision 2031

- It is difficult to distinguish between machine learning and domain sciences
- Data handling fully capable to serve machine learning needs
- Fully supported diagnostic tools via trustworthy AI
- Physical constraints can be represented in deep learning
- Use of machine learning as easy and normal as data re-gridding
- Unsupervised learning and causal discovery used on a regular basis
- Machine learning solutions from end-users integrated in workflow

Objective 1
Explore machine learning applications across the weather and climate prediction workflow and apply them to improve model efficiency and prediction quality.

Objective 2
Expand software and hardware infrastructure for machine learning.

Objective 3
Foster collaborations between domain and machine learning experts with the vision of merging the two communities.

Objective 4
Develop customised machine learning solutions for Earth system sciences that can be applied to various applications and at scale on current and future supercomputing infrastructure.

Objective 5
Train staff and Member and Co-operating State users and organise scientific meetings and workshops.

https://www.ecmwf.int/en/elibrary/19877-machine-learning-ecmwf-roadmap-next-10-years

Step 2: Centre of Excellence in Weather & Climate Modelling

The machine learning project has started end of 2020:

Objective 1: Develop vanilla solutions for the emulation of physical parametrisation schemes.

Objective 2: Develop vanilla solutions for machine learning applications that take the three-dimensional state of the atmosphere on unstructured grids as well as scale interactions in both space and time into account.

Objective 3: Develop vanilla solutions for feature detection in three-dimensional IFS model output.

Objective 4: Develop infrastructure to enable the use of machine learning libraries that are called within IFS on the new HPC.

Objective 5: Enhance data workflow to facilitate machine learning applications that use IFS model output and re-analysis products.

Objective 6: Explore the integration of machine learning with the product-generation workflow.

Step 3: Machine learning benchmark datasets

Benchmark datasets include:

- A problem statement
- Data that is available online
- Python code or Jupyter notebooks
- Quantitative evaluation metrics
- A reference machine learning solution
- Visualisation, diagnostics and robustness tests
- Computational benchmarks

Benchmark datasets are useful because:

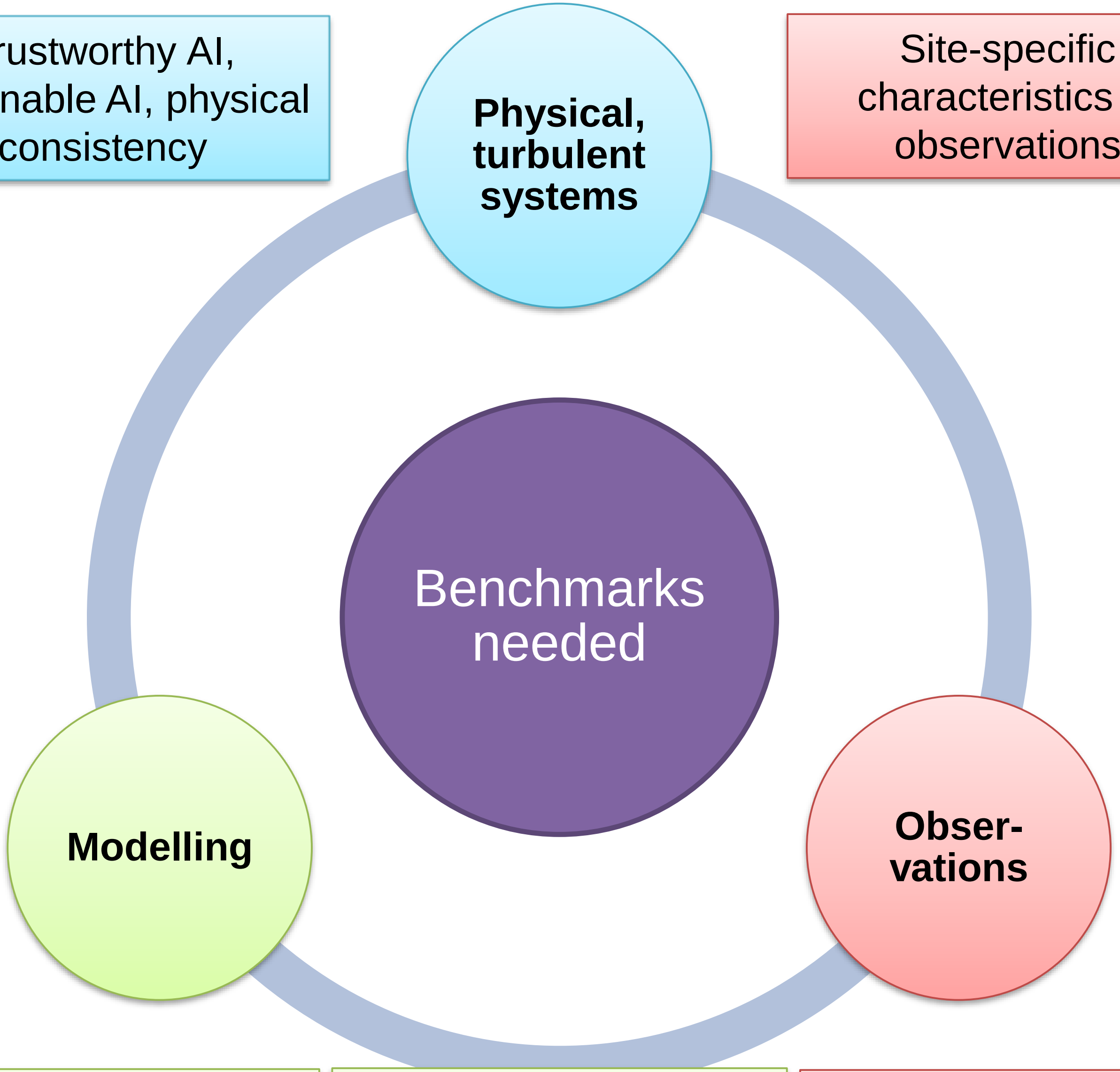
- They allow a quantitative evaluation of machine learning approaches (including efficiency tests with e.g. reduced precision or sparse solutions)
- They reduce data access and help scientists to get access to relevant data
- They allow for a separation of concerns between domain sciences and machine learning experts
- They allow for a separation of concerns between domain sciences and HPC experts, and to learn how to scale

