Machine Learning for Weather and Climate Prediction

Hannah Christensen

Atmospheric, Oceanic and Planetary Physics, University of Oxford, UK

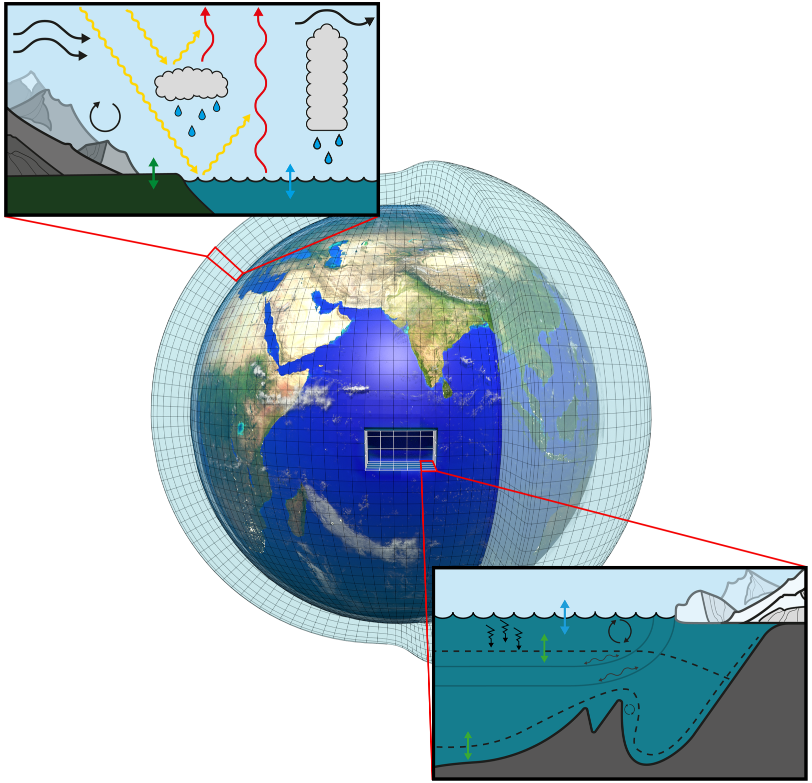
CriticalEarth Meeting, 25-29 April 2022, Utrecht

The problem

What happens next …

- hour?
- day?
- week?
- month?
- season?
- decade?
- century?

The problem



Christensen and Zanna, submitted

The problem

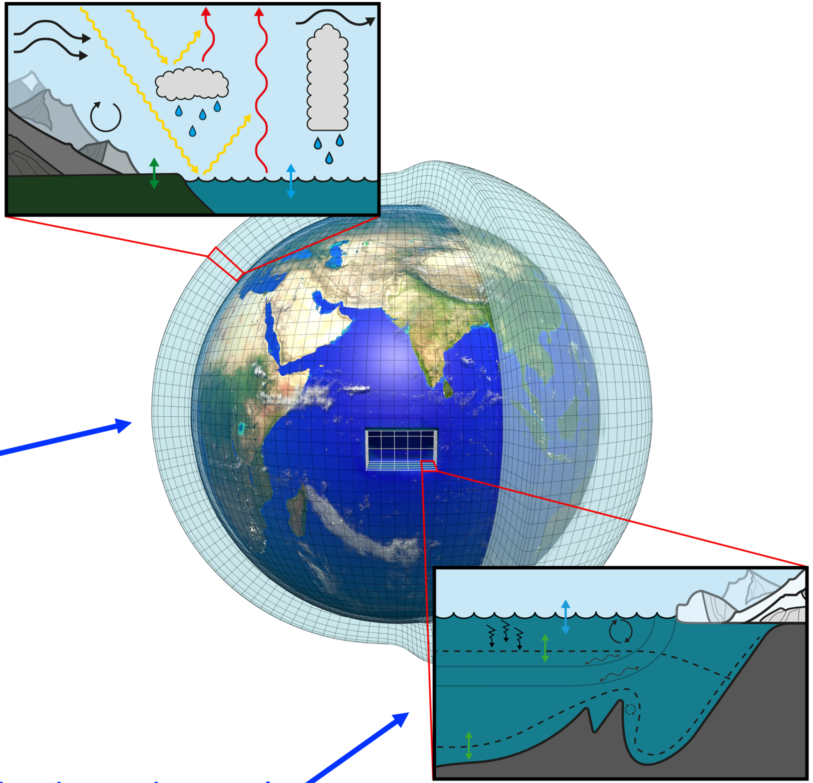
‘parametrisation schemes’

$$\frac{D\mathbf{u}}{Dt} = -\frac{1}{\rho}\nabla p - 2\boldsymbol{\Omega} \times \mathbf{u} - \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) - g\mathbf{k}$$

$$\frac{D\rho}{Dt} + \rho\nabla \cdot \mathbf{u} = 0$$

$$c_p\frac{DT}{Dt} - \frac{1}{\rho}\left(\frac{Dp}{Dt}\right) = Q$$

$$p = R_a T \rho$$

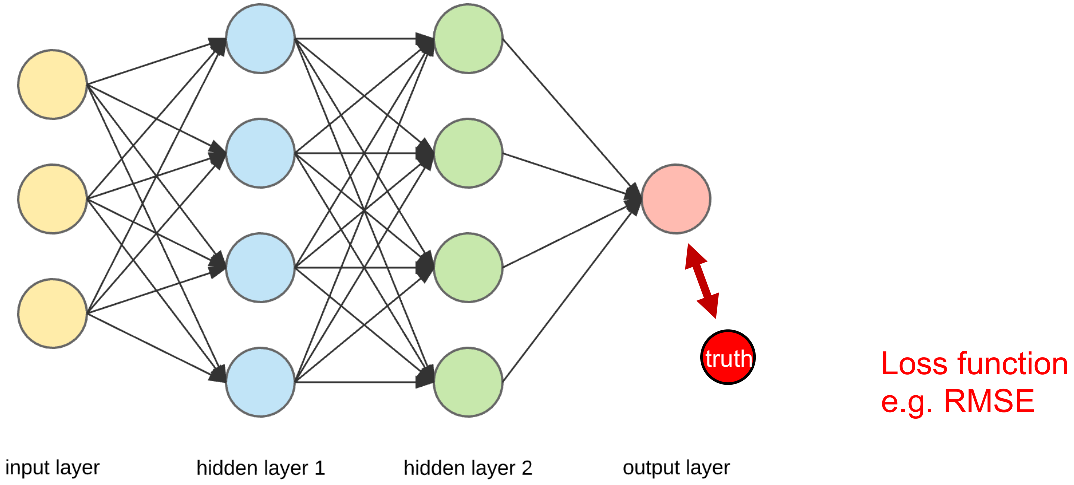


‘parametrisation schemes’

Christensen and Zanna, submitted

How can machine learning help?

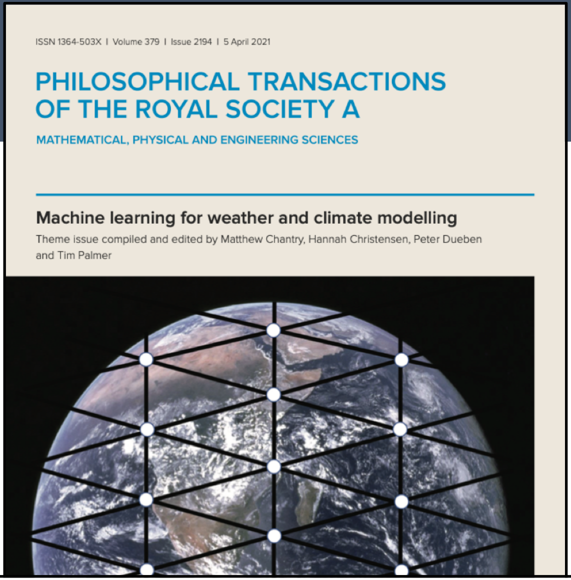
- Tool for modelling general, nonlinear relationships encapsulated in a training dataset



e.g. Neural network

How can machine learning help?

