Machine Learning at ECMWF

Peter Dueben

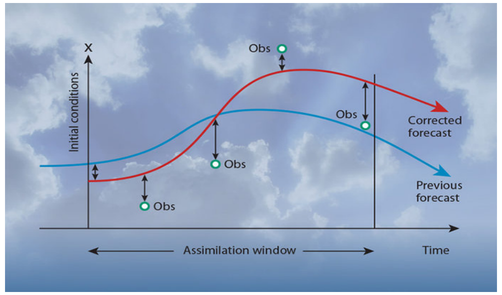
Royal Society University Research Fellow & ECMWF's Coordinator for Machine Learning and AI Activities

The strength of a common goal

The ESIWACE2 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823988.

Machine learning applications across the numerical weather prediction workflow

Observations → **Data-assimilation** → **Numerical weather forecasts** → **Post-processing and dissemination**



Application areas for machine learning are spread over the entire workflow:
weather data monitoring, real-time quality control for observational data, anomaly interpretation, guided quality assignment and decision making, data fusion from different sources, correction of observation error,
learn governing differential equations, non-linear bias correction, bias predictors, learn operational operators, define optical properties of hydrometeors and aerosols, emulate conventional tools improve efficiency,
emulate model components, develop improved parametrisation schemes, build better error models, learn the underlying equations of motion, generate tangent linear or adjoint code from machine learning emulators,
real-time adjustments of forecast products, feature detection, uncertainty quantification, error corrections for seasonal predictions, development of low-complexity models, bespoke products for business opportunities,
and many more…

State-of-play and outline of the talk

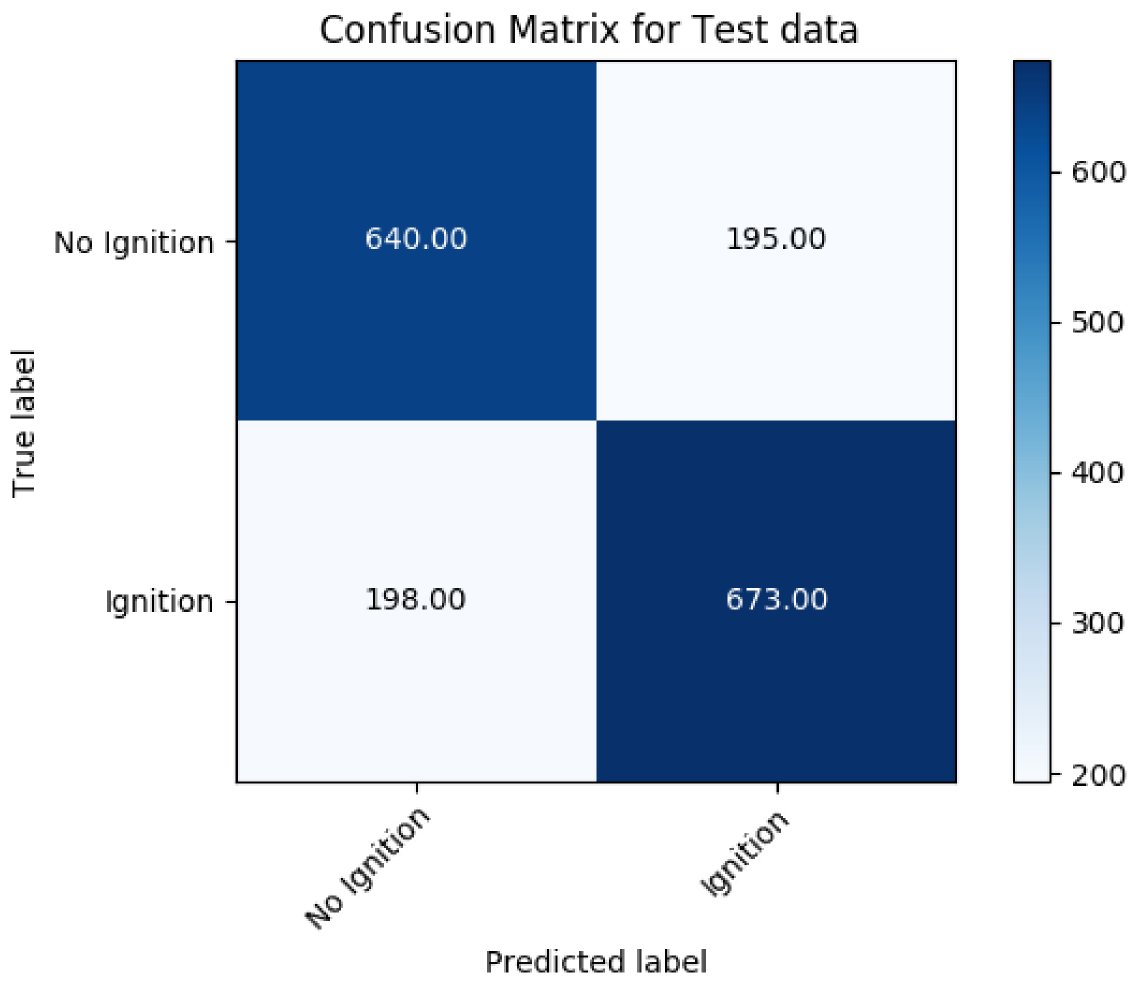
There are many interesting application areas for machine learning to improve weather and climate predictions

We are only at the beginning to explore the potential of machine learning in the different areas

I will present a couple of example applications of machine learning at ECMWF in the following

I will name the main challenges that we are facing when using machine learning today