Machine learning equivalent to ESCAPE dwarfs for machine learning in weather and climate science?

There are two joined special issues in GMD and ESSD:
https://gmd.copernicus.org/articles/special_issue386_1147.html
https://essd.copernicus.org/articles/special_issue1147.html

Step 3: Missing machine learning benchmark datasets for atmospheric sciences

- Transfer learning
- Online and reinforcement learning
- Training of machine learning tools in a changing climate
- Unstructured grids on the sphere
- Multi-scale interactions in space and time
- Dataflow and handling of huge datasets
- Trustworthy AI, explainable AI, physical consistency



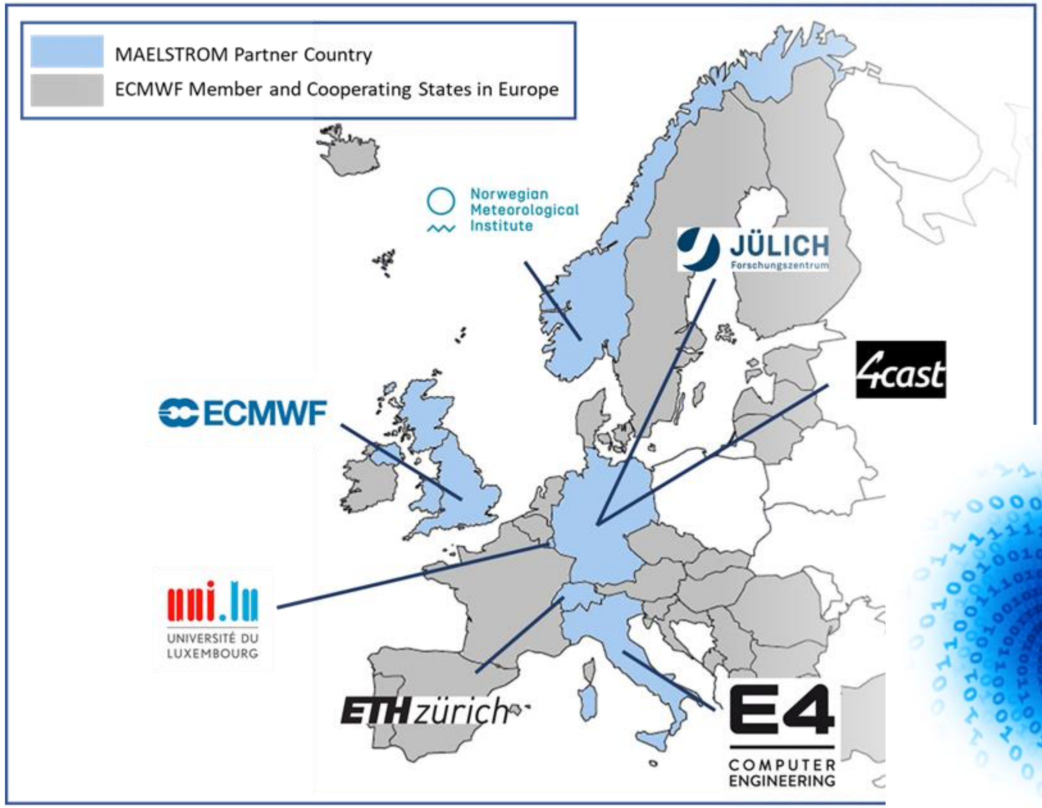
- Hybrid modelling and coupling
- Uncertainty quantification and representation
- Weather and climate predictions based on machine learning
- The emulation of model components
- Post-processing and down-scaling

- Site-specific characteristics of observations
- Composite distributions of observations
- Extreme value predictions
- Missing data and irregular spacing of monitoring sites
- Auto correlation and periodic patterns
- Physical constraints
- Data anonymity
- Fusion of diverse datasets
- Air-quality applications
- Nowcasting applications
- Uncertainty estimates
- Feature detection