- “Hard” AI
  - Replace model used for prediction by AI
- “Medium” AI
  - Improve accuracy of model forecasts using AI
  - Learn improvements from observations or high-resolution model data
- “Soft” AI
  - Improve efficiency of model forecasts using AI
  - Emulate existing model components



PHILOSOPHICAL TRANSACTIONS A

royalsocietypublishing.org/journal/rsta

Introduction

Check for updates

Cite this article: Chantry M, Christensen H, Dueben P, Palmer T. 2021 Opportunities and challenges for machine learning in weather and climate modelling: hard, medium and soft AI. *Phil. Trans. R. Soc. A* **379**: 20200083. https://doi.org/10.1098/rsta.2020.0083

Accepted: 4 August 2020

Opportunities and challenges for machine learning in weather and climate modelling: hard, medium and soft AI

Matthew Chantry[1], Hannah Christensen[1], Peter Dueben[2] and Tim Palmer[1]

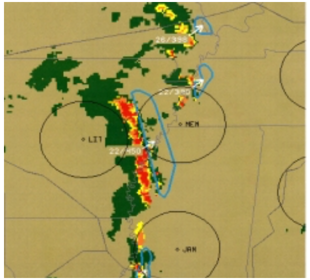
[1]Atmospheric, Oceanic and Planetary Physics, University of Oxford, Oxford, UK

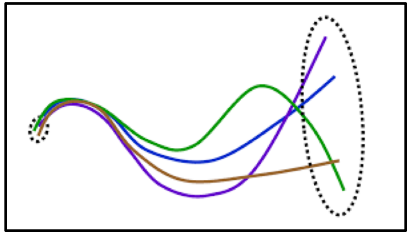
[2]European Centre for Medium Range Weather Forecasts, Reading, UK

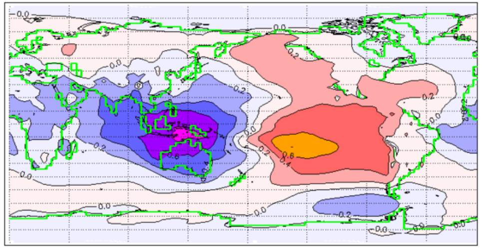
MC, 0000-0002-1132-0961; PD, 0000-0002-4610-3326

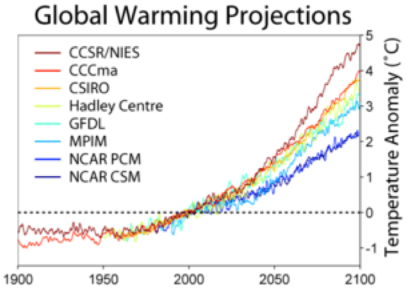
Each timescale has its own challenges

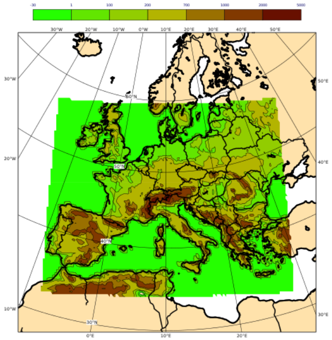
hour — day — week — month — season — decade — century

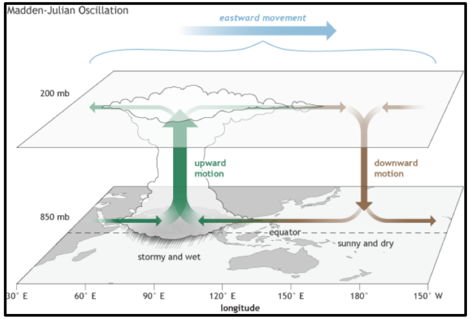


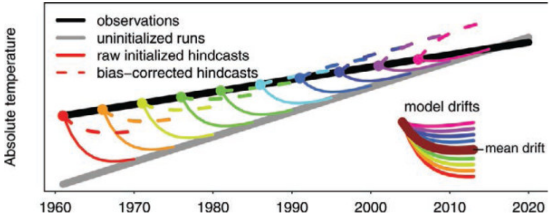






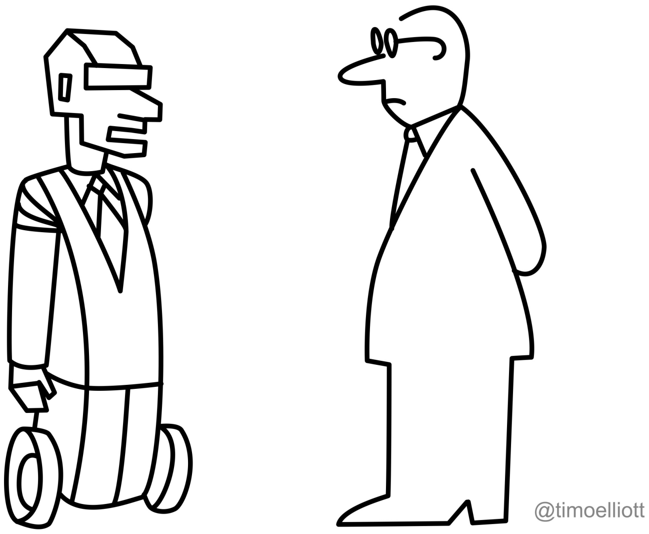






Hard AI

Replace model used for prediction by AI



"The good news is I have discovered inefficiencies. The bad news is that you're one of them."

Nowcasting

Ideal application of Hard AI

- Physical laws not important
- Generally simple approaches used operationally (advection based)
- Ample observational data for training, given short lead times

Ravuri et al, Nature, 2021

Article

Skilful precipitation nowcasting using deep generative models of radar

https://doi.org/10.1038/s41586-021-03854-z

Received: 17 February 2021

Accepted: 27 July 2021

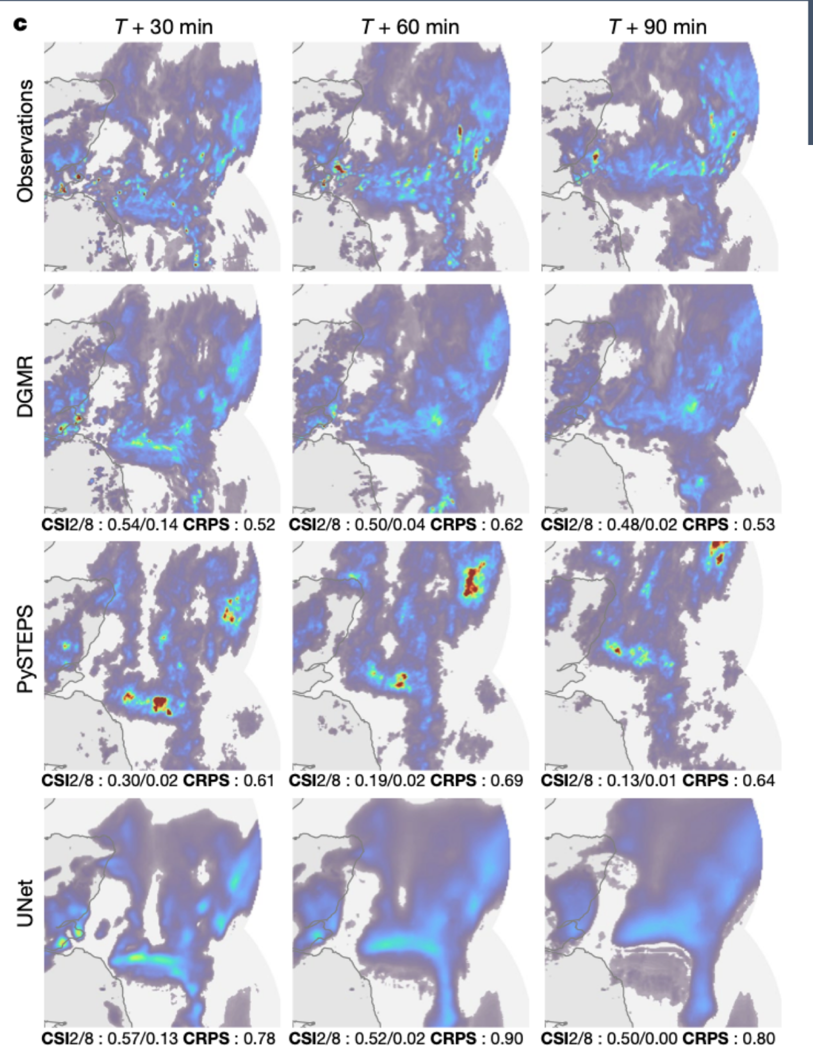
Published online: 29 September 2021

Open access

Check for updates

Suman Ravuri[1,5], Karel Lenc[1,5], Matthew Willson[1,5], Dmitry Kangin[2,3], Remi Lam[1], Piotr Mirowski[1], Megan Fitzsimons[2], Maria Athanassiadou[2], Sheleem Kashem[1], Sam Madge[2], Rachel Prudden[2,3], Amol Mandhane[1], Aidan Clark[1], Andrew Brock[1], Karen Simonyan[1], Raia Hadsell[1], Niall Robinson[2,3], Ellen Clancy[1], Alberto Arribas[2,4] & Shakir Mohamed[1]✉

Precipitation nowcasting, the high-resolution forecasting of precipitation up to two hours ahead, supports the real-world socioeconomic needs of many sectors reliant on weather-dependent decision-making[1,2]. State-of-the-art operational nowcasting methods typically advect precipitation fields with radar-based wind estimates, and struggle to capture important non-linear events such as convective initiations[3,4]. Recently introduced deep learning methods use radar to directly predict future rain



Generative Adversarial Networks (GAN)

Use discriminative model to train a generative model
Originally proposed by Goodfellow et al. (2014)

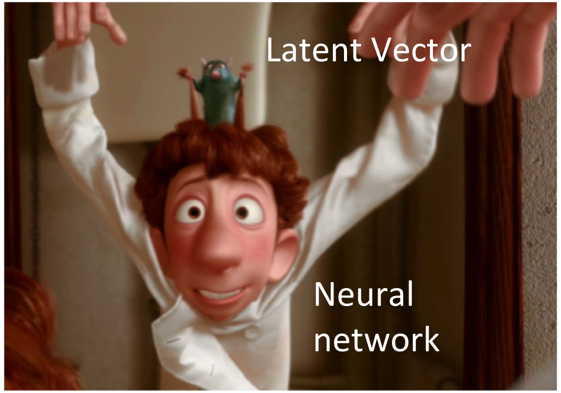
Slide from DJ Gagne

https://upload.wikimedia.org/wikipedia/en/e/e1/Ratatouille-remy-control-linguini.png
http://www.imdb.com/character/ch0009859/mediaviewer/rm988253440

Generative Adversarial Networks (GAN)

Use discriminative model to train a generative model
Originally proposed by Goodfellow et al. (2014)

Generator: Creates synthetic samples to mimic training data, conditioned on latent vector .



