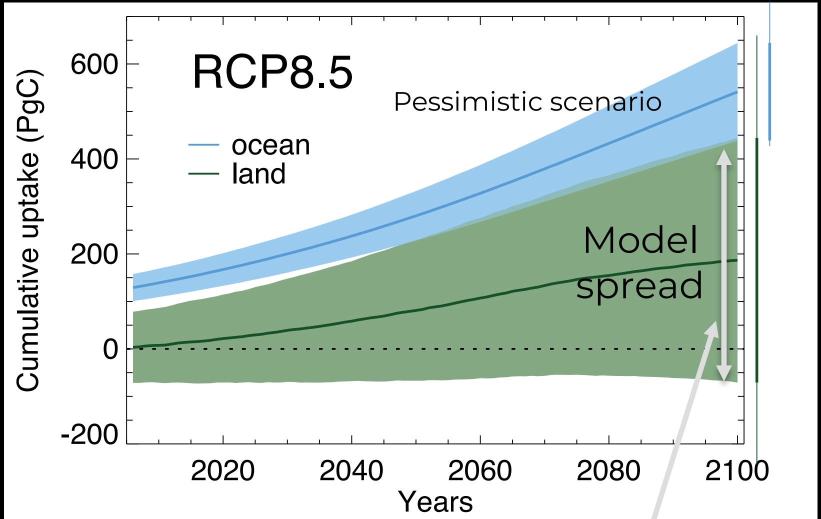
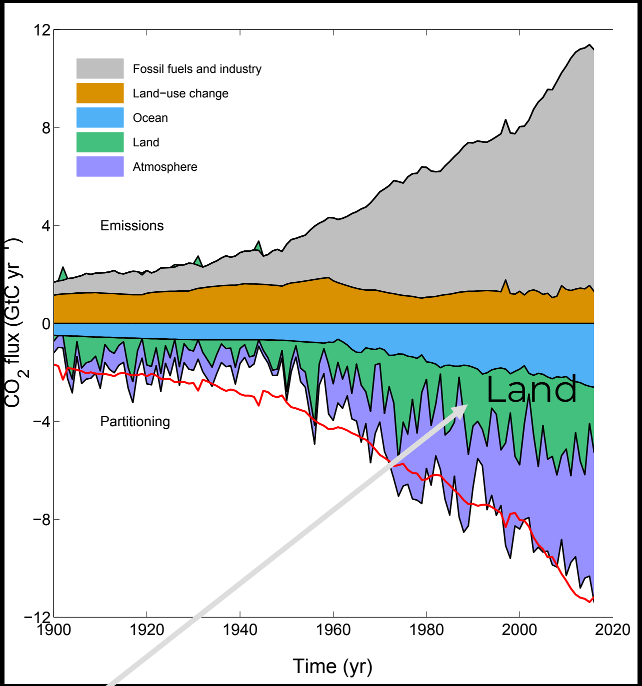
Machine learning models to improve parameterizations of climate models

Pierre Gentine – Columbia University
Mike Pritchard, Tom Beucler, Stephan Rasp, Wenli Zhao

Clivar August 2019

1. Global carbon uptake





Large intermodel spread, large interannual variability

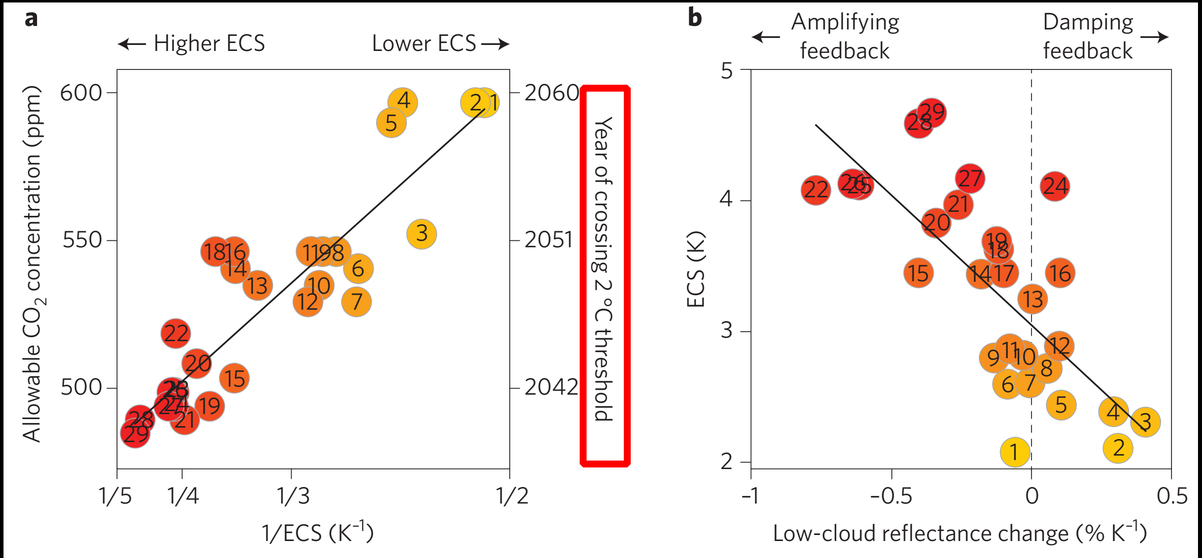
Define concentrations **[CO_2]=f(emission flux)**