Observations:
Detect the risk for the ignition of wild fires by lightnings



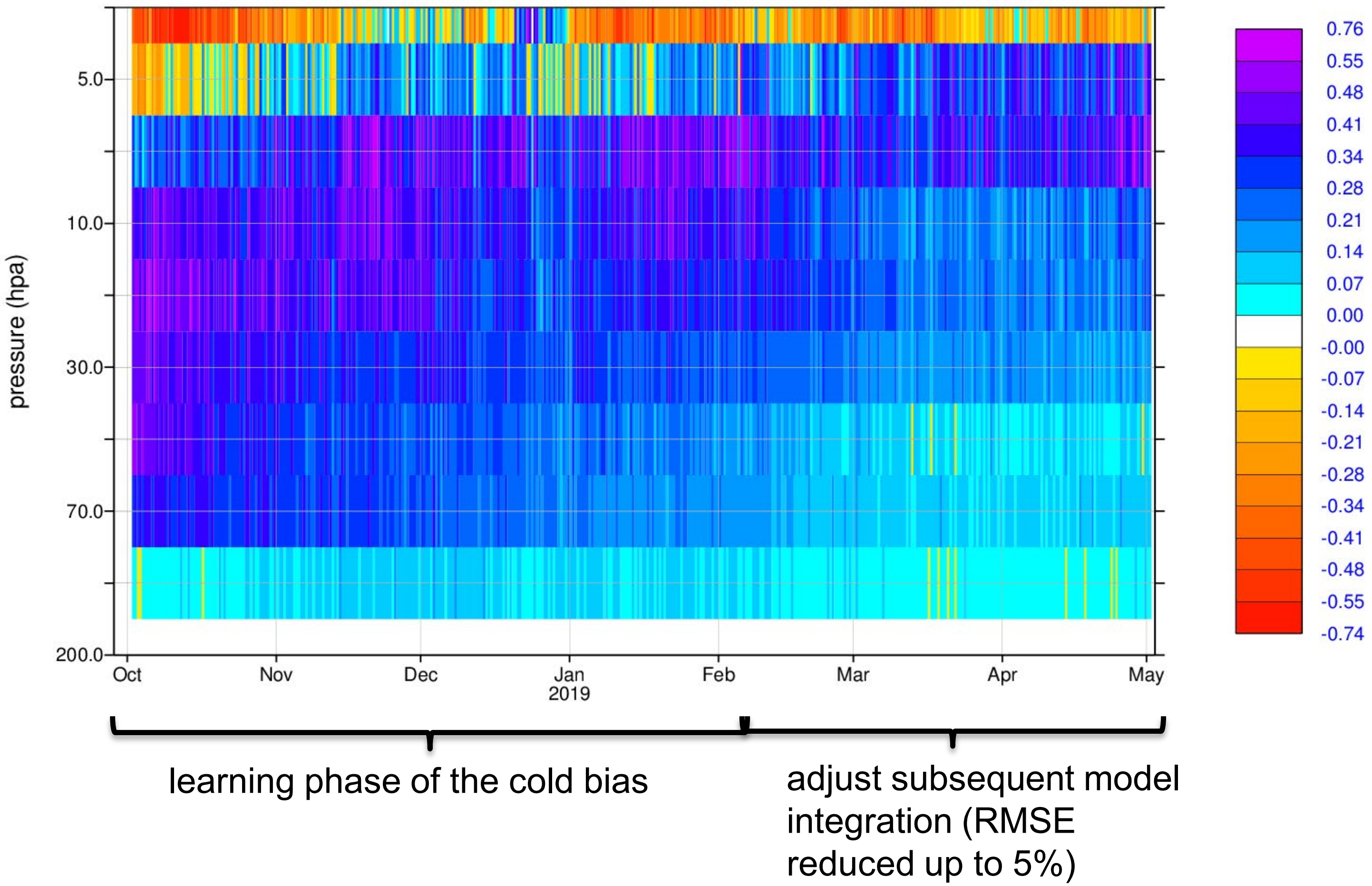
- Observations for 15 variables are used as inputs including soil moisture, 2m temperature, soil type, vegetation cover, relative humidity, and precipitation
- The rate of radiant heat output from the Global Fire Assimilation System (GFAS) was used to generate a "truth"
- 12,000 data points were used for training
- Different machine learning tools (decision trees, random forest and Ada Boost) are used to classify the cases into "ignition" and "no-ignition"
- The best classifier has an accuracy of 77 %

Ruth Coughlan, Francesca Di Giuseppe, Claudia Vitolo and the SMOS-E project

Data assimilation:
Bias-correct the forecast model in 4DVar data assimilation

- Data-assimilation blends observations and the forecast model to generate initial conditions for weather predictions
- This requires estimates of errors of observations and the forecast model
- The new weak-constraint 4D-Var algorithm learns that the model consistently underestimates temperature between 100hPa and 10hPa
- We learn a forcing to correct for the systematic model error
- We still use fairly simple machine learning techniques but we have started to investigate deep learning approaches together with NVIDIA

Mean first-guess departure with respect to GPS-RO temperature retrievals



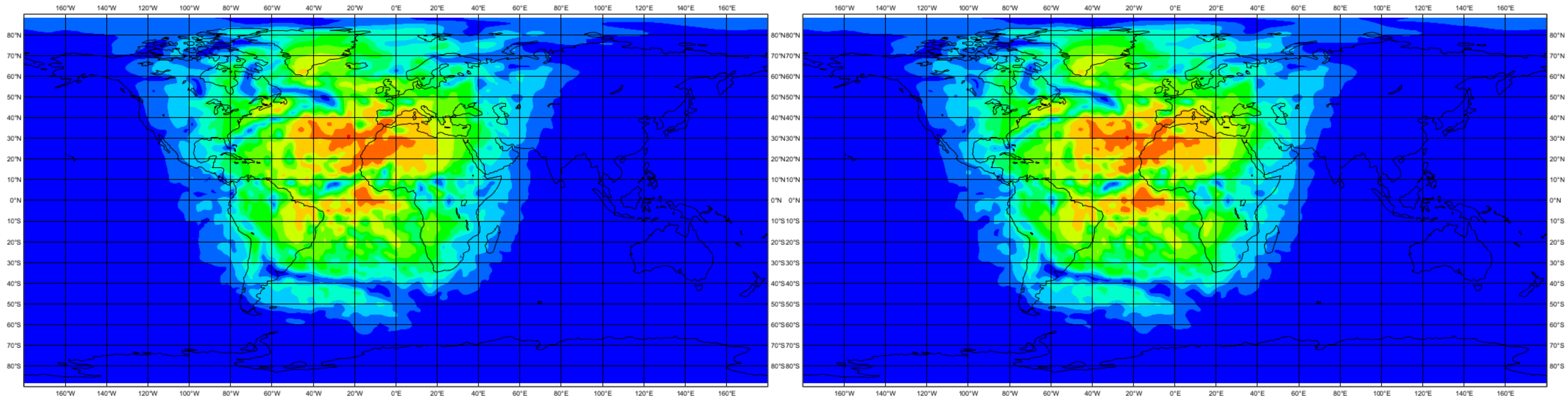
Patrick Laloyaux and Massimo Bonavita

Numerical weather forecasts: To emulate the radiation scheme

- Store input/output data pairs of the radiation schemes
- Use this data to train a neural network
- Replace the radiation scheme by the neural network within the model

This is a very active area of research:
Rasp, Pritchard, Gentine PNAS 2018
Brenowitz and Bretherton GRL 2018
…

Why would you do this?
Neural networks are likely to be much more efficient and portable to heterogenous hardware



Surface downward solar radiation for the original scheme and the neural network emulator (based on a ResNet).