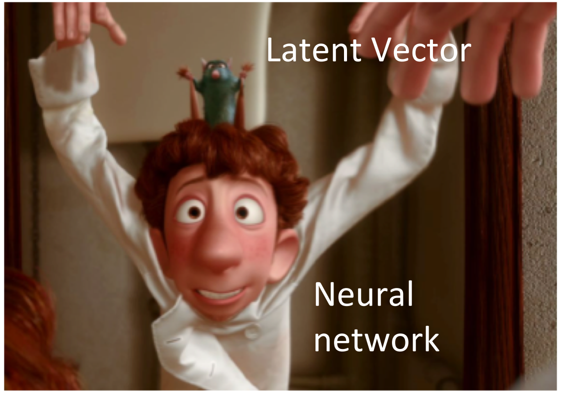
https://upload.wikimedia.org/wikipedia/en/e/e1/Ratatouille-remy-control-linguini.png
http://www.imdb.com/character/ch0009859/mediaviewer/rm988253440

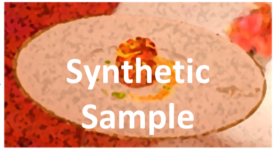
Slide from DJ Gagne

Generative Adversarial Networks (GAN)

Use discriminative model to train a generative model
Originally proposed by Goodfellow et al. (2014)

Generator: Creates synthetic samples to mimic training data, conditioned on latent vector .





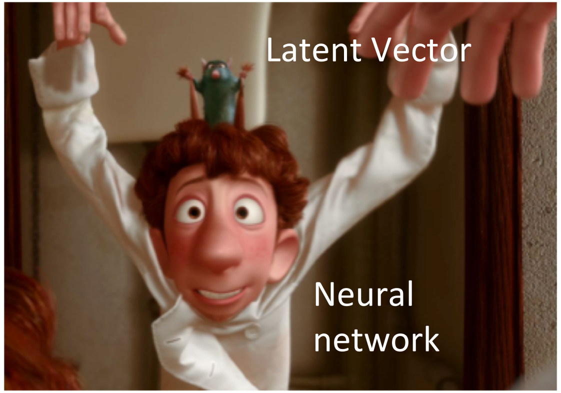
https://upload.wikimedia.org/wikipedia/en/e/e1/Ratatouille-remy-control-linguini.png
http://www.imdb.com/character/ch0009859/mediaviewer/rm988253440

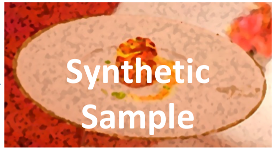
Slide from DJ Gagne

Generative Adversarial Networks (GAN)

Use discriminative model to train a generative model
Originally proposed by Goodfellow et al. (2014)

Generator: Creates synthetic samples to mimic training data, conditioned on latent vector .





Discriminator: Determines which samples are real or synthetic. Adaptive loss

https://upload.wikimedia.org/wikipedia/en/e/e1/Ratatouille-remy-control-linguini.png
http://www.imdb.com/character/ch0009859/mediaviewer/rm988253440

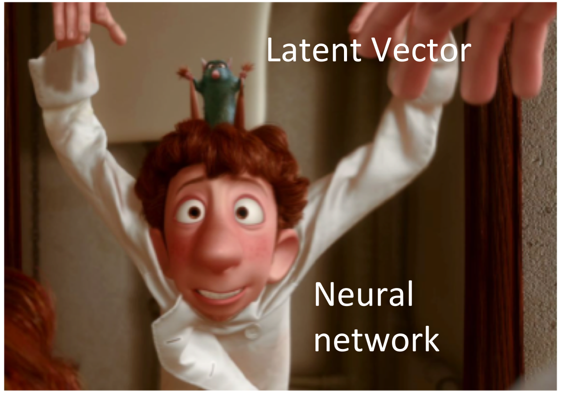
Slide from DJ Gagne

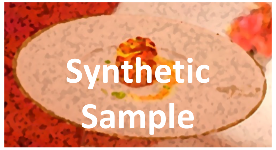
Generative Adversarial Networks (GAN)

Use discriminative model to train a generative model
Originally proposed by Goodfellow et al. (2014)

Generator: Creates synthetic samples to mimic training data, conditioned on latent vector .

Discriminator: Determines which samples are real or synthetic. Adaptive loss





Feedback

https://upload.wikimedia.org/wikipedia/en/e/e1/Ratatouille-remy-control-linguini.png
http://www.imdb.com/character/ch0009859/mediaviewer/rm988253440

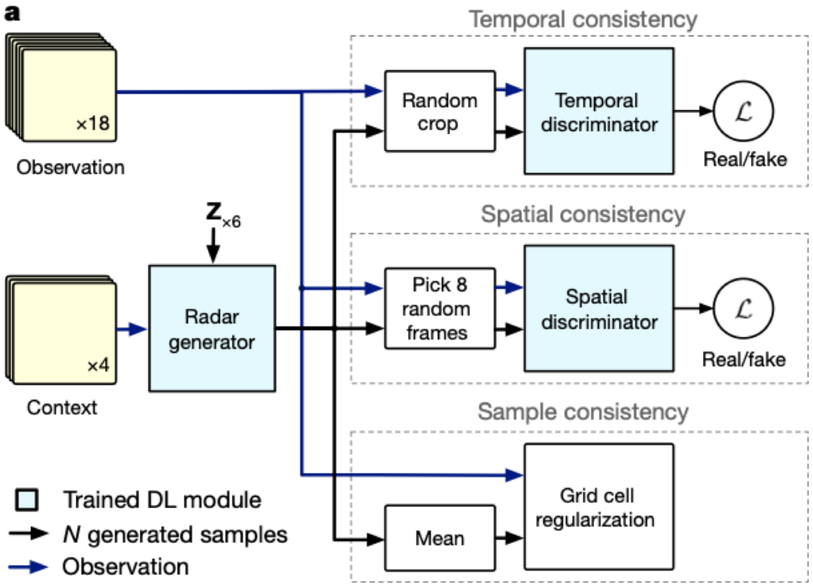
Slide from DJ Gagne

Nowcasting

4 previous UK radar images (5 minutes apart)

Deep generative NN

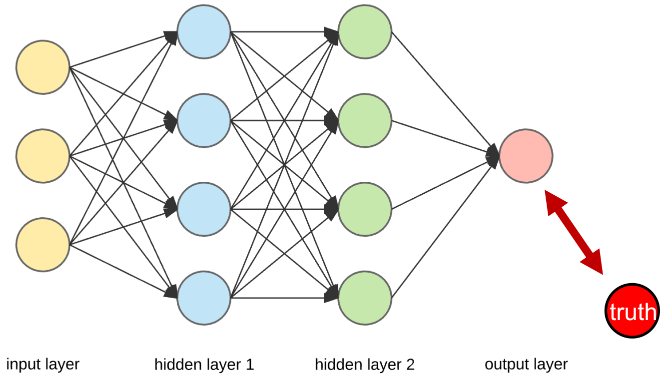
Next 18 UK radar images (5 minutes apart)



Ravuri et al, Nature, 2021

Nowcasting

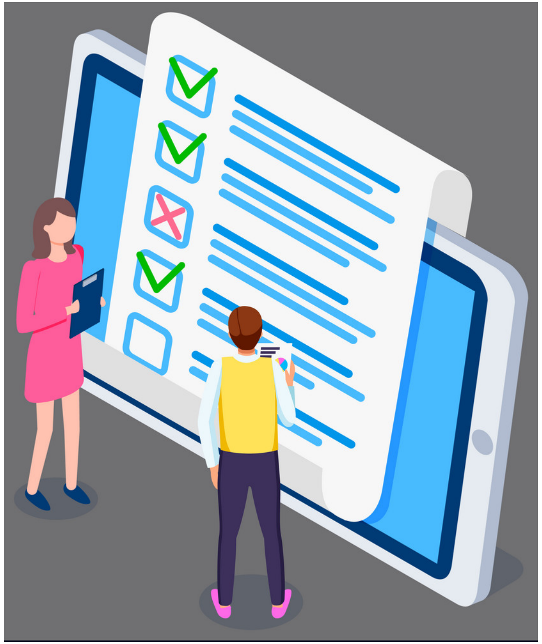
Performance hinges on suitable loss function