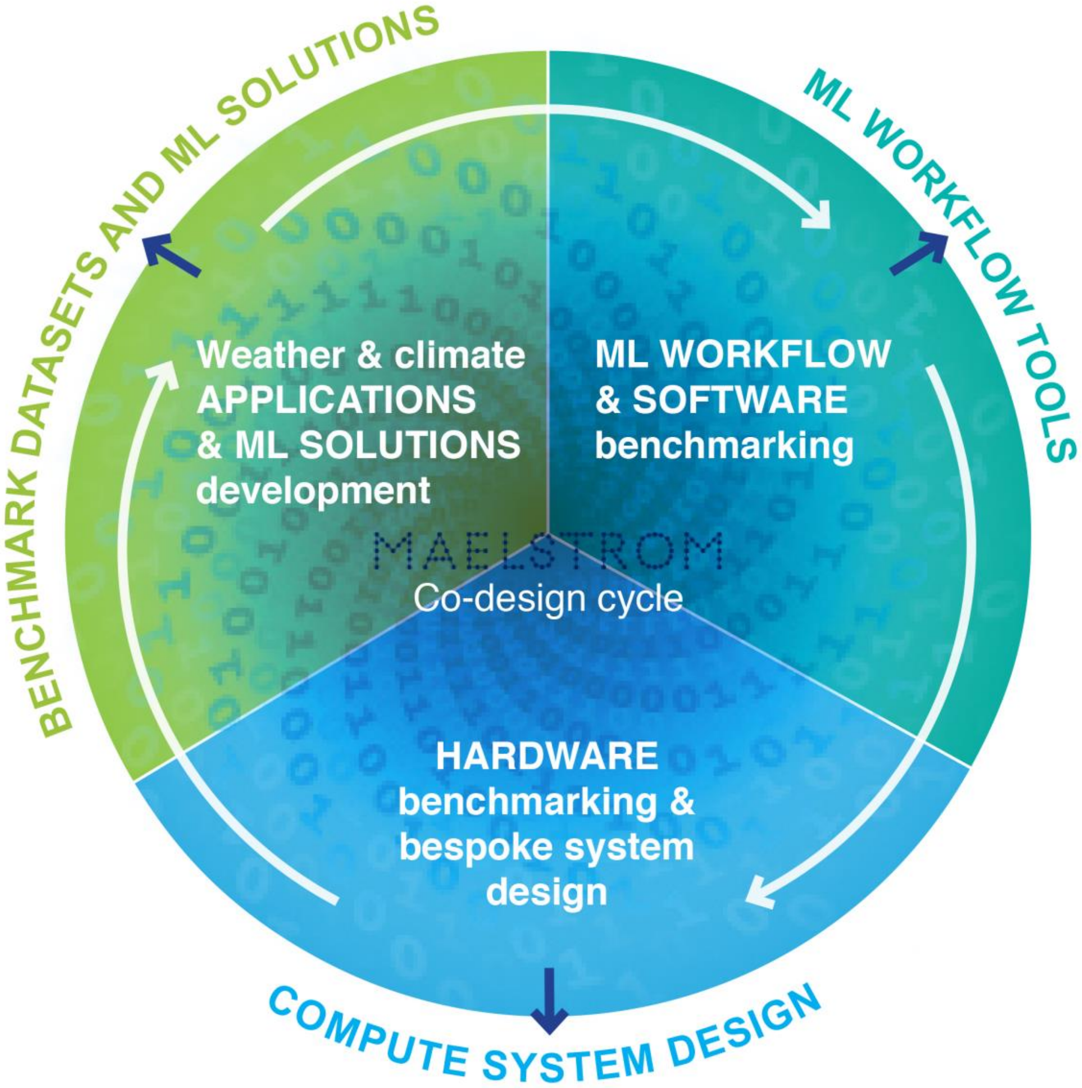
Dueben, Schultz, Chantry, Gagne et al. in preparation

Step 4: The MAELSTROM project

MAchinE Learning for Scalable meTeoROlogy and cliMate



MAELSTROM



The first **benchmark datasets** have been published!
https://www.maelstrom-eurohpc.eu/content/docs/uploads/doc6.pdf