The approach is working and the neural network is ~10 times faster than the original scheme. However, model results are still degraded.

Dueben, Hogan, Bauer @ECMWF and Progsch, Angerer @NVIDIA

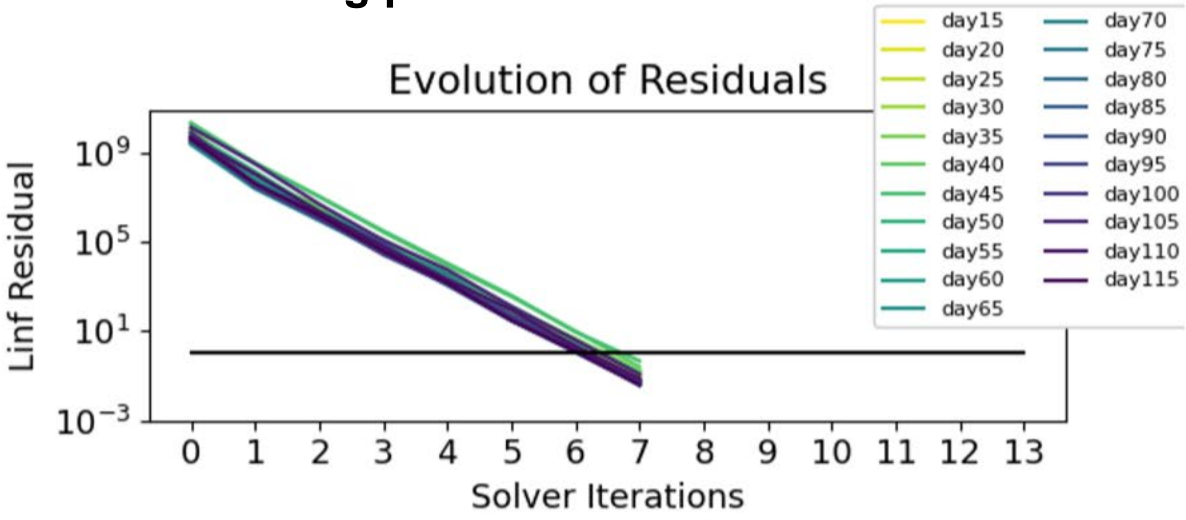
Numerical weather forecasts: To precondition the linear solver

- Linear solvers are important to build efficient semi-implicit time-stepping schemes for atmosphere and ocean models.
- However, the solvers are expensive.
- The solver efficiency depends critically on the preconditioner that is approximating the inverse of a large matrix.

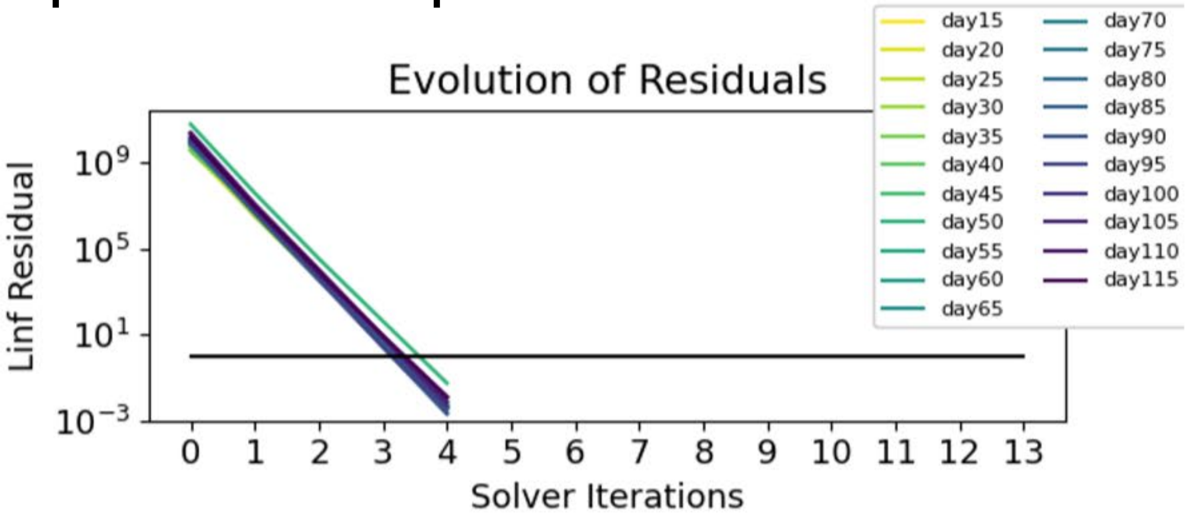
Can we use machine learning for preconditioning, predict the inverse of the matrix and reduce the number of iterations that are required for the solver?

Testbed: A global shallow water model at 5 degree resolution but with real-world topography.
Method: Neural networks that are trained from the model state and the tendencies of full timesteps.

Machine learning preconditioner:



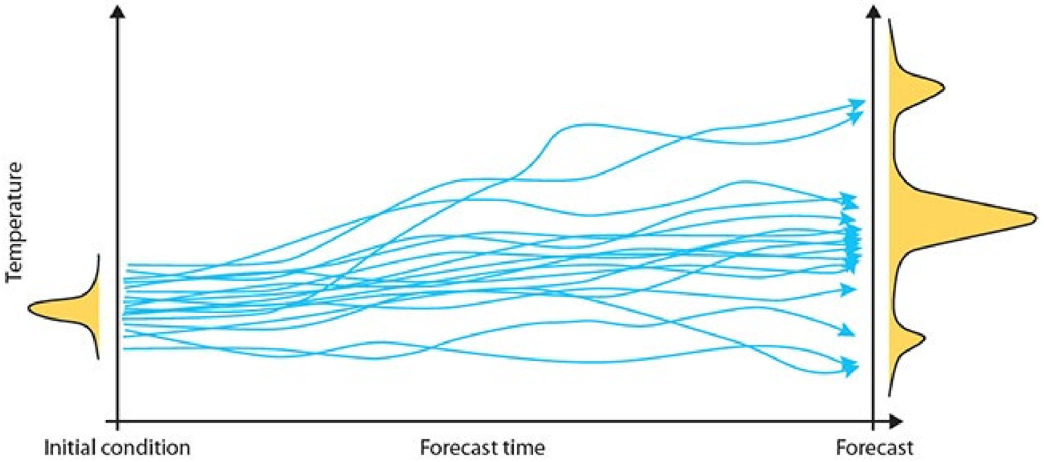
Implicit Richardson preconditioner:

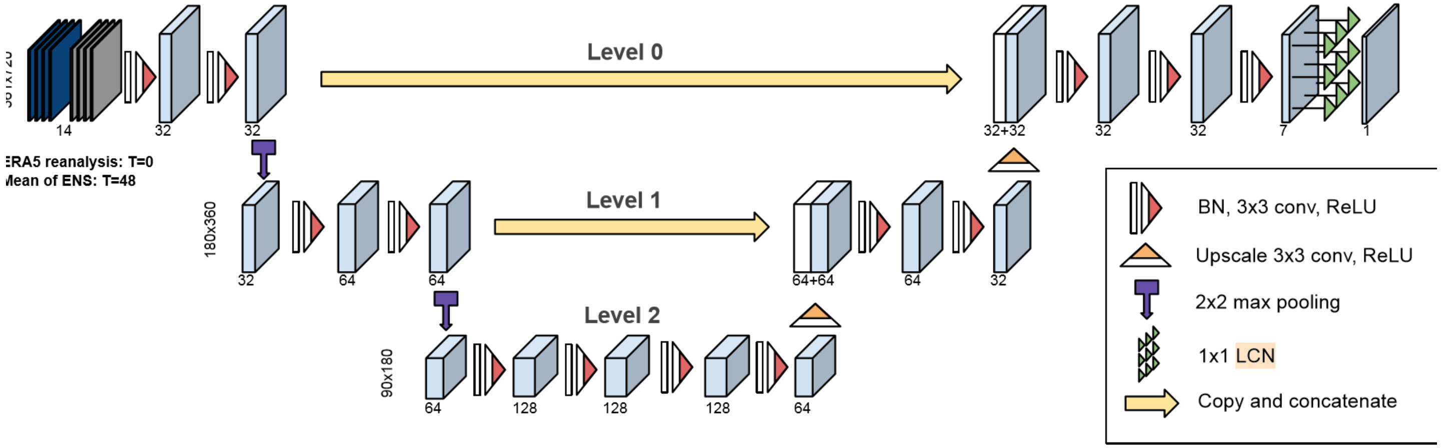


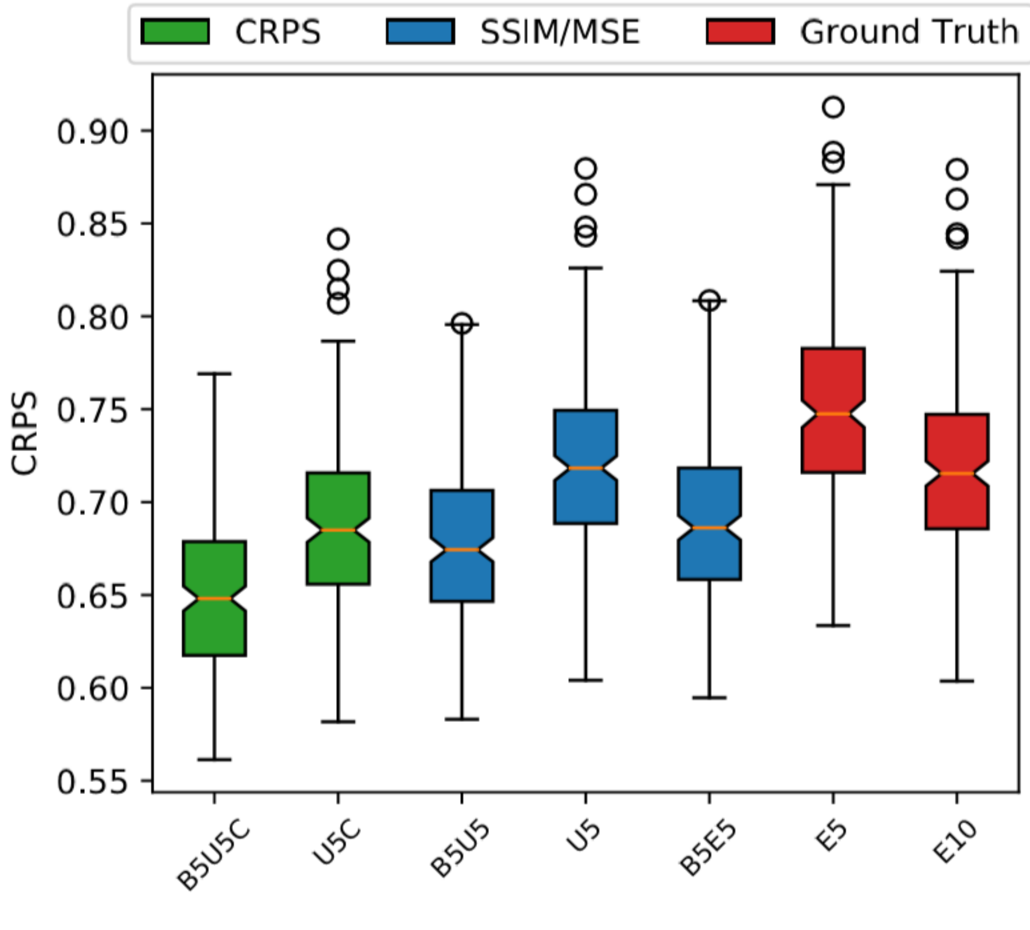
It turns out that the approach (1) is working and cheap, (2) interpretable and (3) easy to implement even if no preconditioner is present.

Ackmann, Dueben, Smolarkieicz and Palmer

Post-processing and dissemination: Improve ensemble predictions







(a) T850

Ensemble predictions are important but expensive.