> Generative Adversarial framework learns the loss function

Ravuri et al, Nature, 2021

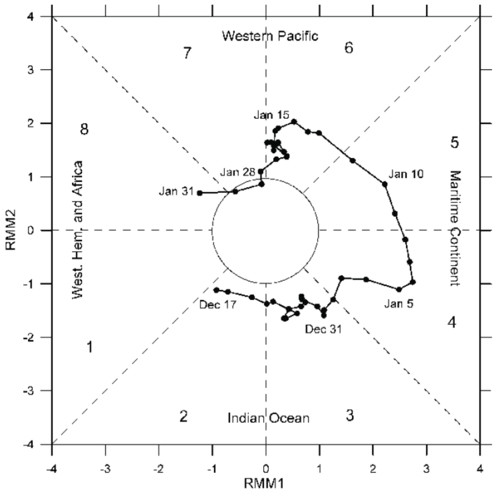
Validation hinges on suitable metrics

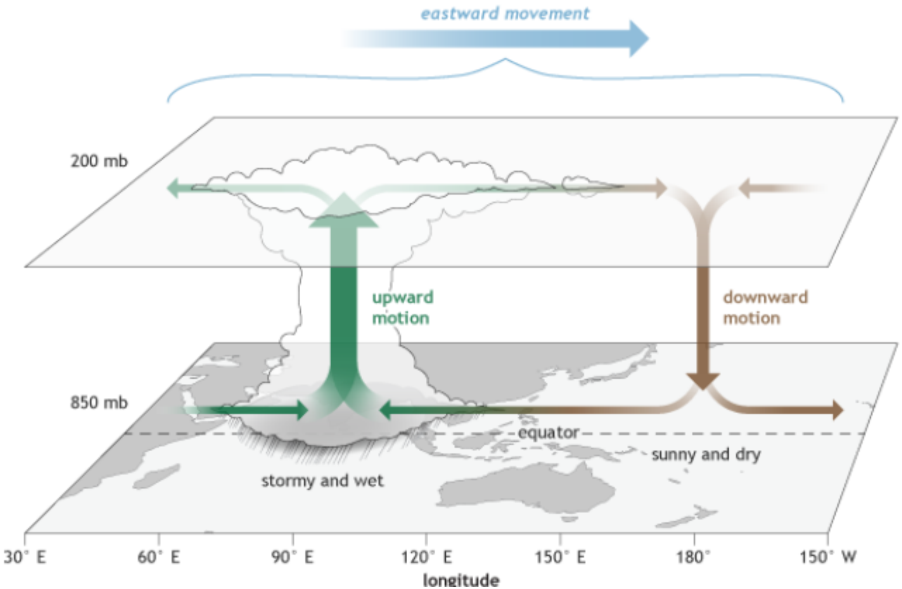


Sub-seasonal prediction: Madden-Julian Oscillation

Ideal application of Hard AI

- Models perform relatively poorly
- Ample observational data for training, given relatively short lead times
- Opportunity to improve understanding of MJO





Sub-seasonal prediction: Madden-Julian Oscillation

- MJO shows high intermittency – poorly understood why this is

- Operational probabilistic forecasts poorly calibrated
  - Forecast spread varies from day to day, but observed error distribution is independent of predicted spread

➢ Science goals:
  - Build well-calibrated probabilistic ML forecast model for MJO
  - Use it to understand sources of predictability for the MJO

Antoine Delaunay

Delaunay and Christensen, GRL, submitted

Convolutional Neural Network

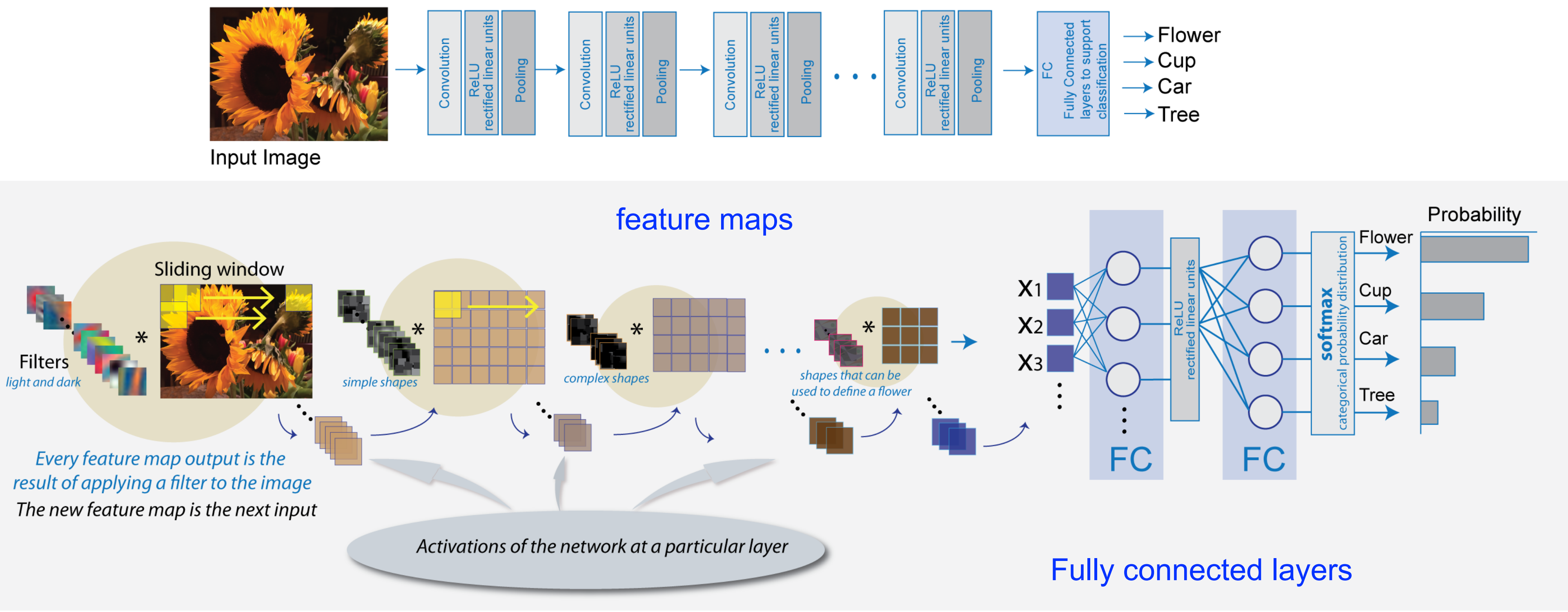
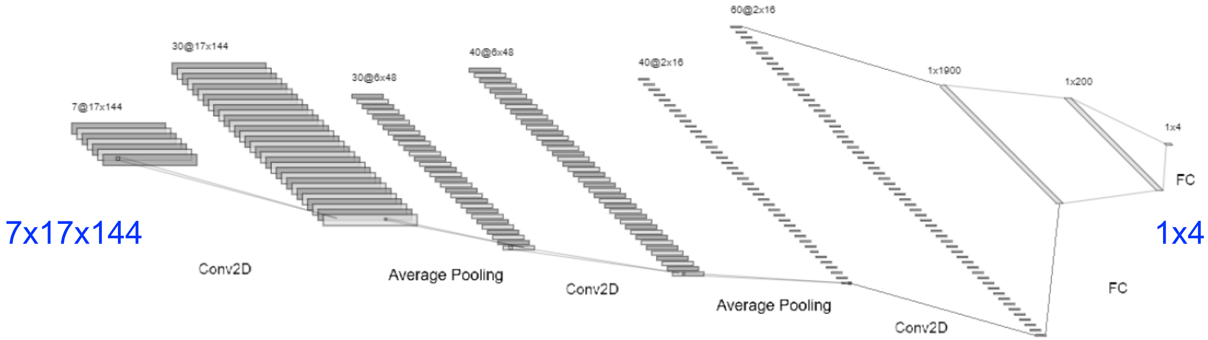