https://www.maelstrom-eurohpc.eu/ @MAELSTROM_EU

Step 5: Science and tool developments

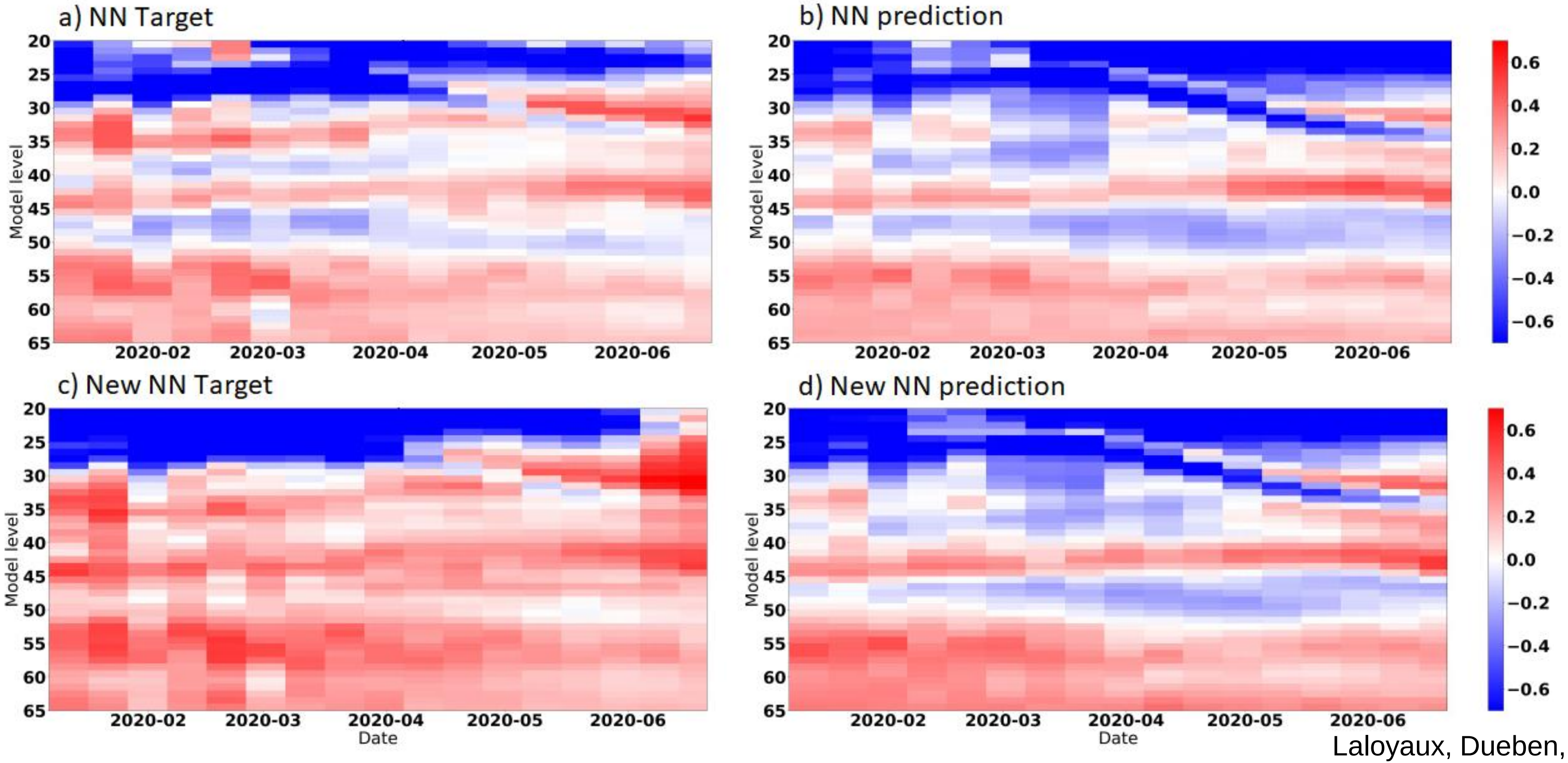
- During data-assimilation the model trajectory is “synchronised” with observations
- Model error can be diagnosed when comparing the model with (trustworthy) observations

Approach: Learn model error for a given model state using machine learning

Benefit: Correct for model error and understand model deficiencies

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

Solution: More work on transfer learning needs to be done



Laloyaux, Dueben, Bonavita @ ECMWF + Kurth and Hall @ NVIDIA

Challenges for machine learning in weather and climate modelling

Different sets of tools for domain (Fortran on CPUs) and machine learning scientists (Python on GPUs)
→ Machine learning roadmap via training and tool development (e.g. CliMetLab)

Off-the-shelf machine learning tools are often not sufficient for weather and climate applications
→ Machine learning roadmap, MAELSTROM and COE via benchmark datasets and tool developments

Training datasets are often not good enough while the data size is huge
→ MAELSTROM via benchmark datasets

We still need to learn how to scale up to petascale supercomputers to make the most of machine learning
→ MAELSTROM via co-design cycle

Integration of machine learning tools into the conventional numerical weather prediction workflow is difficult
→ Science and tool developments, COE, and tool development (e.g. Infero)

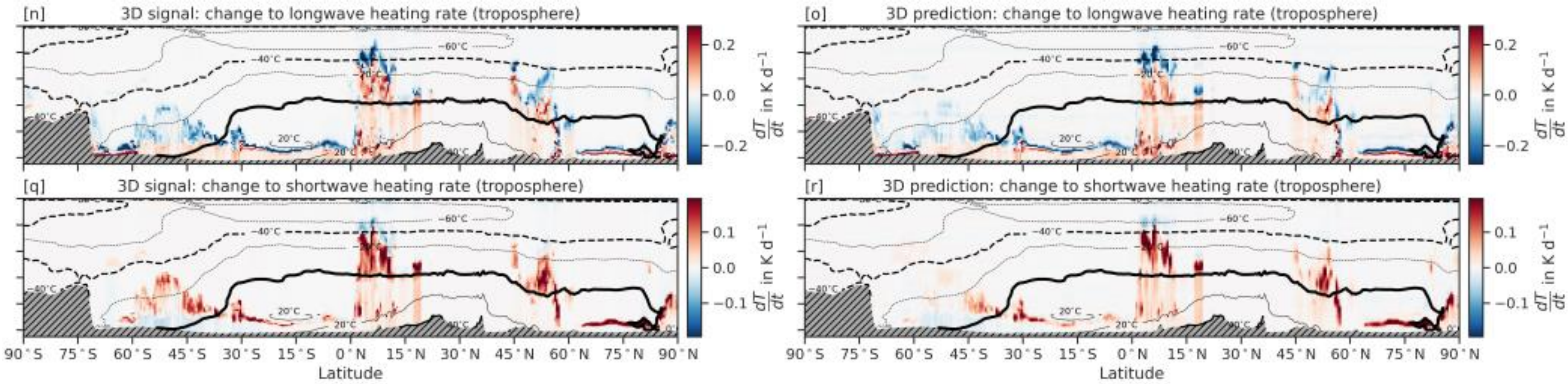
Machine learning tools need to be updated in model cycles
→ Science and tool developments and COE via Transfer Learning

Machine learning tools need to be reliable (extrapolating?) for use in operational predictions
→ Science and tool developments

Research highlight 1: Make models faster

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?

