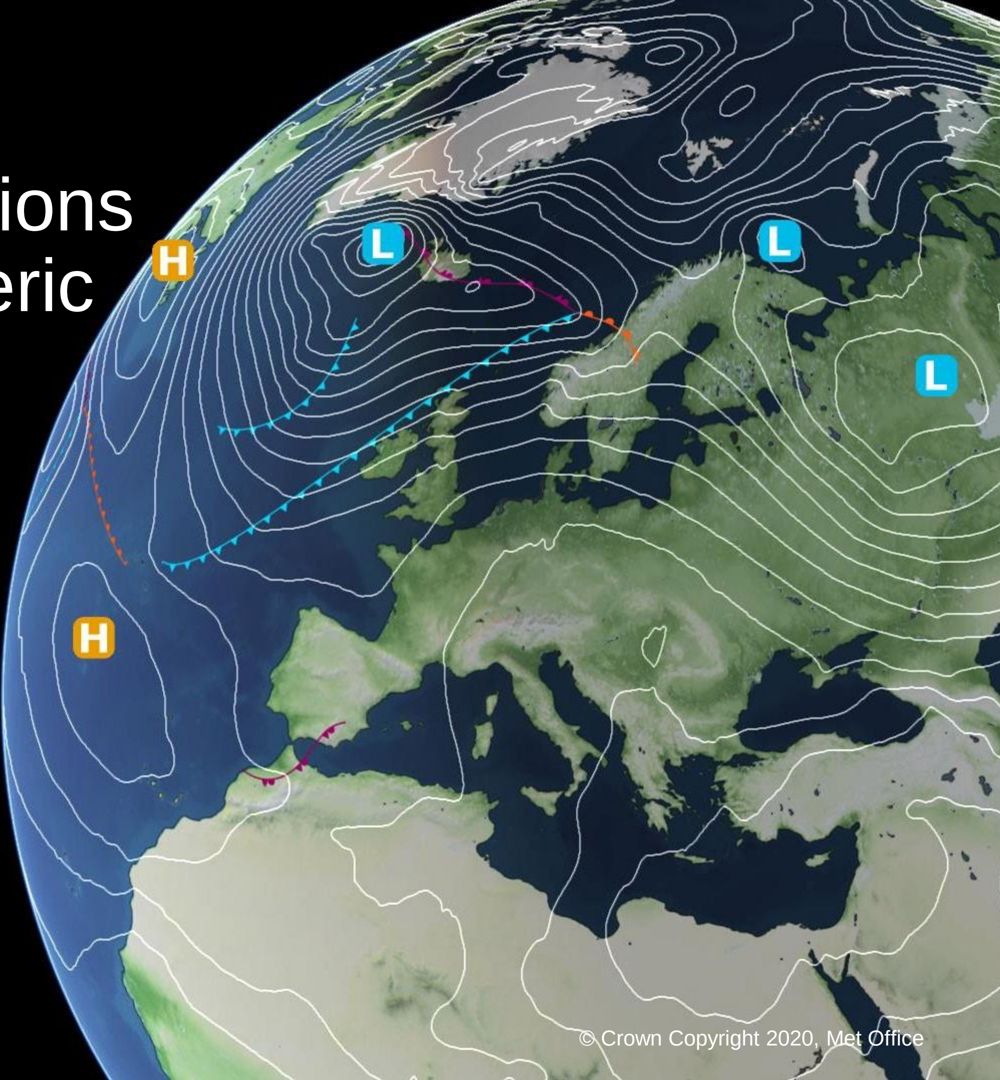
Met Office

# Machine Learning Applications in the Met Office Atmospheric Processes and Parametrizations Team



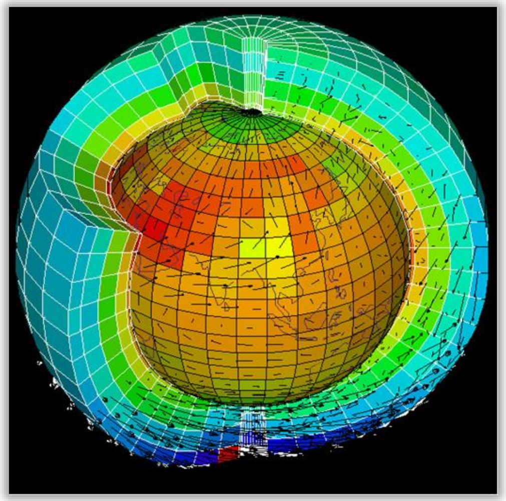

Kwinten Van Weverberg,
Annelize van Niekerk, Tom Dunstan,
Cyril Morcrette
21 April 2022

# Outline

- Weather and climate models and the role of parametrizations
- Examples of machine learning applications in the Atmospheric Processes and Parametrizations team
    - Gravity Wave Drag (Annelize van Niekerk)
    - Radiation (Tom Dunstan)
    - Convection Available Potential Energy (Cyril Morcrette)
    - Clouds (Kwinten Van Weverberg)
- Conclusions

# Weather and climate models and the role of parametrizations

- The atmosphere is divided into millions of "grid-boxes“, with dimensions of ~ 50 x 50 x 0.5 km
- Each grid box carries several “prognostic variables”, such as temperature, wind and pressure for which the primitive equations are solved numerically using a finite elements/volumes time stepping



The Primitive Equations

**Equations of (horizontal) motion**

$$\frac{\partial u}{\partial t}+\frac{u}{a\cos\phi}\frac{\partial u}{\partial \lambda}+\frac{v}{a}\frac{\partial u}{\partial \phi}+w\frac{\partial u}{\partial z}-\frac{uv}{a}\tan\phi = fv-\frac{1}{\rho a\cos\phi}\frac{\partial p}{\partial \lambda}+\mathscr{F}_1$$

$$\frac{\partial v}{\partial t}+\frac{u}{a\cos\phi}\frac{\partial v}{\partial \lambda}+\frac{v}{a}\frac{\partial v}{\partial \phi}+w\frac{\partial v}{\partial z}+\frac{u^2}{a}\tan\phi = -fu-\frac{1}{\rho a}\frac{\partial p}{\partial \phi}+\mathscr{F}_2$$

**Hydrostatic Equilibrium equation**

$$\frac{1}{\rho}\frac{\partial p}{\partial z}=-g=\frac{\partial \Phi}{\partial z}$$