Image: Matlab documentation

Sub-seasonal prediction: Madden-Julian Oscillation

Observed maps of 7 variables, 20S-20N, 0-360E

ERA5 data, 1979-present



μ, σ^2
for each PC

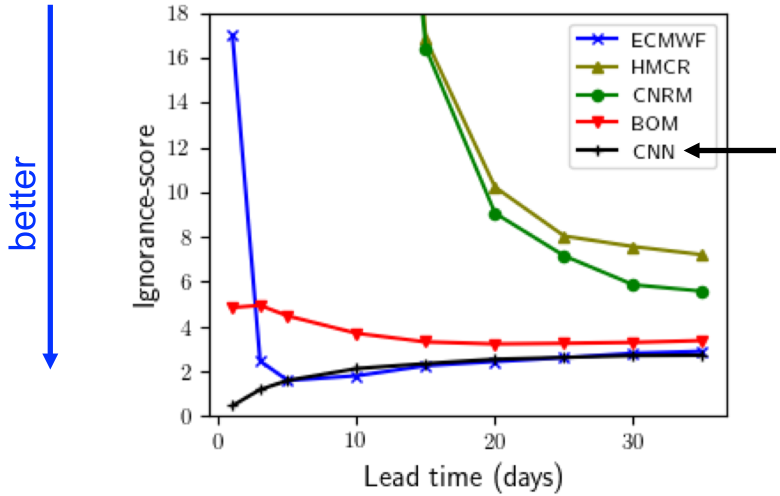
- aleatoric uncertainty

Log-likelihood loss function

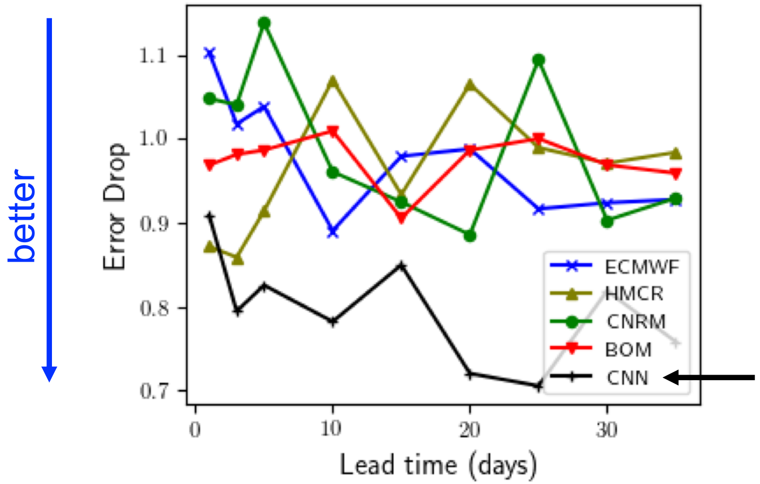
$$L = -\log(\, p(X)\,)$$

Delaunay and Christensen, GRL, submitted

Sub-seasonal prediction: Madden-Julian Oscillation



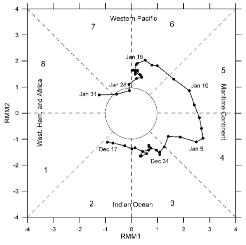
Measures probabilistic skill of forecasts



Measures ability of forecasts to sort days according to their uncertainty

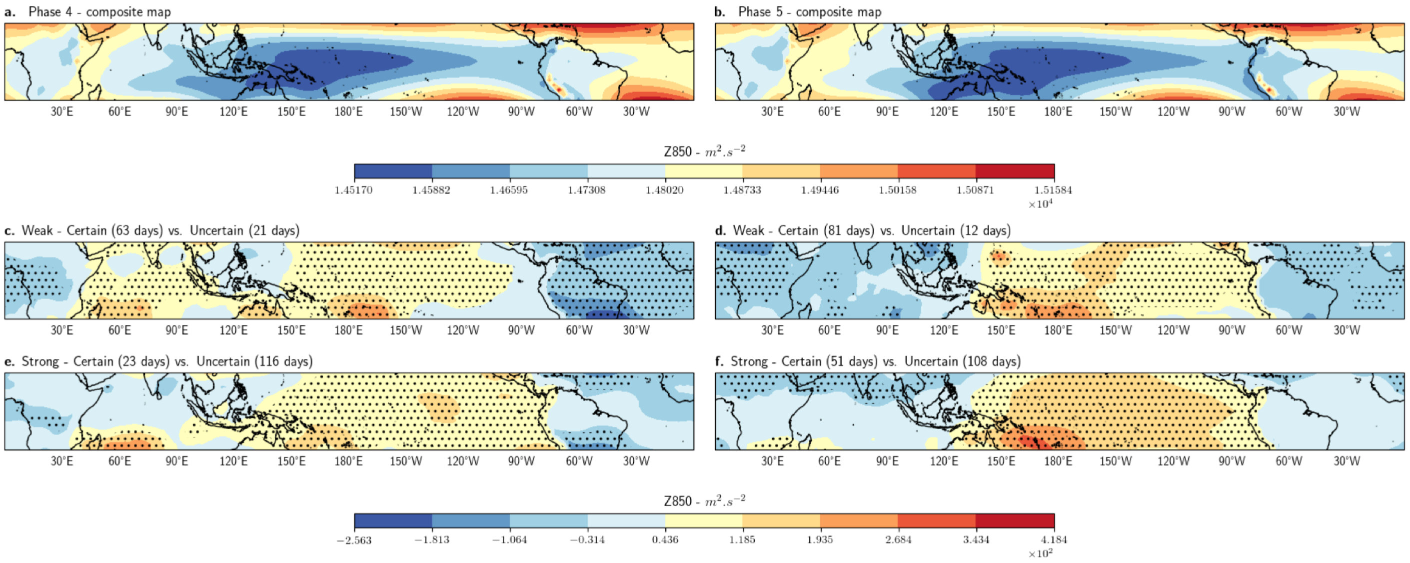
Delaunay and Christensen, GRL, submitted

Sub-seasonal prediction: Madden-Julian Oscillation



What makes an event predictable or not ?

Z850 Composites anomalies for certain & uncertain events
Initial phase 4 (left) & 5 (right)



Delaunay and Christensen, GRL, submitted

Hard AI

Benefits

Outperform traditional models

Gain new understanding about phenomena

Challenges

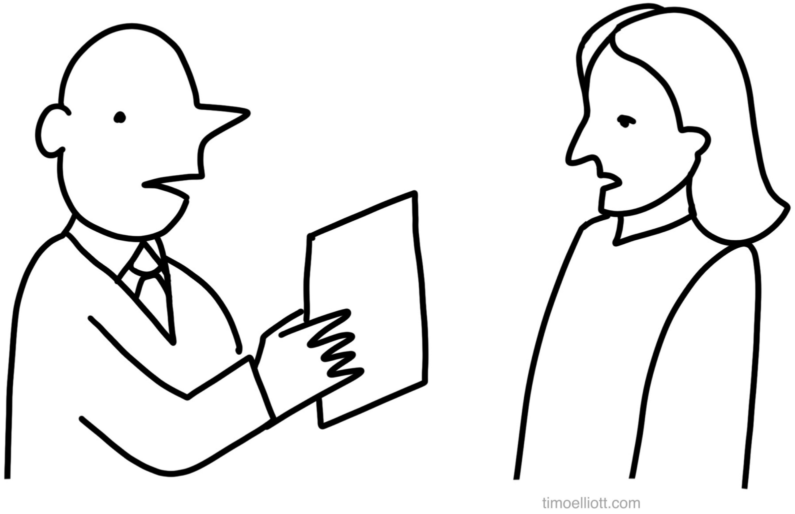
Harder to include (complex) physical constraints

Needs sufficient training data from observations
- restricts timescales that can be considered
- (transfer learning can ameliorate this)

Extrapolation difficult

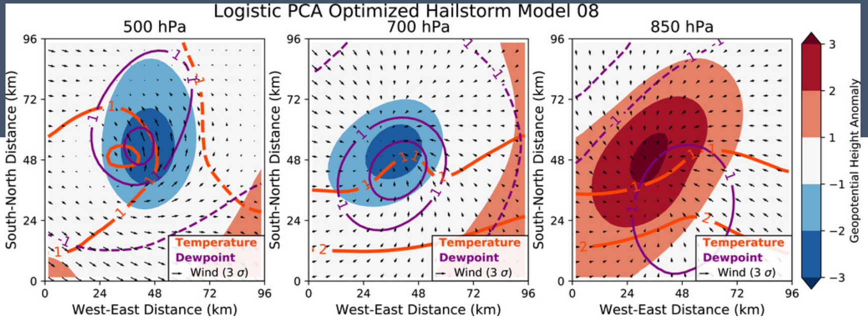
Medium AI

Improve accuracy of
model forecasts using AI



"This article claims computers can now see better than people!"
"Hold on, let me get my glasses so I can read it..."