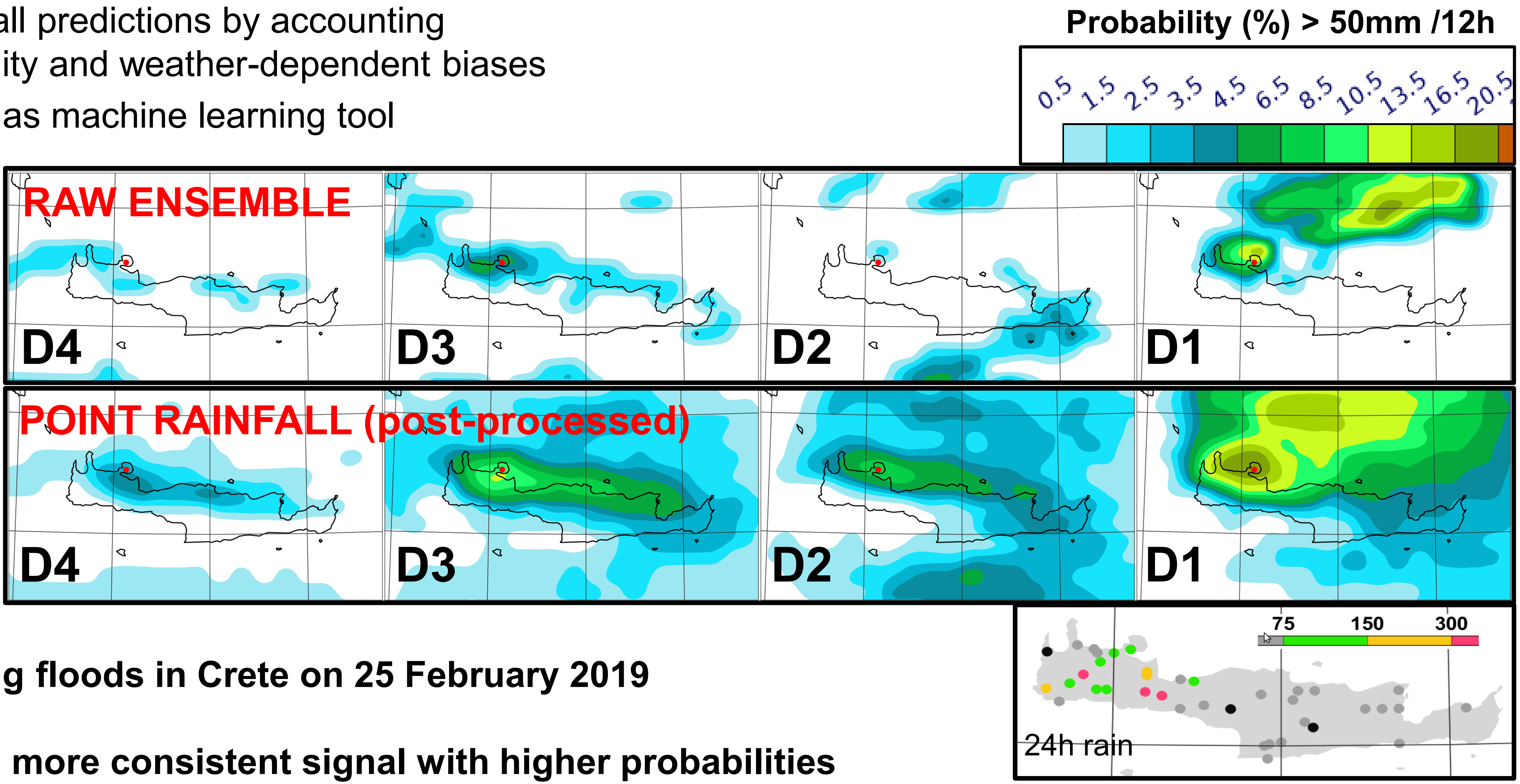
Can we correct ensemble spread calculated from a small number of ensemble members via deep learning?

- Use fields of five ensemble members as inputs.
- Predict ensemble spread of temperature at 850 hPa and Z500 hPa for a 2-day forecast of a full 10 member ensemble forecast.

Grönquist, Yao, Ben-Nun, Dryden, Dueben, Lavarini, Li, Hoefler https://arxiv.org/abs/2005.08748

Post-processing and dissemination: ***ecPoint*** to post-process rainfall predictions

- Use forecast data as inputs
- Train against worldwide rainfall observations
- Improve local rainfall predictions by accounting for sub-grid variability and weather-dependent biases
- Use decision trees as machine learning tool



Example: Devastating floods in Crete on 25 February 2019

Benefits: Earlier and more consistent signal with higher probabilities

Timothy Hewson and Fatima Pillosu

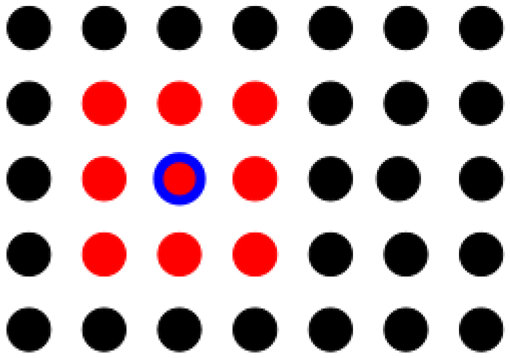
What is the limit?
Can we replace the entire forecast system?

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 210 petabyte of data

A simple test configuration:

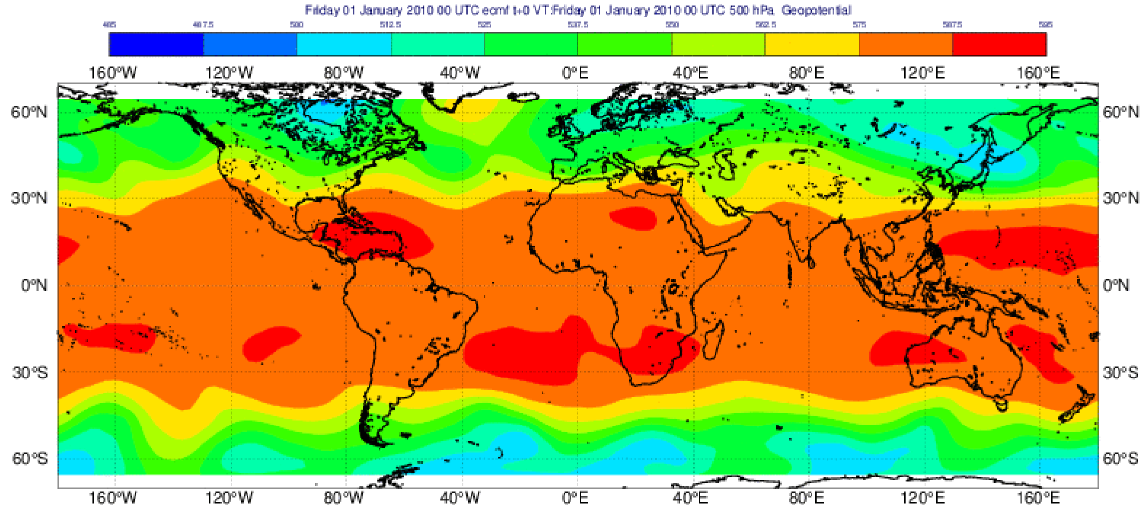
- We retrieve historical data (ERA5) for geopotential at 500 hPa (Z500) for the last decades (>65,000 global data sets)
- We map the global data to a coarse two-dimensional grid (60x31)
- We learn to predict the update of the field from one hour to the next using deep learning
- Once we have learned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