**Continuity Equation**

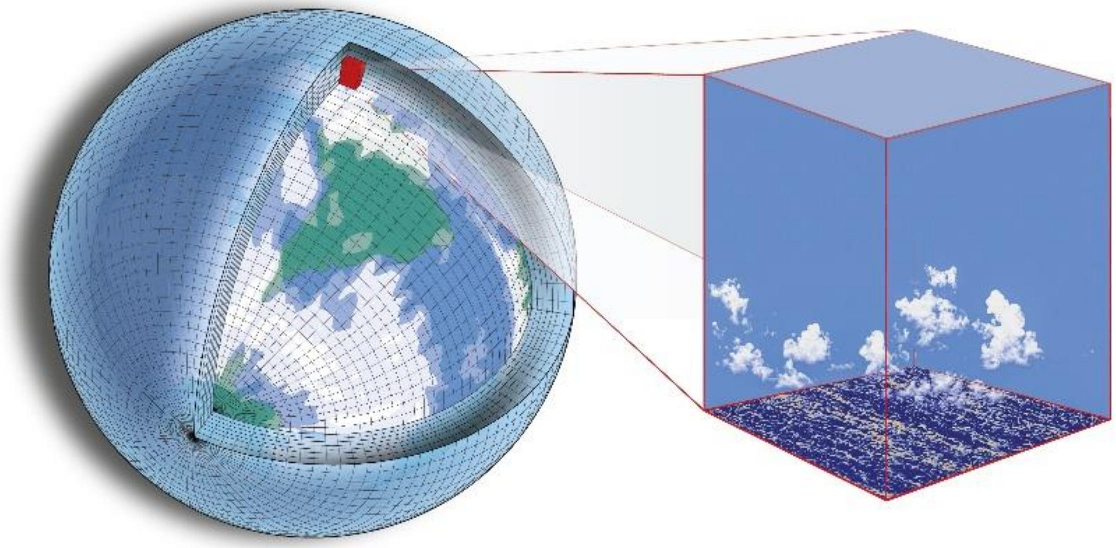
$$\frac{1}{a\cos\phi}\frac{\partial u}{\partial \lambda}+\frac{1}{a\cos\phi}\frac{\partial (v\cos\phi)}{\partial \phi}+\frac{w}{\rho(z)}\frac{\partial\left(\rho(z)w\right)}{\partial z}=0$$

**Thermodynamic equation**

$$\frac{\partial \theta}{\partial t}+\frac{u}{a\cos\phi}\frac{\partial \theta}{\partial \lambda}+\frac{v}{a}\frac{\partial \theta}{\partial \phi}+w\frac{\partial \theta}{\partial z}=l$$

# Weather and climate models and the role of parametrizations

- Many processes happen on scales much smaller than the “resolved” model grid yet have important implications for the large-scale evolution of the atmosphere.
- These processes are typically represented by “parametrizations”, which are statistical, empirical or theoretical sub-models, e.g., representing cloud and precipitation formation, radiative transfer, land-atmosphere interactions etc.



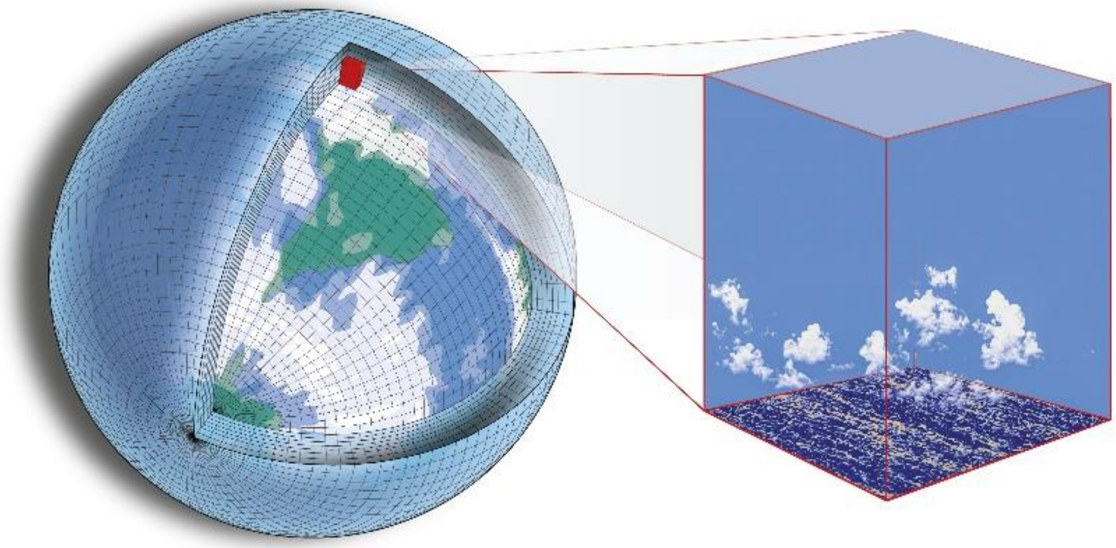
$$q_{liq} = \frac{(\mu_{turb} - \Delta s_{il})}{2}\left(1 + erf\left(\frac{(\mu_{turb} - \Delta s_{il})}{\sqrt{2}\sigma_{turb}}\right)\right) + \frac{\sigma_{turb}}{\sqrt{2\pi}} \exp \frac{-(\mu_{turb} - \Delta s_{il})^2}{2\sigma_{turb}{}^2}$$

$$cf_{liq} = \frac{1}{2}\left(1 + erf\left(\frac{(\mu_{turb} - \Delta s_{il})}{\sqrt{2}\sigma_{turb}}\right)\right)$$

Met Office

# Weather and climate models and the role of parametrizations

- Many processes happen on scales much smaller than the “resolved” model grid yet have important implications for the large-scale evolution of the atmosphere.
- These processes are typically represented by “parametrizations”, which are statistical, empirical or theoretical sub-models, e.g., representing cloud and precipitation formation, radiative transfer, land-atmosphere interactions etc.



Large-scale state
(water content and cloud fraction in grid box)

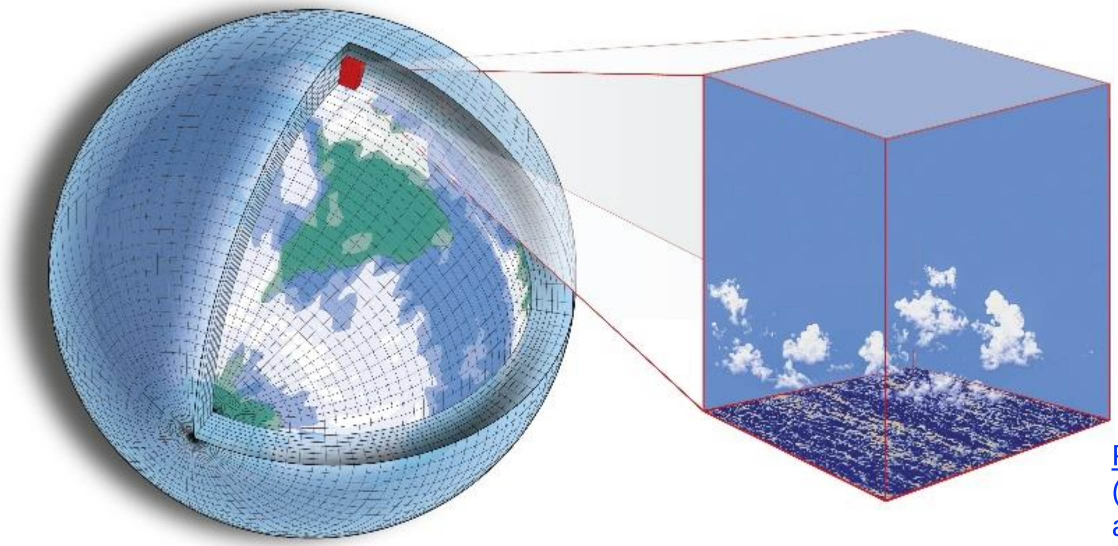
$$\boxed{q_{liq}} = \frac{(\mu_{turb} - \Delta s_{il})}{2}(1 + erf(\frac{(\mu_{turb} - \Delta s_{il})}{\sqrt{2}\sigma_{turb}})) + \frac{\sigma_{turb}}{\sqrt{2\pi}} \exp \frac{-(\mu_{turb} - \Delta s_{il})^2}{2\sigma_{turb}^2}$$

$$\boxed{cf_{liq}} = \frac{1}{2}(1 + erf(\frac{(\mu_{turb} - \Delta s_{il})}{\sqrt{2}\sigma_{turb}}))$$