Postprocessing

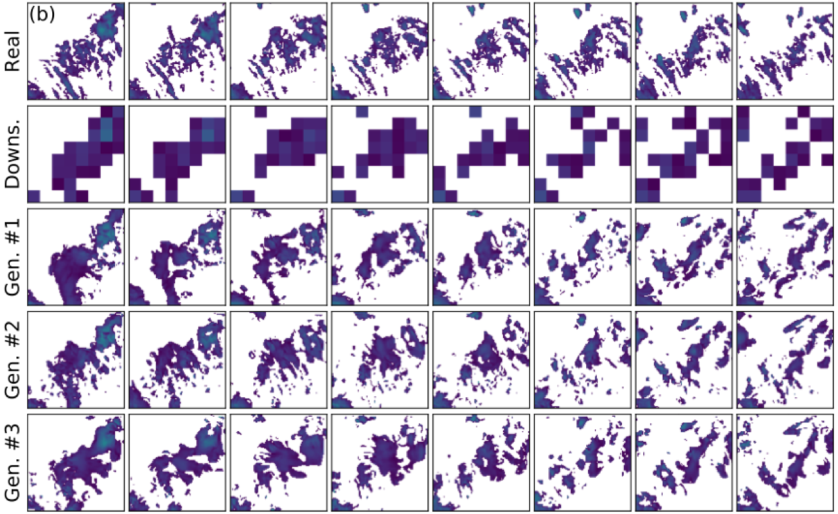


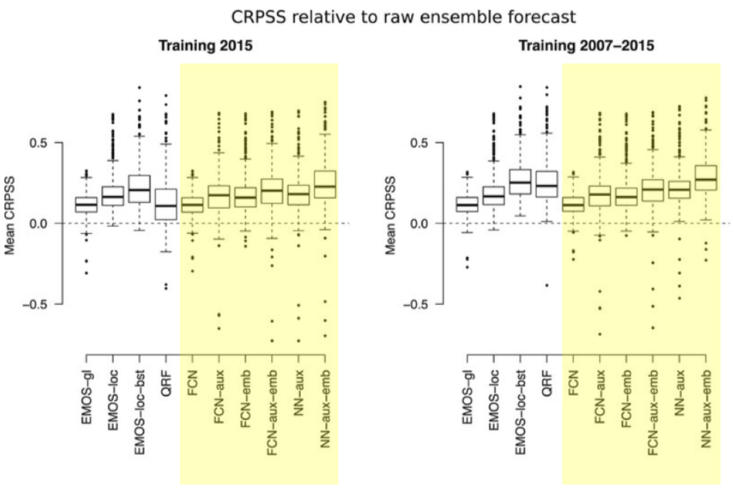
e.g. understanding ML hail forecasts (CNN): Gagne et al, 2019

Ideal application of Medium AI

- Uncontroversial – long history
- Build on past data-driven approaches
- Add value to forecasts
- User oriented
- Possibly develop new understanding

e.g. downscaling to add detail (GAN): Leinonen et al, 2021





< e.g. improving calibration ensemble forecasts (shallow NN): Rasp and Lerch, 2018

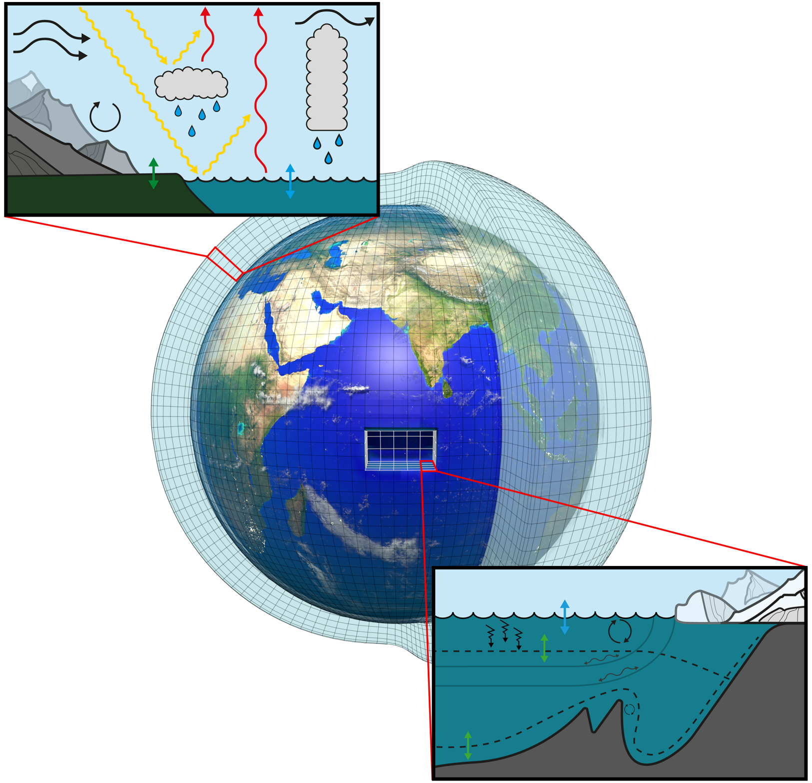
Improving model parametrisations

Replace parametrisation schemes with a ML model

- Makes use of physical laws encoded in dynamical core
- Large databases of high-resolution simulations for training
- Many parametrisations are complex and for poorly understood processes

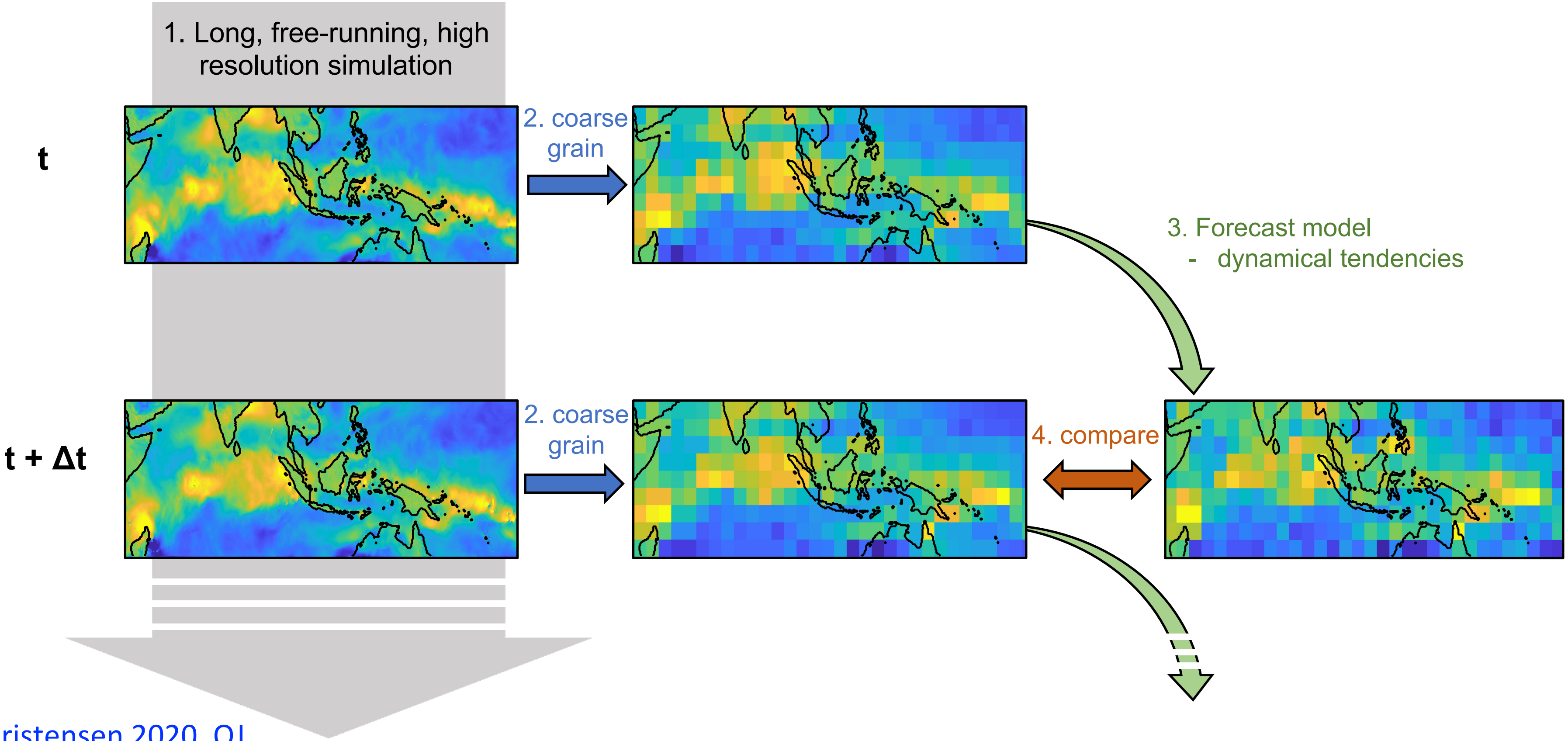
BUT

- Optimal coupling between parametrisations and dynamics poorly understood
- Many challenges: conservation, stability, extrapolation, …