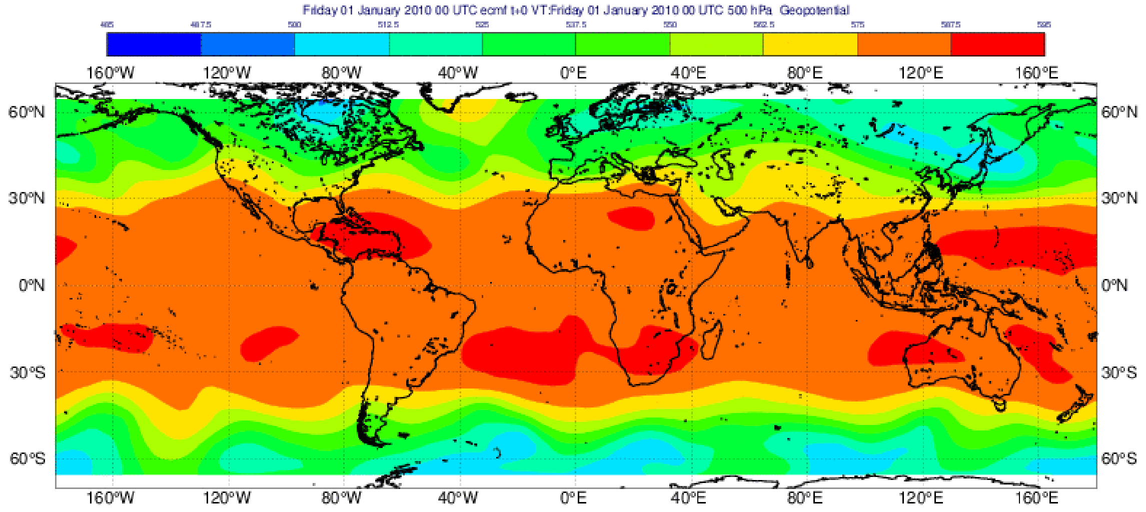
No physical understanding is required!



Dueben and Bauer GMD 2018

What is the limit? Can we replace the entire forecast system?



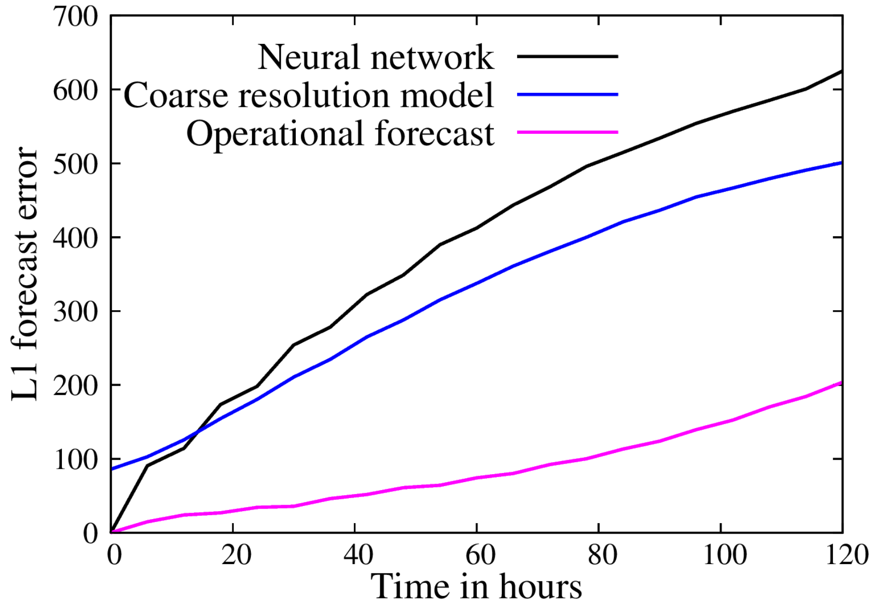


Time evolution of Z500 for historic data and a neural network prediction.
Can you tell which one is the neural network?

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

Unlikely…
The simulations are unstable.
It is unknown how to increase complexity.
There are only ~40 years of data available.

However, there is a lot of progress at the moment:
Scher and Messori GMD 2019; Weyn, Durran, and Caruana JAMES 2019; **…**



Dueben and Bauer GMD 2018

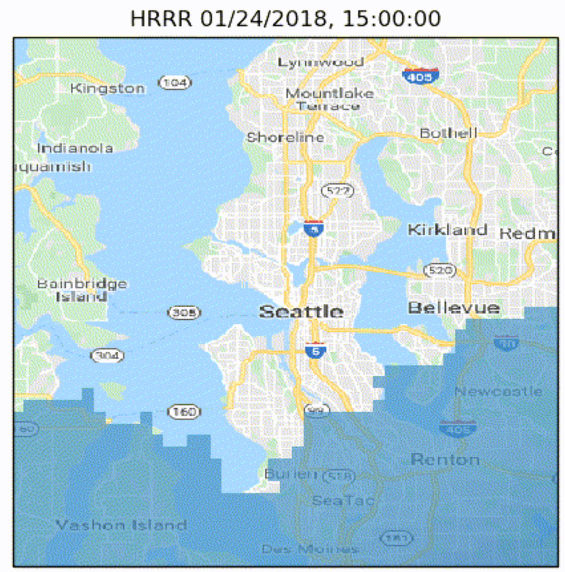
What is the limit? Can we replace the entire forecast system?

However, machine learning models are very promising for **now-casting applications** that provide weather predictions for a couple of hours lead time.

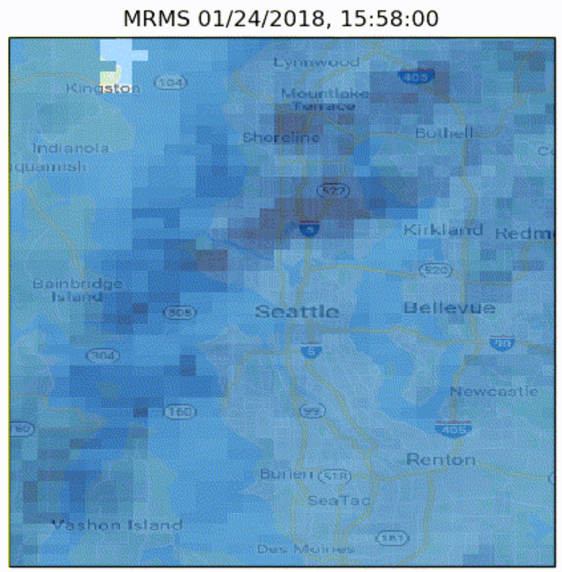
Here…
…conservation is not important as errors have no time to accumulate.
…interactions between weather phenomena are not important (more advection, less physics).
…only local predictions are required → more independent datasets are available for training.