Met Office

# Weather and climate models and the role of parametrizations

- Many processes happen on scales much smaller than the “resolved” model grid yet have important implications for the large-scale evolution of the atmosphere.
- These processes are typically represented by “parametrizations”, which are statistical, empirical or theoretical sub-models, e.g., representing cloud and precipitation formation, radiative transfer, land-atmosphere interactions etc.



Large-scale state
(water content and cloud fraction in grid box)

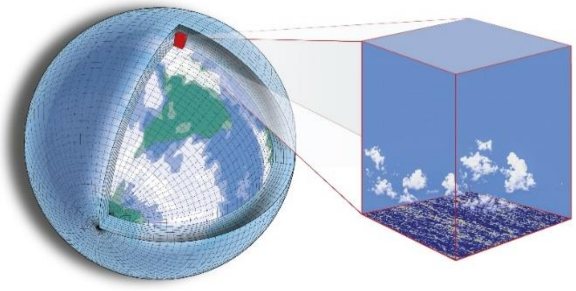
$$q_{liq} = \frac{(\mu_{turb} - \Delta s_{il})}{2}(1 + erf(\frac{(\mu_{turb} - \Delta s_{il})}{\sqrt{2}\sigma_{turb}})) + \frac{\sigma_{turb}}{\sqrt{2\pi}} \exp \frac{-(\mu_{turb} - \Delta s_{il})^2}{2\sigma_{turb}^2}$$

$$cf_{liq} = \frac{1}{2}(1 + erf(\frac{(\mu_{turb} - \Delta s_{il})}{\sqrt{2}\sigma_{turb}}))$$

Parametrization
(linking large-scale state to sub-grid-turbulence, and a Gaussian sub-grid moisture distribution)

# Weather and climate models and the role of parametrizations

- Many processes happen on scales much smaller than the "resolved" model grid yet have important implications for the large-scale evolution of the atmosphere.
- These processes are typically represented by "parametrizations", which are statistical, empirical or theoretical sub-models, e.g., representing cloud and precipitation formation, radiative transfer, land-atmosphere interactions etc.



→ Parametrizations can add considerable **cost (~ 35%)** to the model integration time, and are typically a large source of **uncertainty**, due to their crude approximations, empiricism with limited observational data, or lack of theoretical framework.

→ Parametrizations are **indispensable** to correctly represent how the atmosphere evolves over time in climate and weather forecasting models.

→ We could use **machine learning** to replace complete parametrizations:

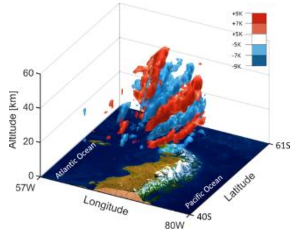
  → Learn on data from existing parametrizations to just speed up the model

  → Learn on observational data sets (relating various observed variables)

  → Learn on high-resolution model data (where parametrizations should start having a smaller impact)

→ We could use **machine learning** to replace parts of parametrizations

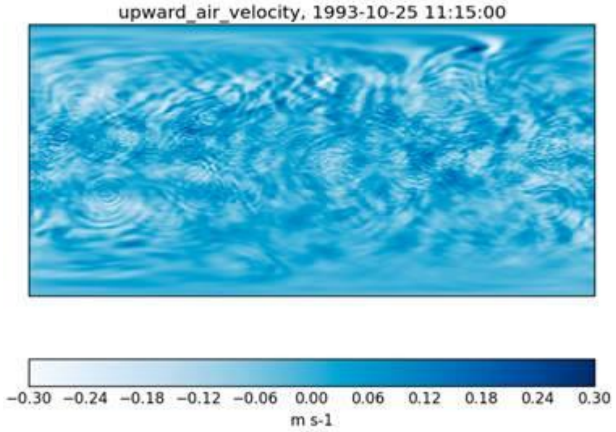
# Examples of machine learning application in the Atmospheric Processes and Parametrizations Team



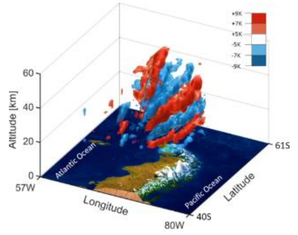

# Gravity wave drag (van Niekerk, Hardiman, Prudden, Scaife)

*Data-Wave: Five-year international project, funded by the Virtual Earth System Research Institute (VESRI), to enhance and improve gravity wave representation in climate simulations using machine learning.*

Aim to:
- emulate existing non-orographic gravity wave scheme
- improve on existing scheme (i.e. descent of the Quasi Biennal Oscillation – a common and important bias in climate models).

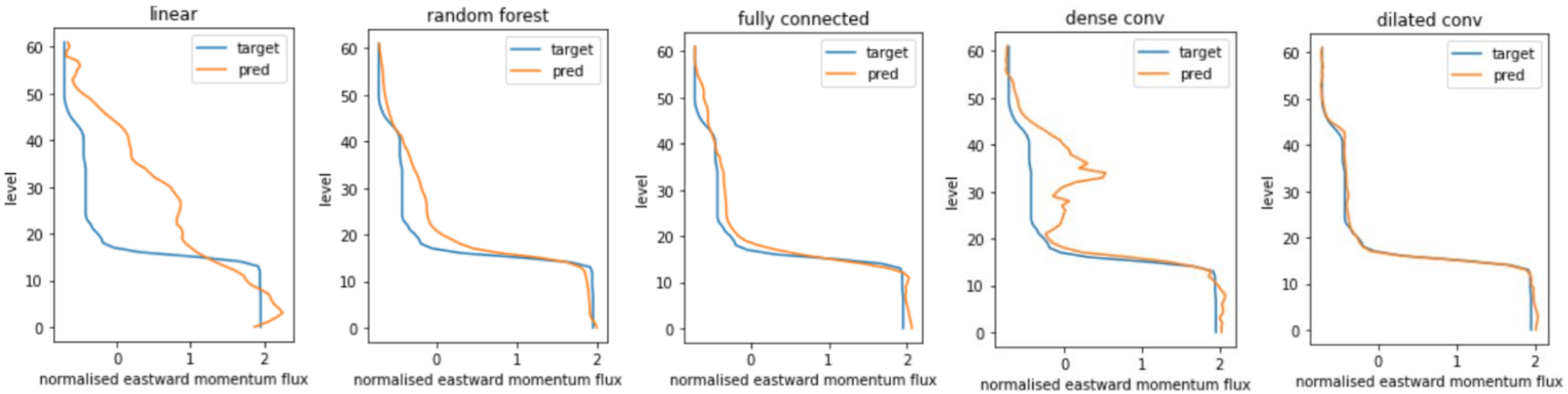



*Gravity waves seen in the vertical velocity field of the Met Office climate model*