Christensen and Zanna, submitted

Improving model parametrisations

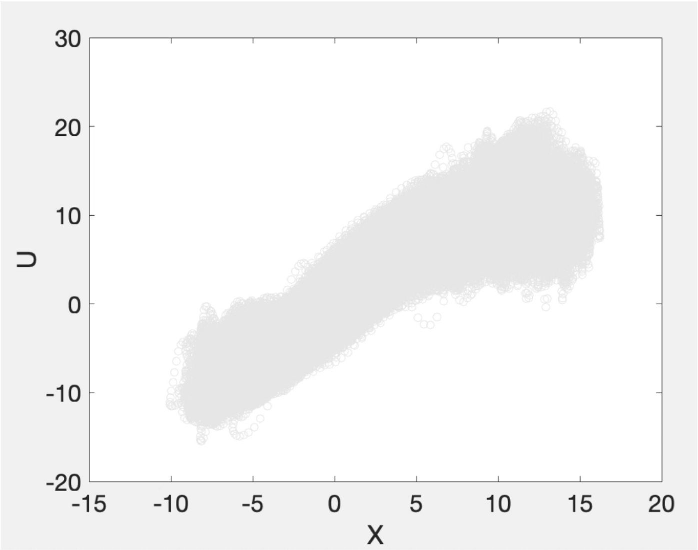


Christensen 2020, QJ

Improving *stochastic* parametrisations

- Stochastic parametrisations
  - Account for uncertainty in parametrisation process
  - Use stochasticity to represent one possible evolution of the subgrid

- Sub-grid tendencies must be correlated in space and time
  - Represents real structure of term to be parametrised

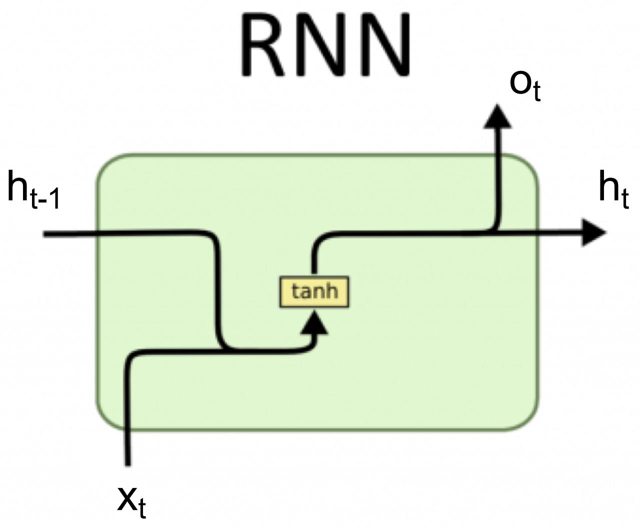


➢ Science goals:
  - Use ML to learn improved stochastic parametrisations

Recurrent Neural Network

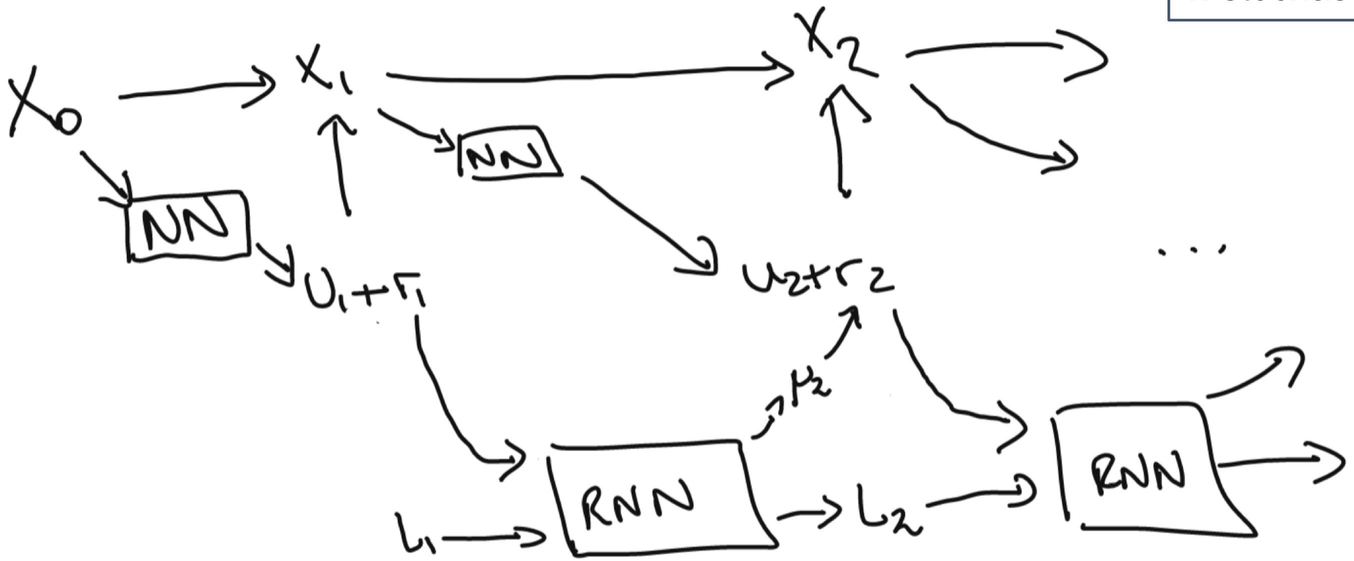
- RNNs are optimal for timeseries problems
  - Trained on how well they reproduce a timeseries
  - Track a hidden state in time

- Training involves “unrolling” the RNN
  - Can lead to issues with vanishing gradients resulting in short term memory issues
  - Solution came in form of LSTMs (Long Short-Term Memory) and GRU (Gated Recurrent Units)



Recurrent Neural Networks generate correlated timeseries

X: state
U: deterministic tendency
r: stochastic correction



Parthipan, Christensen, Hoskings, Wischik, JAMES, submitted

Raghul Parthipan

Medium AI

Benefits

Outperform traditional models

Easier to include (complex) physical constraints

Generally access sufficient model data

Challenges

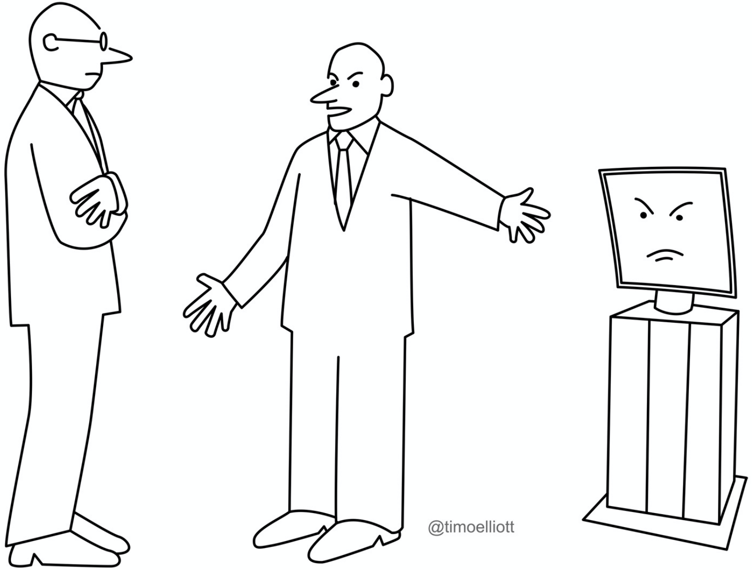
Harder to gain new understanding about phenomena

Extrapolation difficult

Limitations with available observational data

Soft AI

Improve efficiency of model forecasts using AI



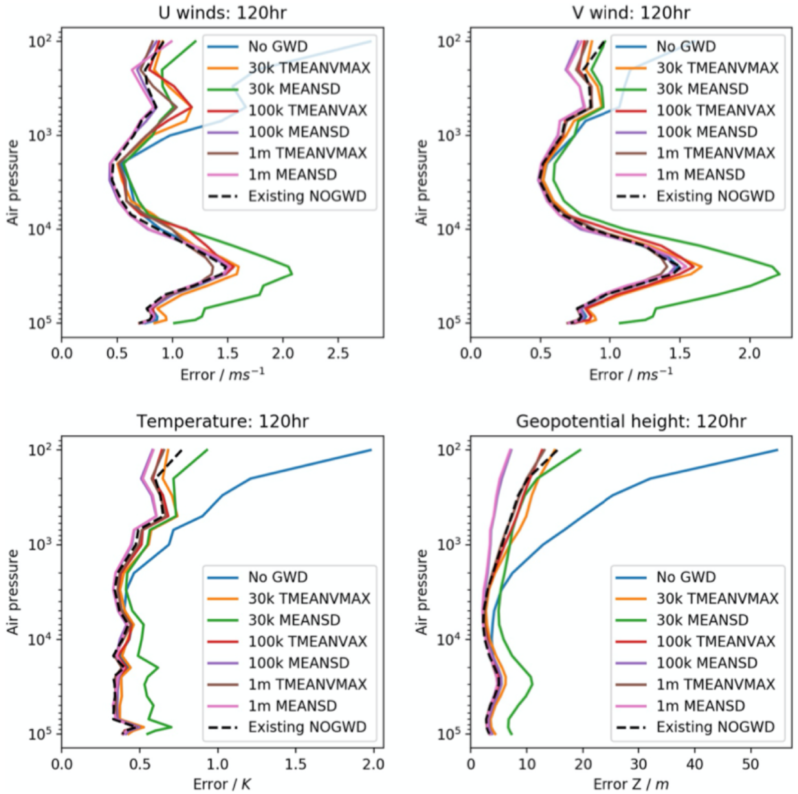
His decisions aren't any better than yours — but they're WAY faster...