2. Climate sensitivity

Still **substantial spread in model climate sensitivity global T=f(greenhouse gases)**:

Limits our climate mitigation and management capacity and increases cost

Mostly due to **clouds**



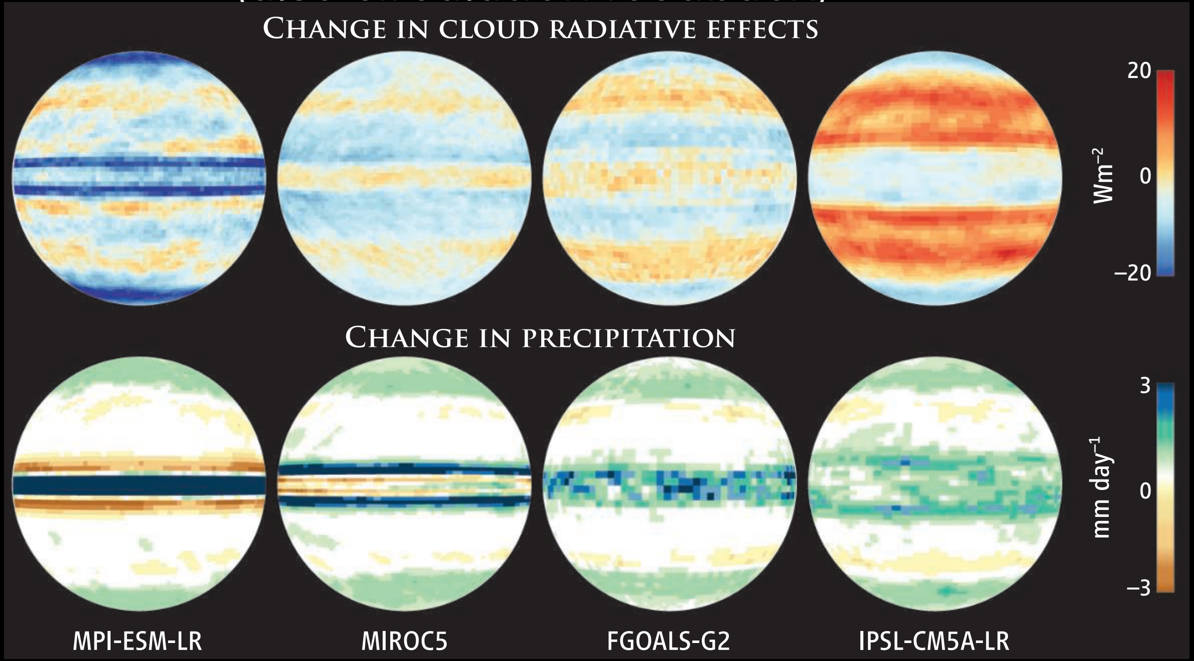
ECS = Equilibrium climate sensitivity (T response do CO_2 doubling)

Schneider, T., et al. (2017). *Nature Climate Change*

3. Regional climate sensitivity

Cloud impact is not just global but also regional
(also circulation feedback)

Aquaplanet
+4K
(no SST
feedback!)