# Gravity wave drag: Sensitivity to ML algorithm

(van Niekerk, Hardiman, Prudden, Scaife)




**test MSE**

linear: 0.1876

dense convolution: 0.0962

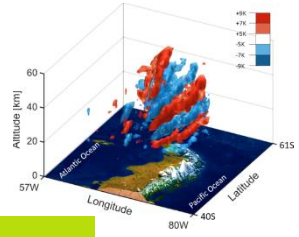
random forest: 0.0380

fully connected: 0.0238

dilated convolution: 0.0092

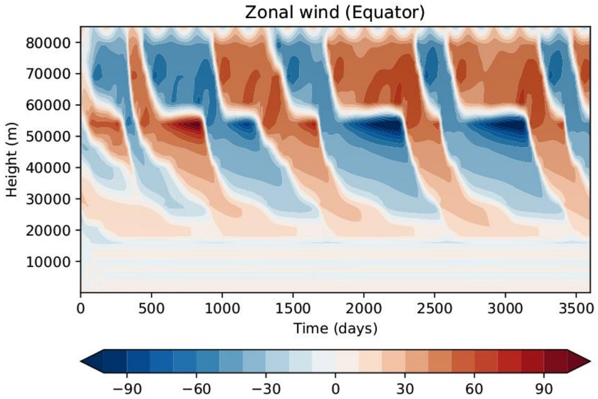
Trained convolutional neural network to output non-orographic GW acceleration, with zonal wind as input.

# Gravity wave drag: QBO (van Niekerk, Hardiman, Prudden, Scaife)

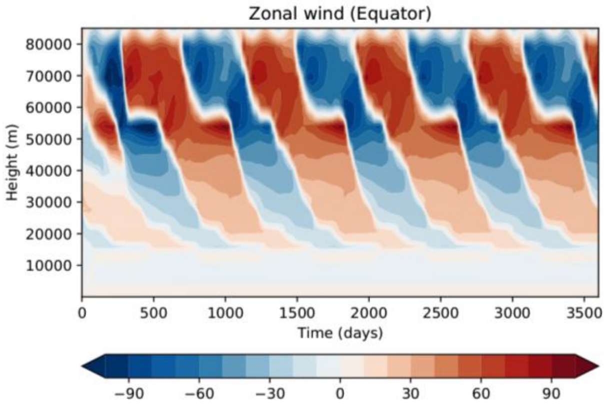



- Simulate QBO in simple 1D model
  - enables us to optimize the network
  - gives hope that scheme will work when coupled to climate model
- Train on two years of data, containing QBOE and QBOW phases, and NH winter stratosphere with SSW and strong polar vortex.

- Neural net coupled to 1D model reproduces a QBO with correct period, amplitude and structure.

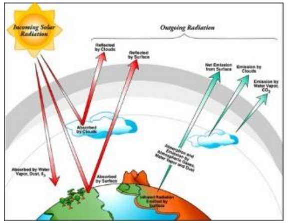



*QBO simulated by offline NOGWD scheme coupled to 1D model.*

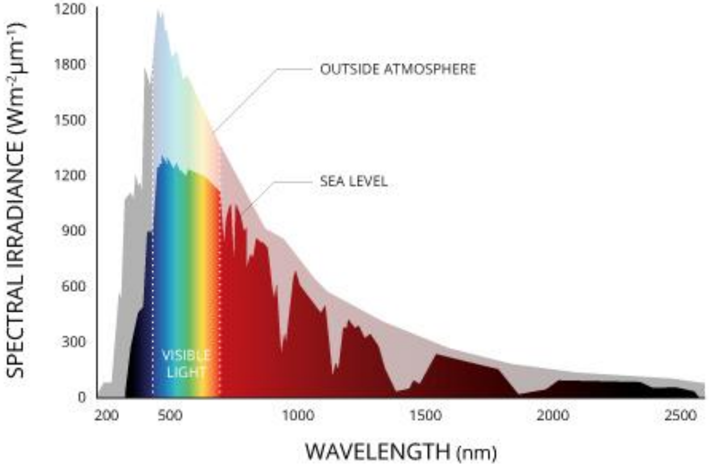



*QBO simulated by neural network coupled to 1D model.*

# Radiation (Dunstan, Manners)



Full radiation calculations are very expensive and hence are highly simplified in weather and climate models:

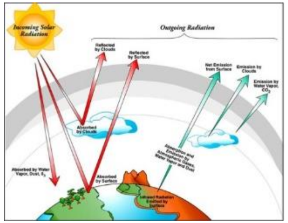



- Two-stream fluxes:
  - up, down (diffuse / direct), net, heating rates
- Gaseous absorption:
  - Correlated-K sorting of spectrum for limited number of bands
- Cloud representation:
  - Monte-Carlo Independent Column Approximation (McICA)

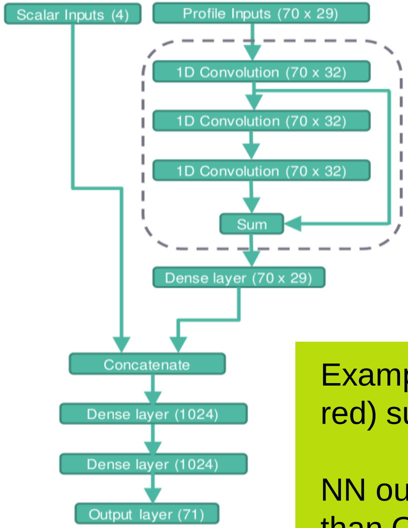
Two configurations:

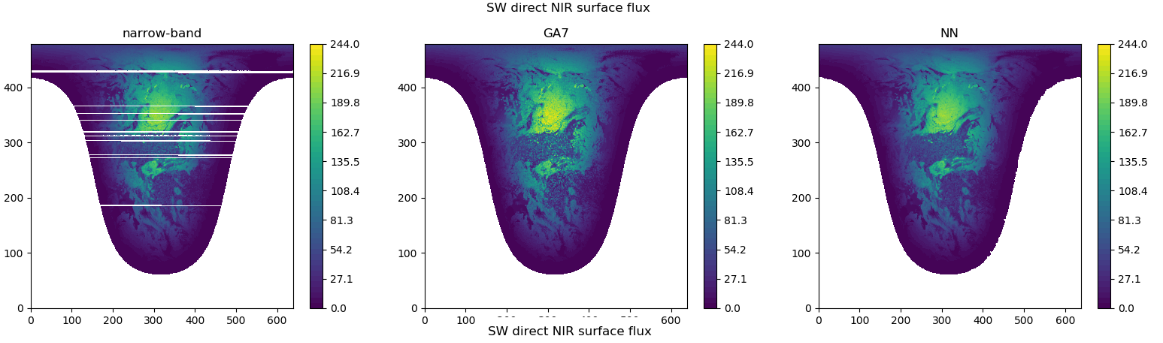
- GA7 (operational weather climate model, broad-band)
  - 9 bands in longwave
  - 6 bands in shortwave
  - 81 McICA sub-columns for clouds
- NB (narrow-band, expensive 'truth')
  - 300 bands in longwave
  - 260 bands in shortwave
  - 5897 McICA sub-columns for clouds