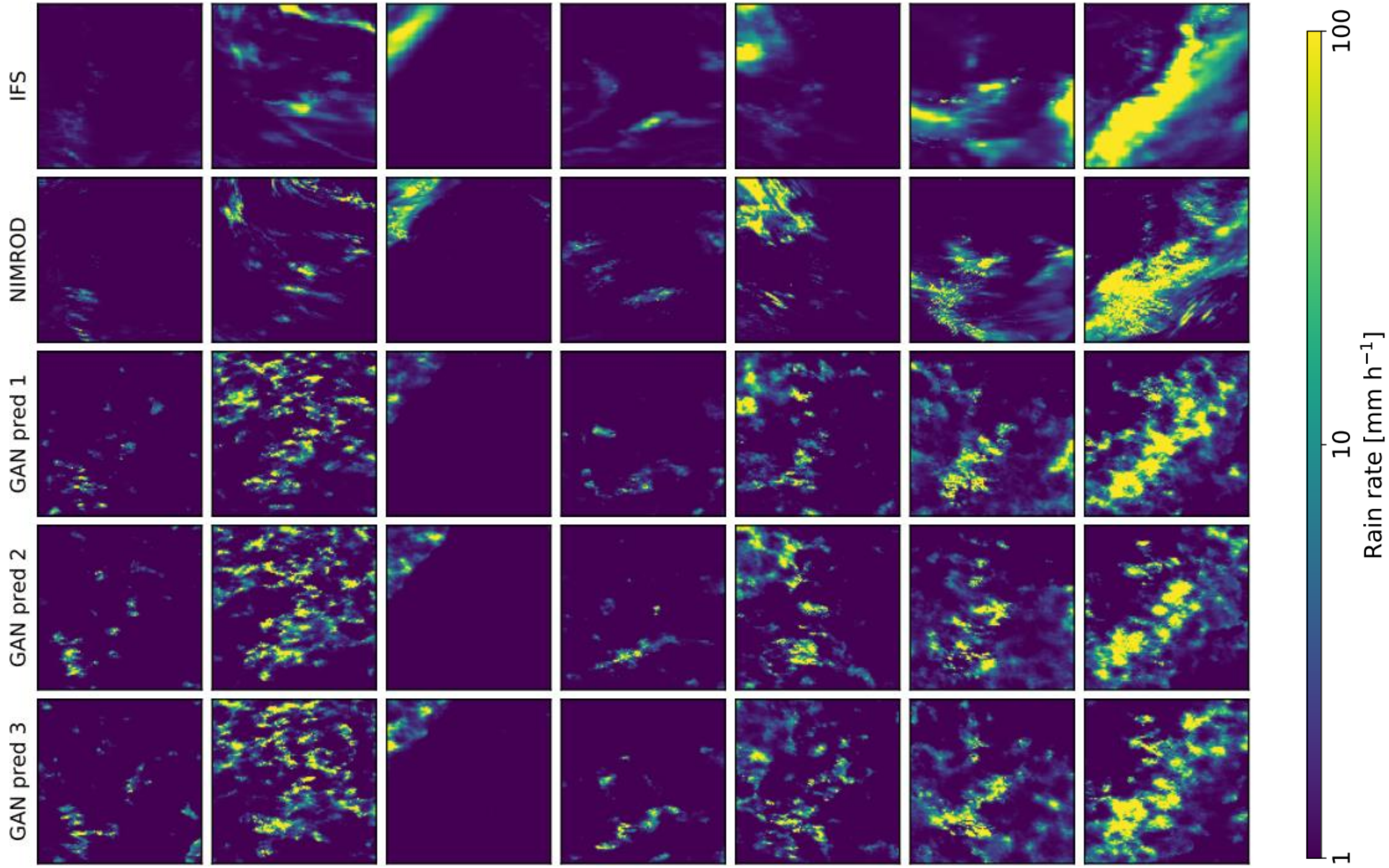


	Tripleclouds	SPARTACUS	Neural Network	Tripleclouds+Neural Network
Relative Cost	1.0	4.4	0.003	1.003

Also see the next talk by Atos.

Meyer, Hogan, Dueben, Mason https://arxiv.org/abs/2103.11919

Research highlight 2: Make models better



- Map IFS model data at ~10 km resolution to NIMROD precipitation observations at ~1 km resolution
- Test Generative Adversarial Networks (GANs) and Variational Autoencoders (VAs)
- Generate ensembles to represent the uncertainty of the mapping.

Harris, McRae, Chantry, Dueben, Palmer

Conclusions

- There are a large number of application areas throughout the prediction workflow in weather and climate modelling for which machine learning can make a difference.
- The weather and climate community is still only at the beginning to explore the potential of machine learning (and in particular deep learning) at scale and there are challenges to be faced.
- However, an approach that combines collaborations, meetings, scientific studies, targeted projects, shared datasets, software and hardware developments should allow us to overcome most of the challenges in the medium-term future.

Please do not forget to register for the ESA-ECMWF Workshop on Machine Learning for Earth System Observation and Prediction – 15-18 November – https://www.ml4esop.esa.int/

Many thanks!