Example: 1-hour predictions of precipitation by Google:

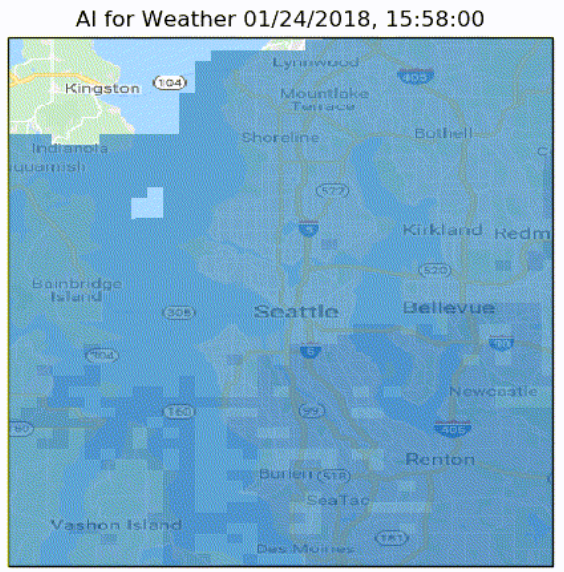
NOAA forecast:



Ground truth:



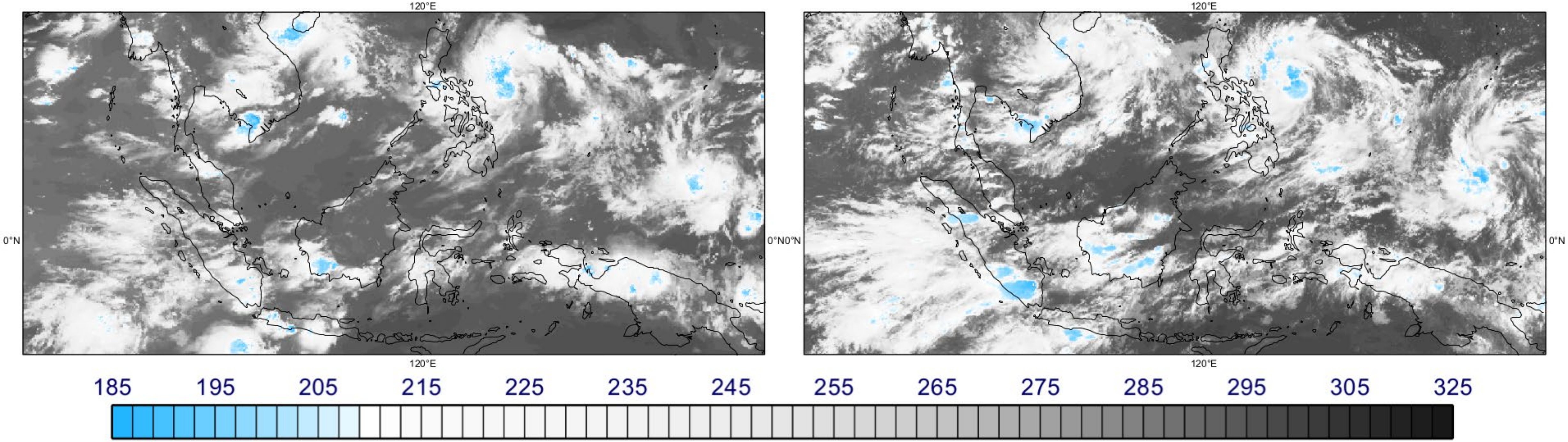
Machine learning solution:



https://ai.googleblog.com/2020/01/using-machine-learning-to-nowcast.html

Why is hard for machine learning tools to compete?

Because our models are astonishing!



Top-of-the-atmosphere cloud brightness temperature [K] for satellite observations and a simulation of the atmosphere with 1.45 km resolution.

Dueben, Wedi, Saarinen and Zeman JSMJ 2020

A weather forecast simulation has O(1,000,000,000) degrees-of-freedom.

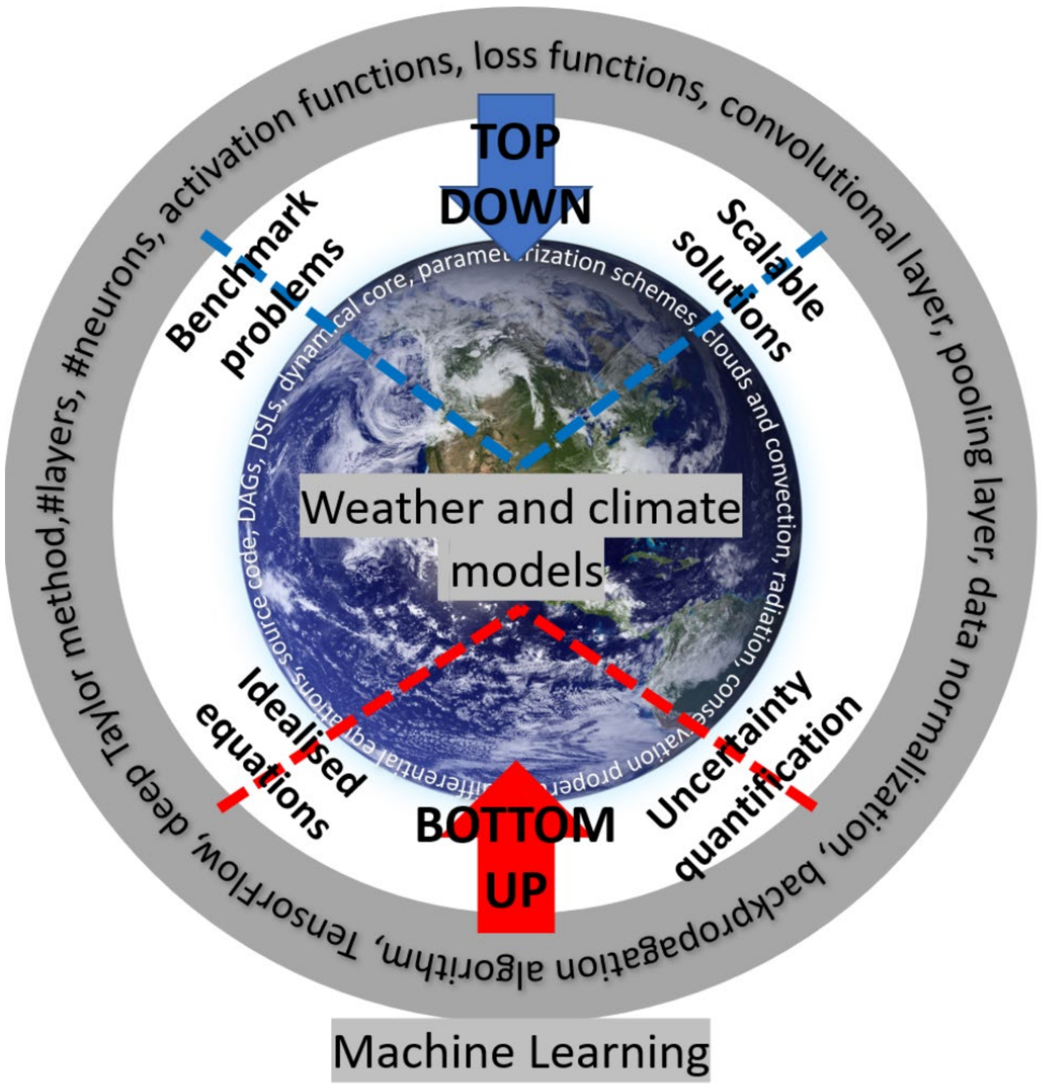
Scientific challenges for machine learning in numerical weather predictions