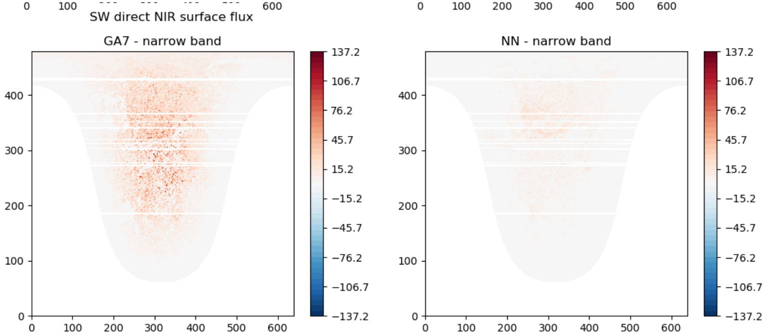
# Radiation (Dunstan, Manners)



Full radiation calculations are very expensive and hence are highly simplified in weather and climate models:



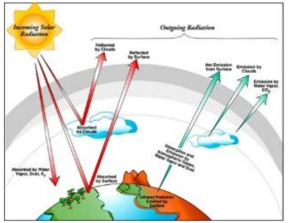



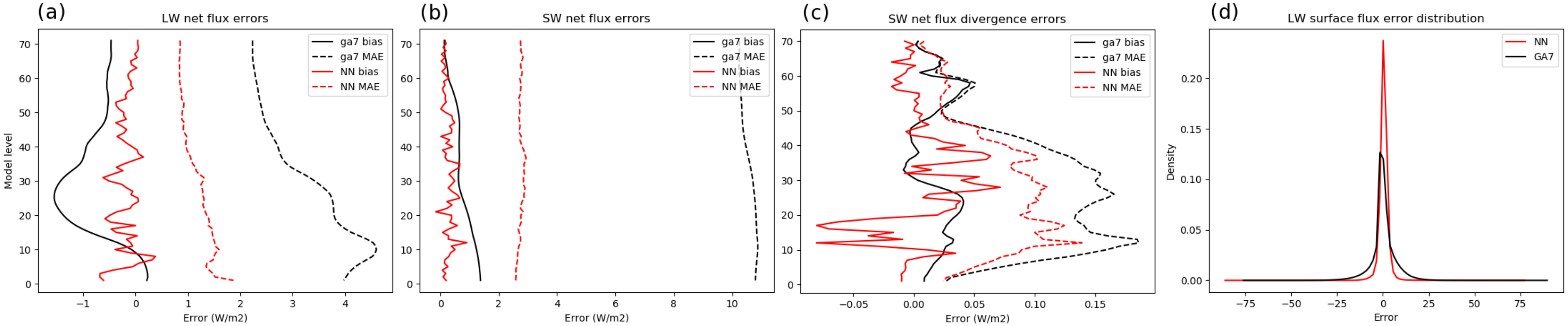

Example of shortwave direct NIR (near infra-red) surface fluxes

NN output is more accurate and less noisy than GA7 configuration than NB output

# Radiation (Dunstan, Manners)



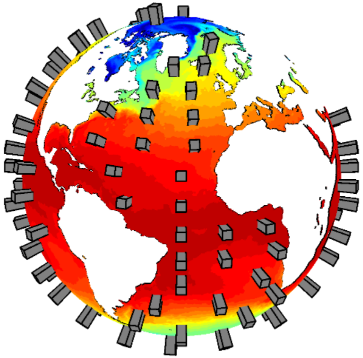
Full radiation calculations are very expensive and hence are highly simplified in weather and climate models:

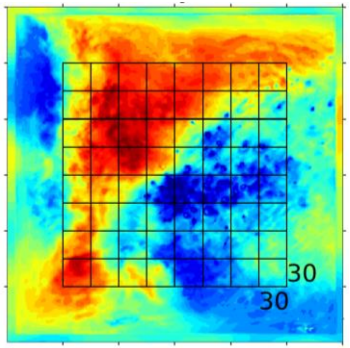



Error profiles (a-c) for mean error (bias) and mean absolute error (MAE) of net flux and net flux divergence
Error distribution for LW net surface fluxes (d) (All errors are w.r.t. narrow-band SOCRATES output)

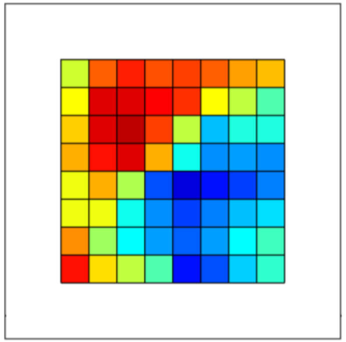
# Convection (Morcrette, Giles, Guillas, Van Weverberg, Xie, Zhang, Lin)

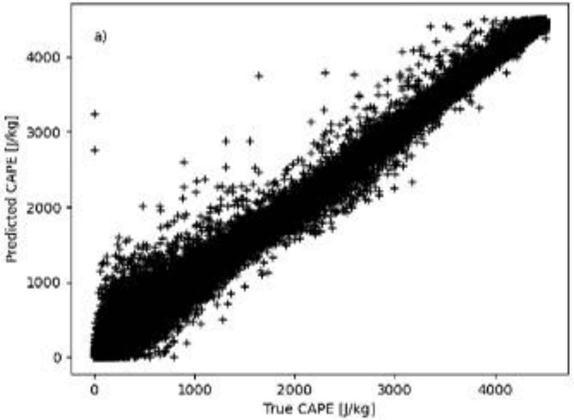
Machine learning of the PDF of Convection Available Potential Energy (CAPE) from lots of coarse-grained kilometre-scale numerical weather prediction model simulations, distributed around the globe. This PDF can be used in convective parametrizations



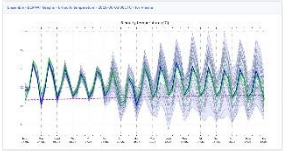



Use e.g. 30x30 averaging to coarse-grain 1.5 km data to 45 km






Use convoluted neural network to learn mean and PDF of CAPE in the global model from large-scale state