Peter.Dueben@ecmwf.int **@PDueben**

ECMWF | EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

The strength of a common goal

Destination Earth at the horizon…

Check for updates

comment

A digital twin of Earth for the green transition

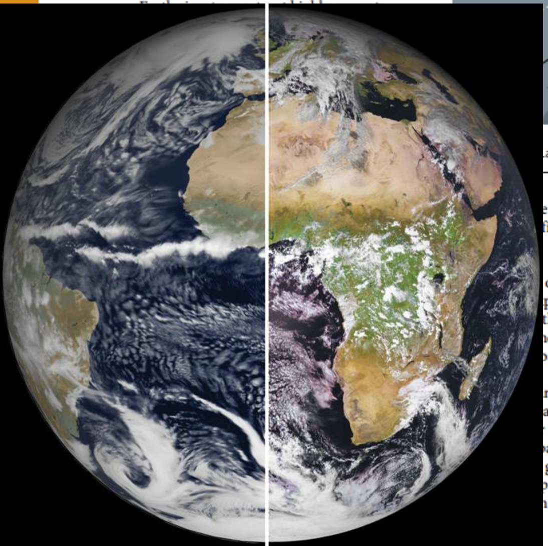
For its green transition, the EU plans to fund the development of digital twins of Earth. For these twins to be more than big data atlases, they must create a qualitatively new Earth system simulation and observation capability using a methodological framework responsible for exceptional advances in numerical weather prediction.

Peter Bauer, Bjorn Stevens and Wilco Hazeleger

The European Union (EU) intends to become climate neutral by 2050, and the set of policies designed to bring about this green transition — the European Green Deal — was announced in December 2019 (ref. [1]). Accompanied by €1 trillion of planned investment, Green Deal policies aim to help the world's second-largest economy sustainably produce energy, develop carbon-neutral fuels and advance circular products in energy-intensive industrial sectors with zero waste and zero pollution.

A key element of the Green Deal is its dependence on the 'digital transformation' — an openly accessible and interoperable European dataspace as a central hub for informed decision making. The EU identified two landmark actions to support the necessary information systems: GreenData4All[2] and Destination Earth[3]. Whereas GreenData4All will develop the European approach to discover, manage and exploit geospatial information, Destination

ayerace / Freepik



e coordinated development
fic disciplines.

of Earth is an information
poses users to a digital
the state and temporal
he Earth system constrained
bservations and the laws of

re familiar with a plethora of
ased monitoring tools that
impact on the environment,
ased simulation models
grasp the causes of change
ptions for future adaptation
actions. The ongoing step

change in the physical content of Earth system models is making them amenable to approaches that harmonize the physical laws they encode with ever more extensive observations to provide the best possible estimate of the state of our planet. Hence, digital twins must focus exactly on how best to realize this convergence of the modelling and observation worlds.

A methodological framework for the twin's architecture already exists in the form of data assimilation, which numerical weather prediction has developed with success over decades[10]. Data assimilation combines data from different observational sources with physical Earth system model simulations to derive an estimate of the state

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

Check for updates

The digital revolution of Earth-system science

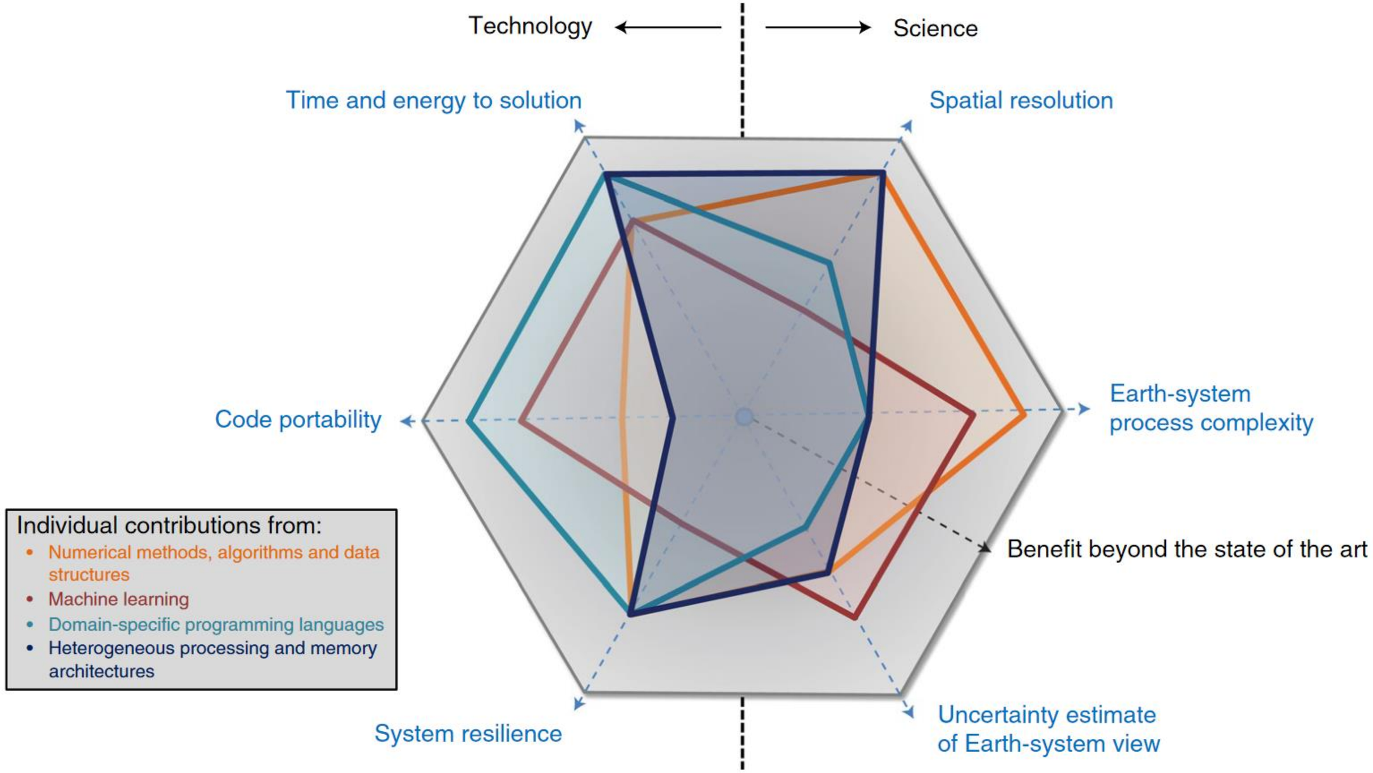
Peter Bauer[1]✉, Peter D. Dueben[1], Torsten Hoefler[2], Tiago Quintino[3], Thomas C. Schulthess[4] and Nils P. Wedi[1]