ML to emulate parametrisations

ML parametrisations (can be) **cheaper** than conventional approaches

So we can …
- increase resolution
- include more complexity
- run models for longer
- run more ensemble members …

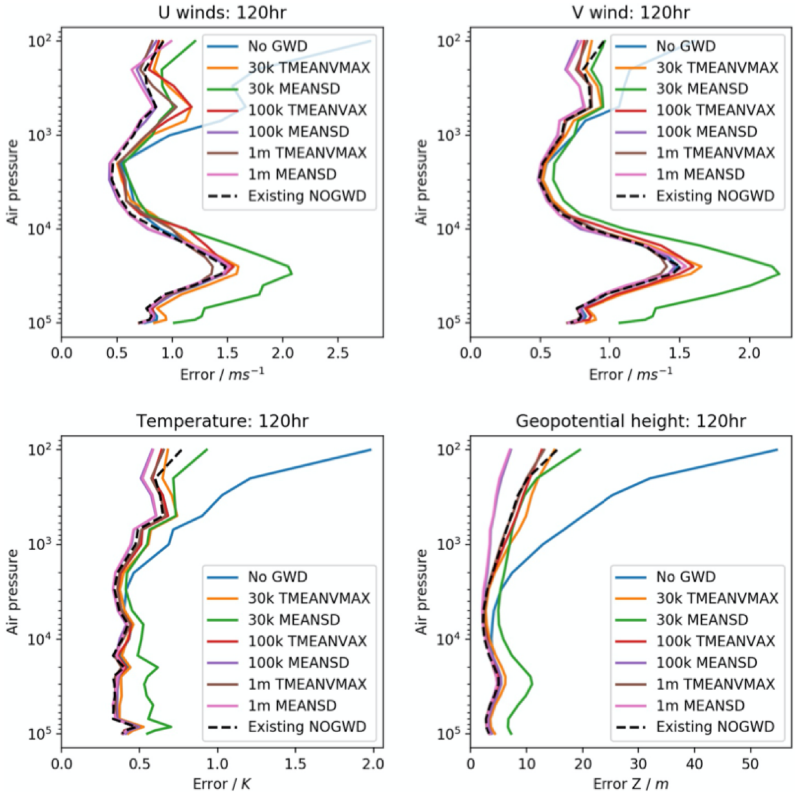
- improve models by training on unaffordable schemes



e.g. Gravity wave drag. Chantry et al, 2021, JAMES

ML to emulate parametrisations

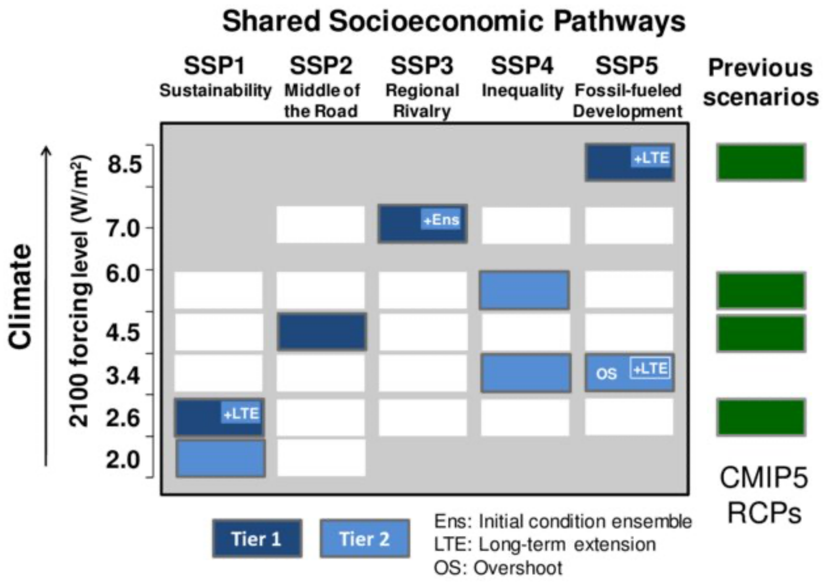
“Using the current operational CPU hardware, our emulators have a similar computational cost to the existing scheme, but are heavily limited by data movement. On GPU hardware, our emulators perform 10 times faster than the existing scheme on a CPU.”

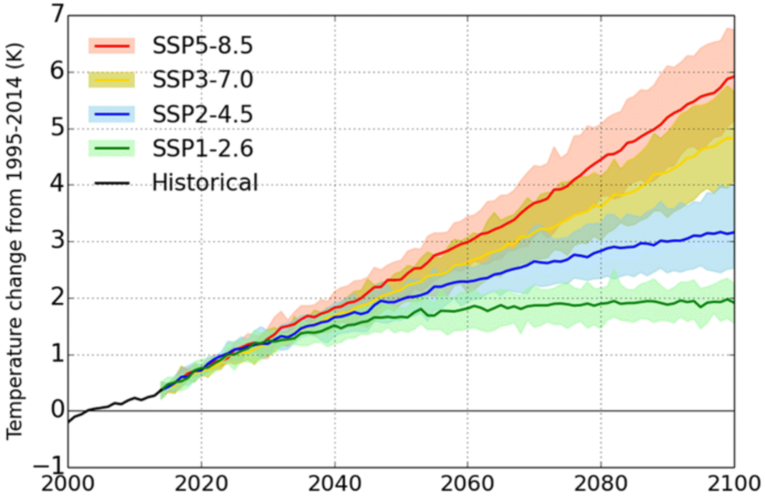


e.g. Gravity wave drag. Chantry et al, 2021, JAMES

What about *climate* prediction

- Fundamentally different problem to initialised forecasts

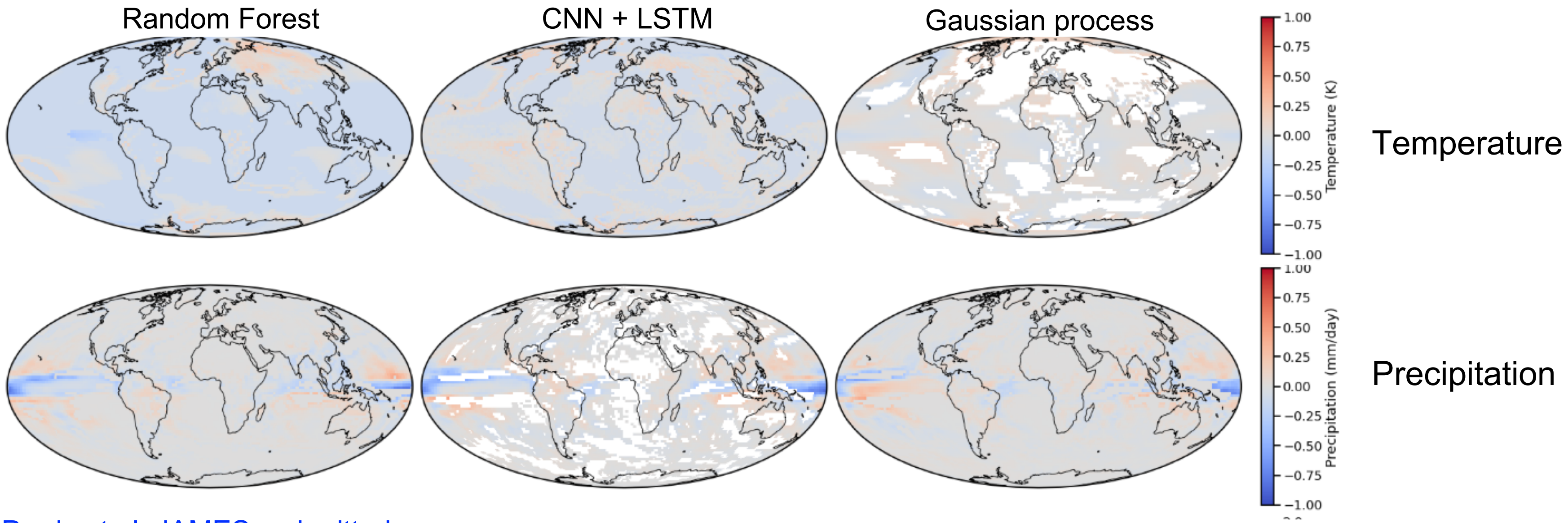




What about *climate* prediction: a soft AI approach?

Yearly global mean CO_2, CH_4 + maps of aerosol → **ML** → Maps of T and precipitation averaged over 2051-2100

E.g. train to predict SSP2-4.5 given training data from other scenarios



Watson-Parris et al, JAMES, submitted

Soft AI

Benefits

Outperform traditional models (speed and/or complexity)

Easier to include (complex) physical constraints

Generate ample model data

Challenges

Harder to gain new understanding about phenomena

Summary

- Focused on goal of using ML to improve Weather and Climate *Prediction*
- Each timescale has its own challenges and opportunities
- Hard, Medium, and Soft AI each have own pros and cons

Thanks for listening!

Hannah.Christensen@physics.ox.ac.uk