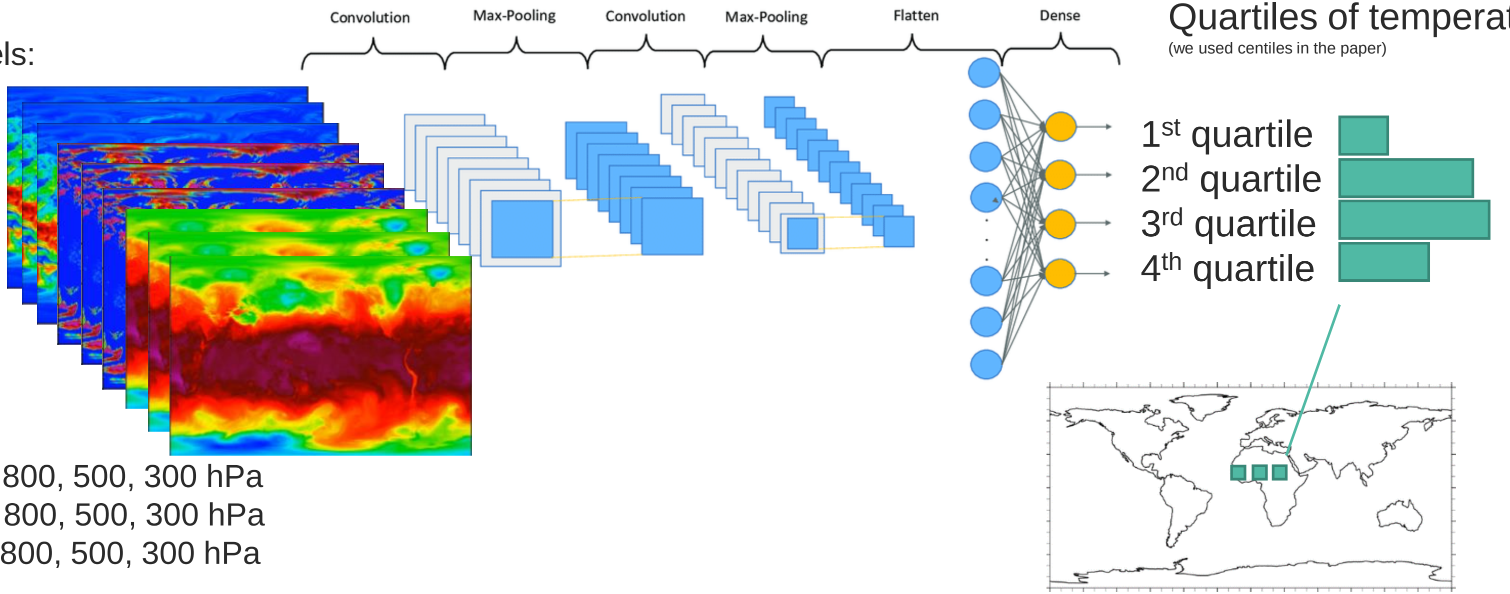
# Ensemble Forecasting (Clare, Jamil, Morcrette)

Data-driven (no model involved). Learning uncertainty 3 and 5 days ahead in Z500 and T850 using ResNets

n channels:

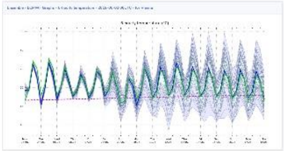
Quartiles of temperature

(we used centiles in the paper)

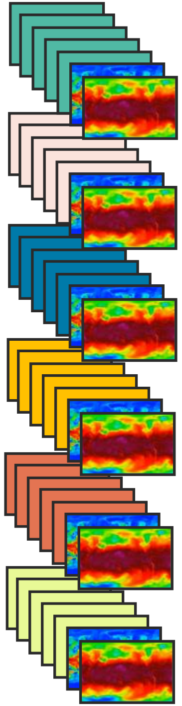



T, 800, 500, 300 hPa
u, 800, 500, 300 hPa
v, 800, 500, 300 hPa

Hence can produce global maps (5 deg lat/lon grid) of
PDF of centiles of T850 or Z500 at day+3 and day+5



# Ensemble Forecasting (Clare, Jamil, Morcrette)



… geopotential

… temperature

… zonal wind

… specific humidity

… potential vorticity

… orog,lat,,land-sea mask

CNN

CNN

CNN

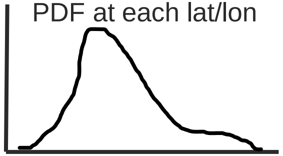
CNN

CNN

CNN

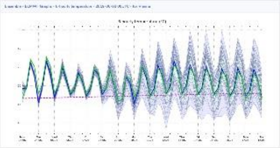
A meta-learner, that combines the output from each of the other CNNs.

CNN




Six separate CNNs trained using 500 hPa geopotential and 800 hPa T and several other fields of …

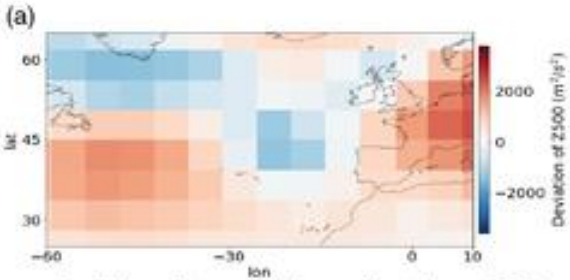
Hence can produce global maps (5 deg lat/lon grid) of PDF of centiles of T850 or Z500 at day+3 and day+5



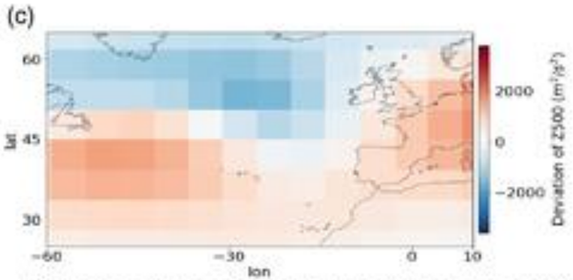
# Ensemble Forecasting (Clare, Jamil, Morcrette)

Hind-cast for **Storm Ophelia**
*(00 UTC on 17 Oct 2017)*

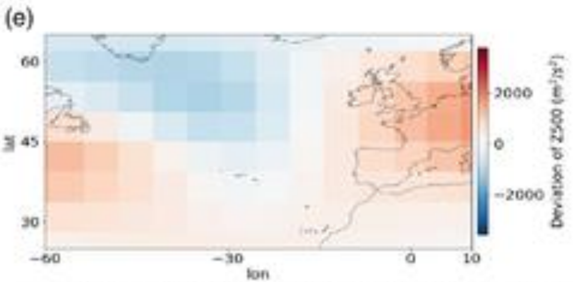
*Using expectation (mean) value of geopotential from the predicted PDF.*




Deviation of true value from climatology - Z500.




Deviation of 3-day hindcast from climatology - Z500.



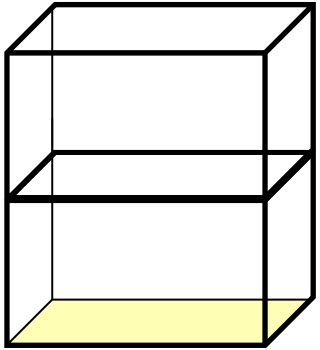

Deviation of 5-day hindcast from climatology - Z500.