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 stag-

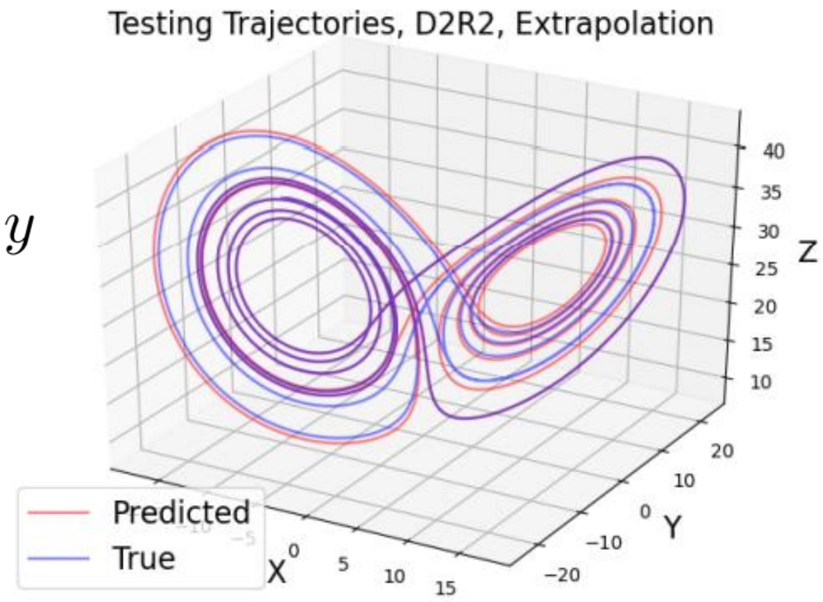
commodity parallel processing. Moore's law drove the economics of computing by stating that every 18 months, the number of transistors on a chip would double at approximately equal cost. However, the cost per transistor starts to grow with the latest chip generations, indicating an end of this law. Therefore,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

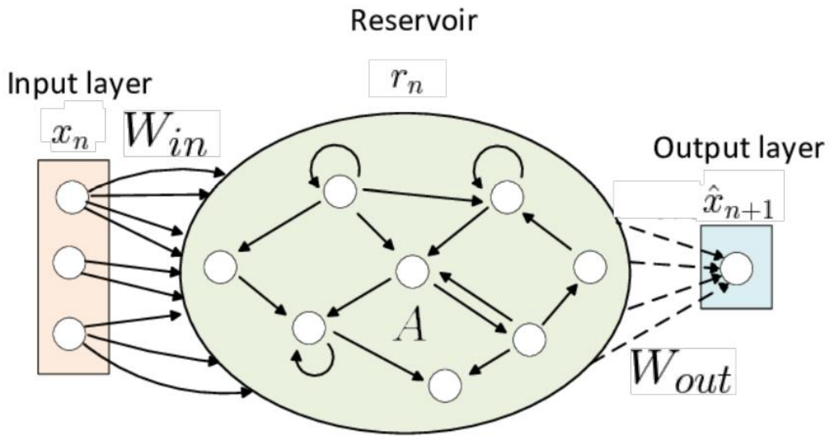
In this Perspective, we will present potential solutions to adapt our current algorithmic framework to best exploit what new digital technologies have to offer, thus paving the way to address the aforementioned challenges. In addition, we will propose the concept of