There is no fundamental reason not to use a black box within weather and climate models but there are unanswered questions.

- Can we use our knowledge about the Earth System to improve machine learning tools?
- Can we diagnose physical knowledge from the machine learning tools?
- Can we remove errors from neural networks and secure conservation laws?
- Can we guarantee reproducibility?
- Can we find the optimal hyper-parameters?
- Can we efficiently scale machine learning tools to high performance computing applications?
- Can we interface machine learning tools with conventional models?
- Can we design good training data (short time steps and high resolution, labelled datasets)?
- Can we explore the full phase space (all weather regimes) during training?

Many scientists are working on these challenges as we speak.

My personal vision of the way forward...



Idealised equations: To study known differential equations to learn how to derive blueprints for neural network architectures.

Uncertainty quantification: To study the representation of variability and the correction of systematic errors for neural networks.

Scalable solutions: To learn how to scale neural networks to millions of inputs for 3D fields on the sphere.

Benchmark problems: To build benchmark problems similar to ImageNet (*see WeatherBench in Rasp, Dueben, Scher, Weyn, Mouatadid and Thureey 2020*)

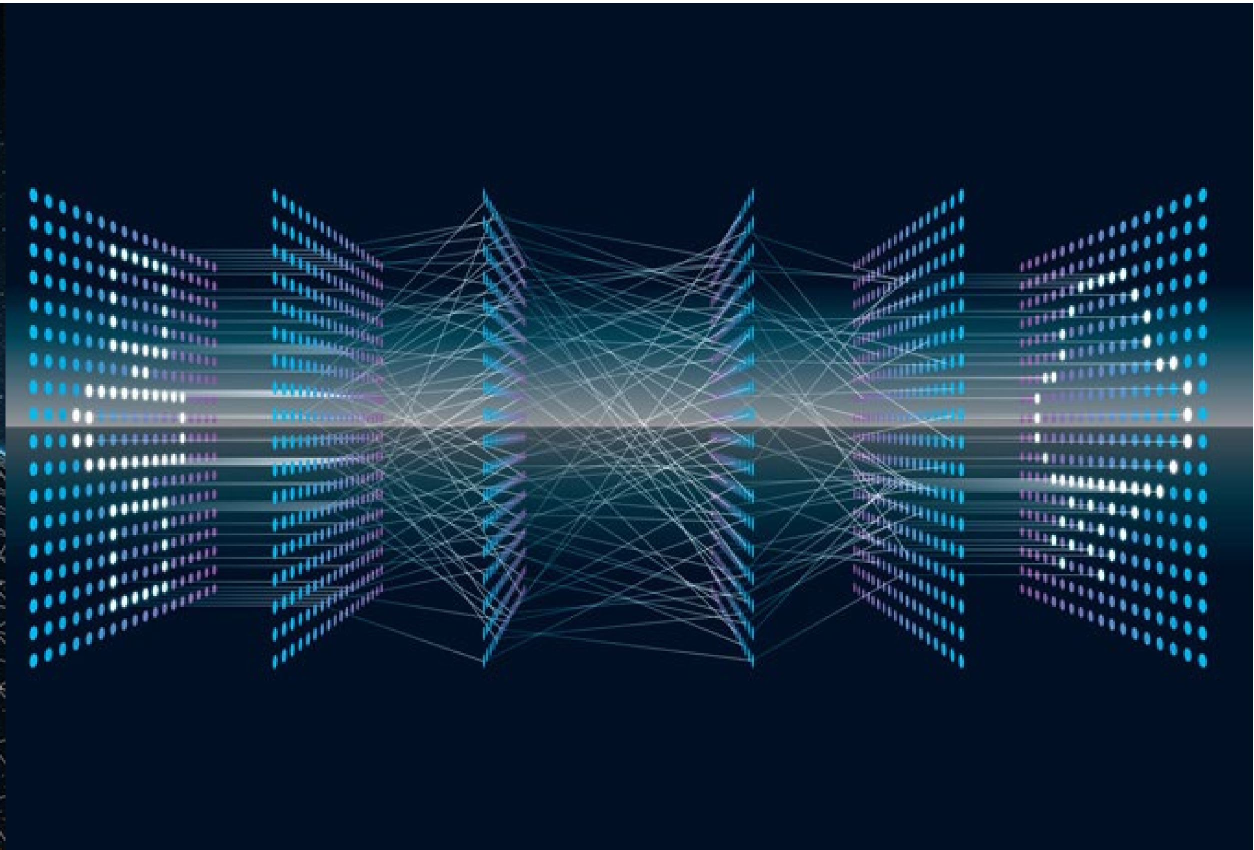
This will require machine learning solutions that are customised to weather and climate models.

We also need to learn how to use ML accelerators for conventional models *Hatfield, Chantry, Dueben, Palmer Best Paper Award PASC2019*

ML workshop and seminar series at ECMWF

ECMWF-ESA Workshop on Machine Learning for Earth System Observation and Prediction at ECMWF 5-8 October 2020. More information is [here](#).

We have also started a special [seminar series](#) on Machine Learning that is broadcasted.



Conclusions

- There are a large number of application areas throughout the prediction workflow in weather and climate modelling for which machine learning can really make a difference.

- The weather and climate community is still at the beginning to explore the potential of machine learning (and in particular deep learning).

- Machine learning could not only be used to improve models, it could also be used to make them more efficient on future supercomputers.

- Machine learning accelerators could be useful to speed-up components of weather and climate models.

- However, there are limitations for the application of black-box solutions within weather and climate models and challenges that need to be addressed.

Many thanks.

ECMWF | EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

The strength of a common goal