# Clouds: sub-grid organisation (Van Weverberg et al)

Cloud parametrization’s task is to diagnose the fraction of the grid box that will be cloudy

Without a cloud parameterisation: cloud fraction in a grid box is 0 or 1


**Sub-saturated mean state**

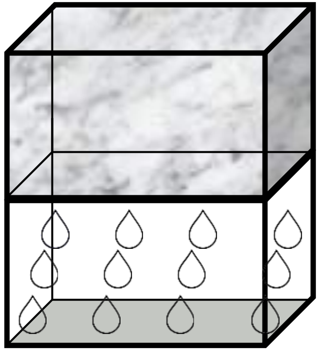


# Clouds: sub-grid organisation (Van Weverberg et al)

Cloud parametrization’s task is to diagnose the fraction of the grid box that will be cloudy

Without a cloud parameterisation: cloud fraction in a grid box is 0 or 1

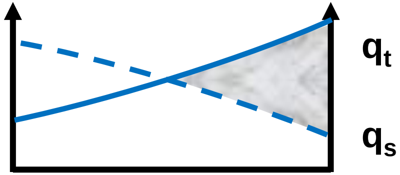

**Saturated mean state**



# Clouds: sub-grid organisation (Van Weverberg et al)

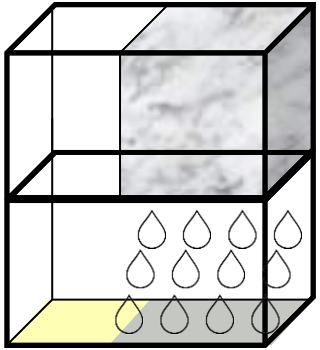
Cloud parametrization’s task is to diagnose the fraction of the grid box that will be cloudy

Without a cloud parameterisation: cloud fraction in a grid box  is 0 or 1




**Need for sub-grid variability**

→Important for radiation
→Important for precipitation



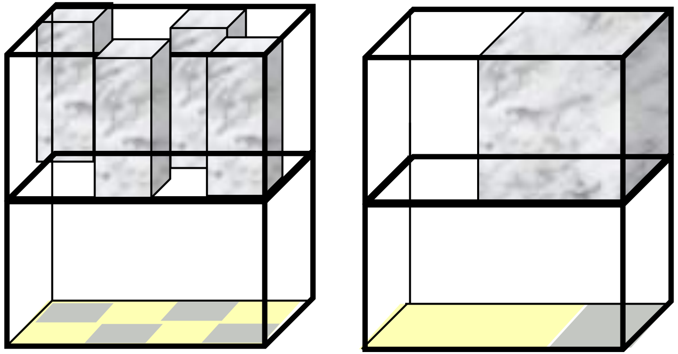
# Clouds: sub-grid organisation (Van Weverberg et al)

Cloud parametrization's task is to diagnose the fraction of the grid box that will be cloudy

Without a cloud parameterisation: cloud fraction in a grid box is 0 or 1

**Organisation of clouds is also critical, but currently not taken into account**

→The more cloud edge, the faster cloud evaporation



# Clouds: sub-grid organisation (Van Weverberg et al)

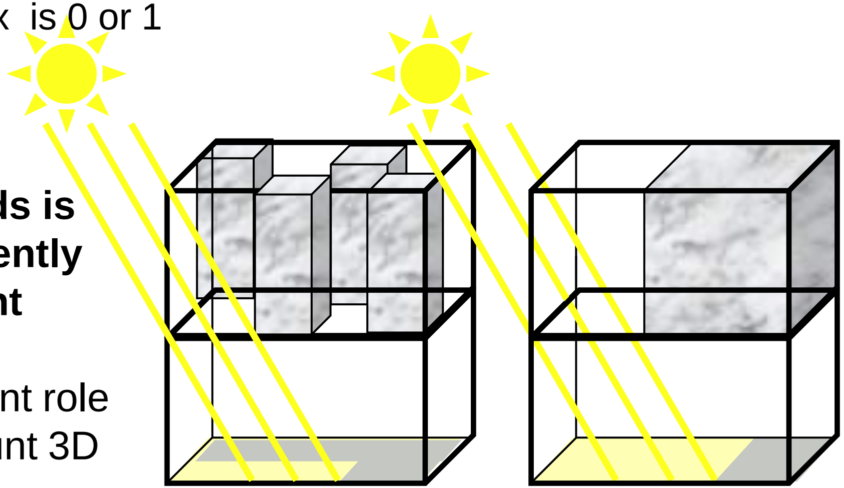
Cloud parametrization’s task is to diagnose the fraction of the grid box that will be cloudy

Without a cloud parameterisation: cloud fraction in a grid box is 0 or 1

**The organisation depends on lots of environmental conditions, e.g. turbulence, surface properties, atmospheric state**

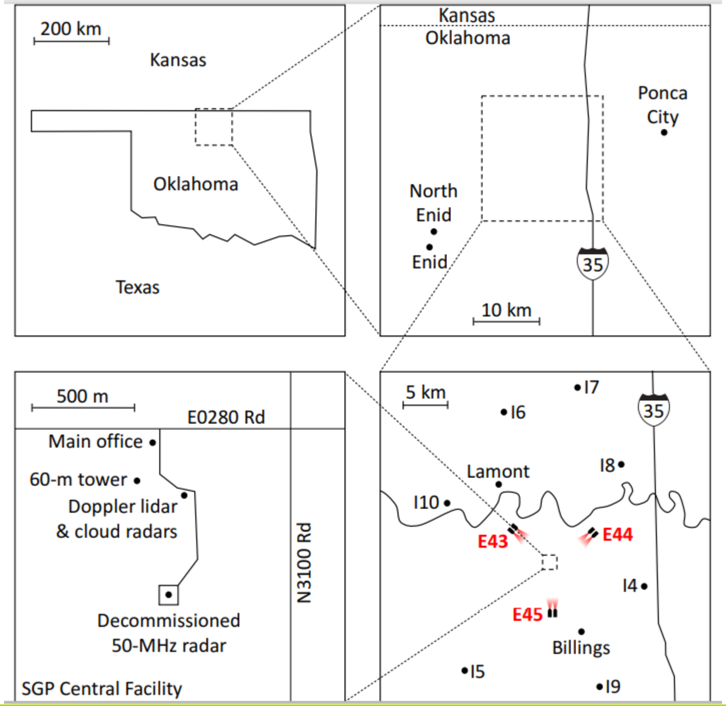
**Organisation of clouds is also critical, but currently not taken into account**

→Plays a very important role when taking into account 3D radiation



# Clouds: sub-grid organisation (Van Weverberg et al)

## US Department of Energy Clouds Optically Gridded by Stereo (COGS) Product




State of the art product using 3D reconstructed cloud locations over US Great Plains supersite, using digital cameras and stereophotogrammetry.

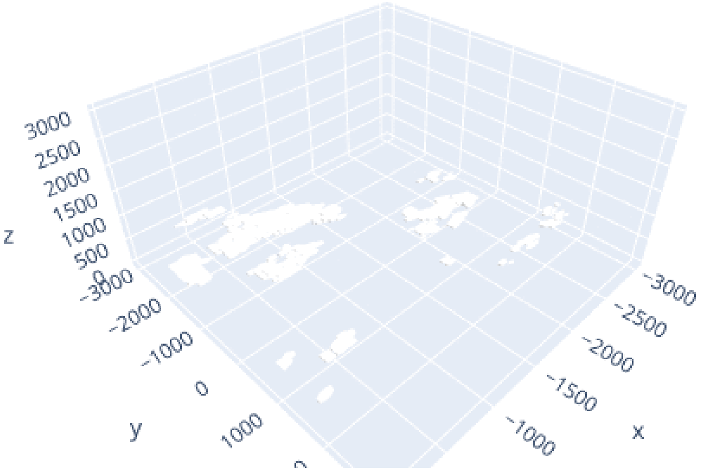
6 digital cameras installed around supersite, covering a 6 by 6 km area with a 50m grid spacing and samples every 20s.