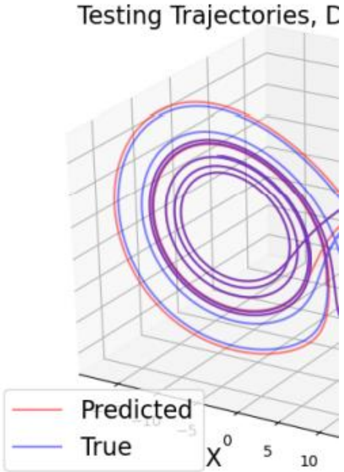


Step 5: Science and tool developments

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



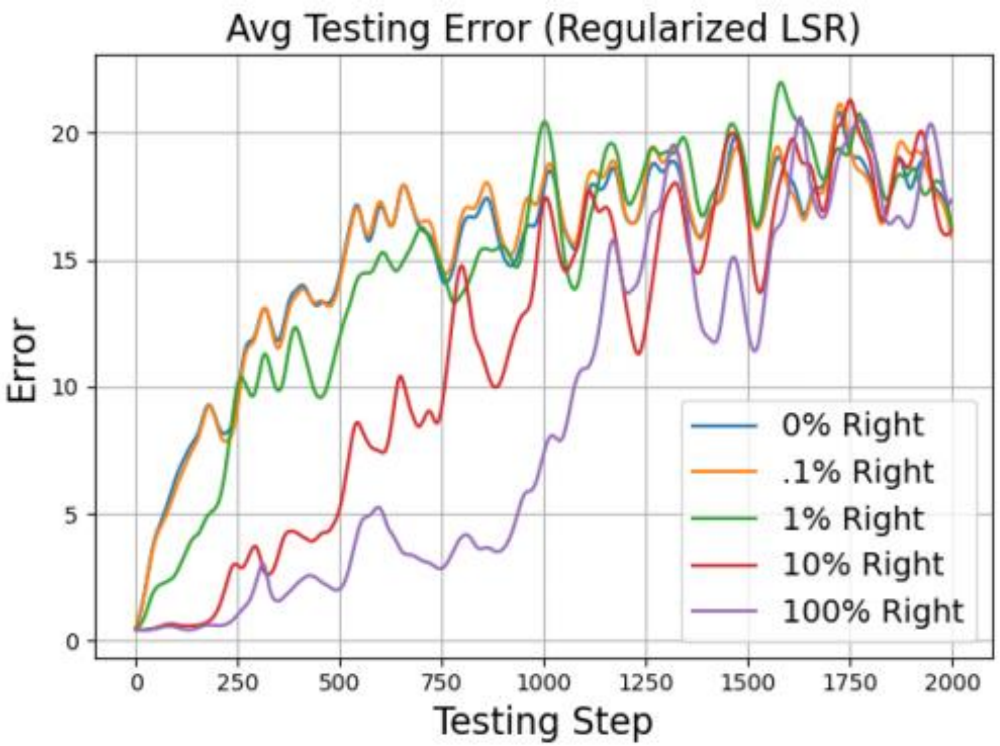




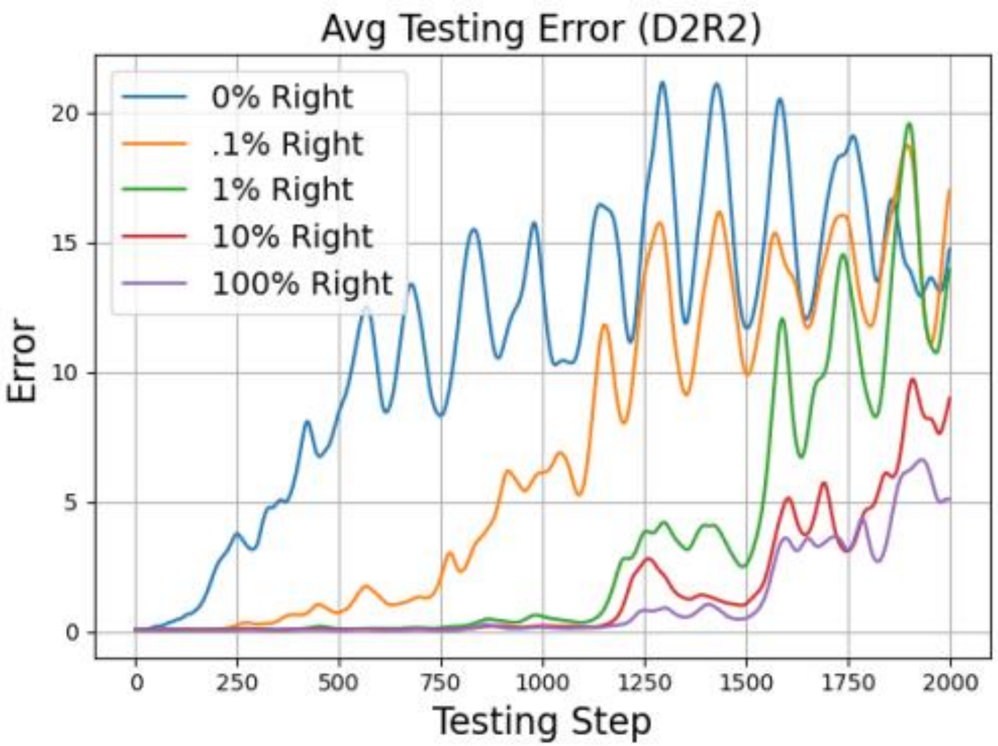
- Let's train a machine learning tool in a changing climate
- Let's start simple to be able to make clear statements → The Lorenz'63 model
- Let's take two different approaches to learn the model from a truth trajectory:
  1. Echo State Networks (Vlachas et al. 2020 and Chattopadhyay et al. 2020)
  2. Domain-Driven Regularized Regression (D2R2; Pyle et al. 2021)
- Let's assume that today's climate is the "left-lobe regime" and that climate change is kicking us into the "two-lobe regime".
- What if we only train from 1%, 2%, 5%… of the training data from the right lobe?

Pyle, Chantry, Palem, Palmer, Dueben, Patel

Step 5: Science and tool developments



Echo State Network



Regression Technique (D2R2)

- The Echo State Network performs horrible unless you provide at least 10% of the data of the right lobe.
- The regression technique needs a very small amount of the right lobe to perform well.

Physics informed machine learning, explainable AI and trustworthy AI need to be explored.

Pyle, Chantry, Palem, Palmer, Dueben, Patel