Available for few years running now, with collocated information about the atmospheric state (temperature, humidity, turbulence …)

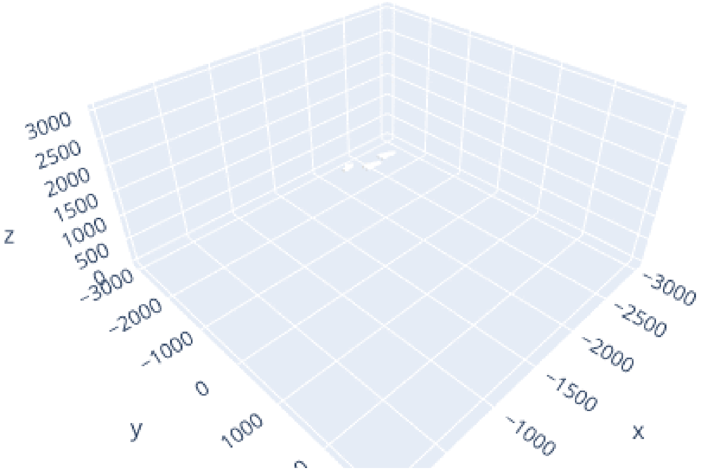
# Clouds: sub-grid organisation (Van Weverberg et al)

## US Department of Energy Clouds Optically Gridded by Stereo (COGS) Product

**Stratocumulus (15 October 2018)**




**Shallow cumulus (5 May 2019)**

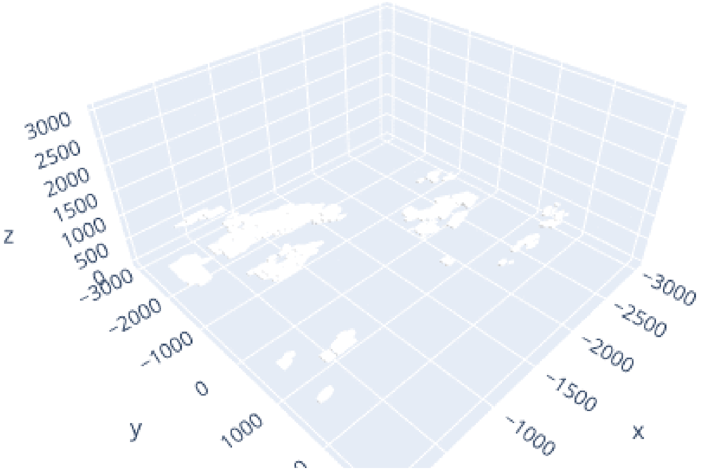


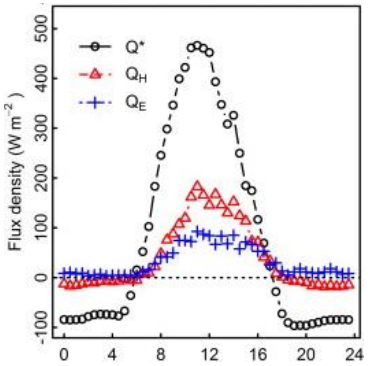
# Clouds: sub-grid organisation (Van Weverberg et al)

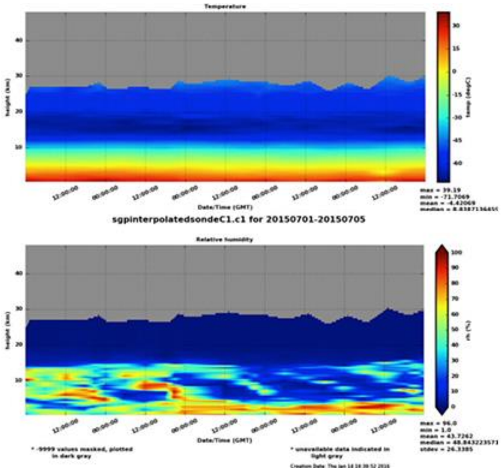
## US Department of Energy Clouds Optically Gridded by Stereo (COGS) Product

**Stratocumulus (15 October 2018)**

Calculate cloud perimeter for each "grid box" and use deep learning to link this with other observed variables in the domain:

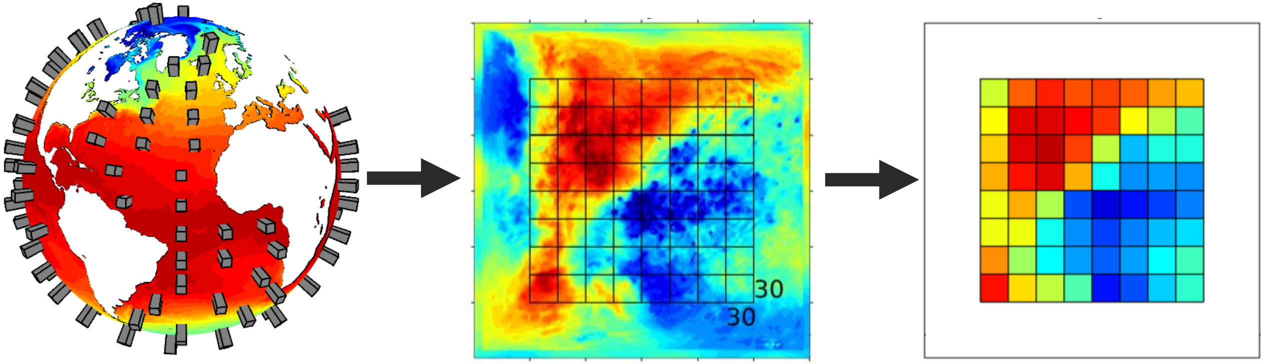


# Clouds: parametrization learning (Van Weverberg et al)

Machine learning of the cloud cover from lots of coarse-grained kilometre-scale numerical weather prediction model simulations, distributed around the globe. Replace the cloud fraction parametrization in coarse-resolution global circulation model




Use e.g. 30x30 averaging to coarse-grain 1.5 km data to 45 km

# Conclusions

- **Machine learning has many potential applications in parametrization development for Numerical Weather Prediction and climate modelling**
  - Direct learning of parametrization to expedite the model integration time
  - Learning of parts of the parametrization that are not well founded in theory (parameter optimisation)
  - Learning on high-resolution models or observations to improve coarse-resolution models

- **Many projects ongoing within the Atmospheric Processes and Parametrizations team at the Met Office:**
  - Learning of the radiation and gravity wave drag parametrization
  - Learning of Convective Available Potential Energy
  - Learning of Ensemble Forecasts
  - Learning of cloud organisation and the cloud parametrization

Met Office

# Thank you

Kwinten.vanweverberg@metoffice.gov.uk