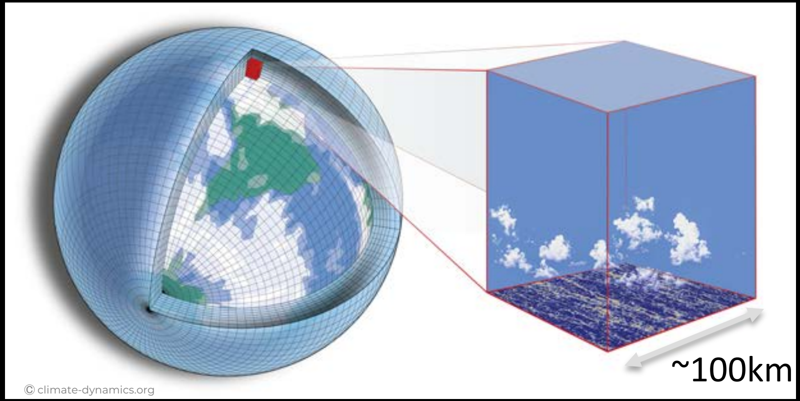


Regional climate prediction is too uncertain

Stevens, B., & Bony, S. (2013). *Science*

1. Using ML for climate: (deep) clouds

Parameterization: represent (physically or statistically) a physical process that cannot be resolved (e.g. clouds)
Typically physically based



$$\frac{\partial \bar{X}}{\partial t}_{|\text{clouds}} = f(\bar{X})$$ with $\bar{X}$ coarse-scale average of X

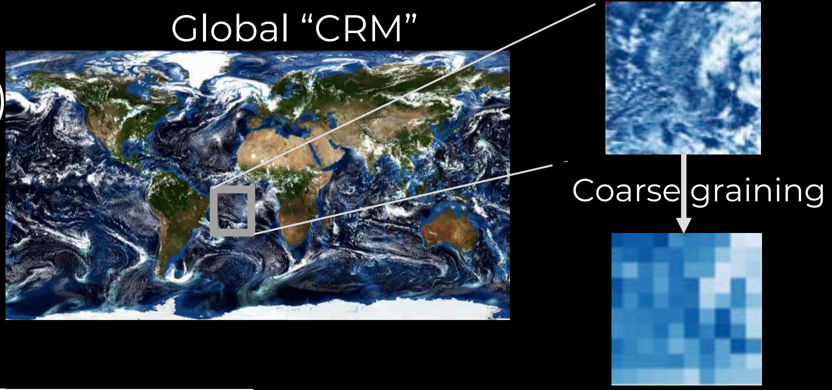
However: it has failed for ~40 years (Randall et al. 2003)
This largely **explains intermodal spread in climate prediction**

Randall, D. et al.. (2003)., *Bulletin Of The American Meteorological Society,*

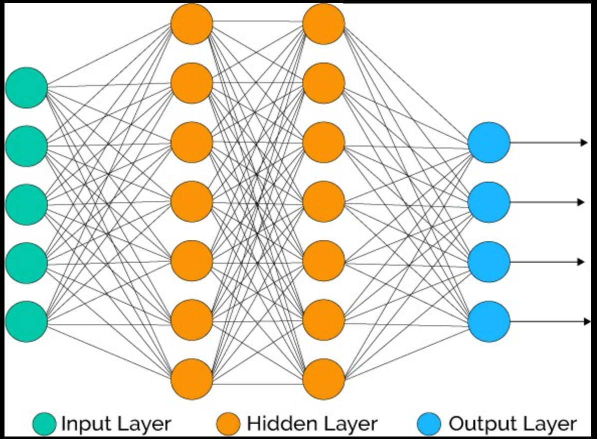
1. Using ML for climate: (deep) clouds

How can we solve this issue?
Take advantage of **cloud-resolving simulations** (~1km, **alleviate most biases** but very expensive)

Not "physical" but
Data-driven approach
(informed by cloud-resolving simulations)



Temperature $\bar{T}(z)$
Specific humidity $\bar{q}(z)$
Surface sensible heat flux $\overline{H}$
Surface evaporation $\overline{E}$
Surface pressure $\overline{P}_s$



$\frac{\partial \bar{T}}{\partial t}_{|\text{convection}}$
$\frac{\partial \bar{q}}{\partial t}_{|\text{convection}}$
Precipitation

Deep Neural Net or Convolutional NN

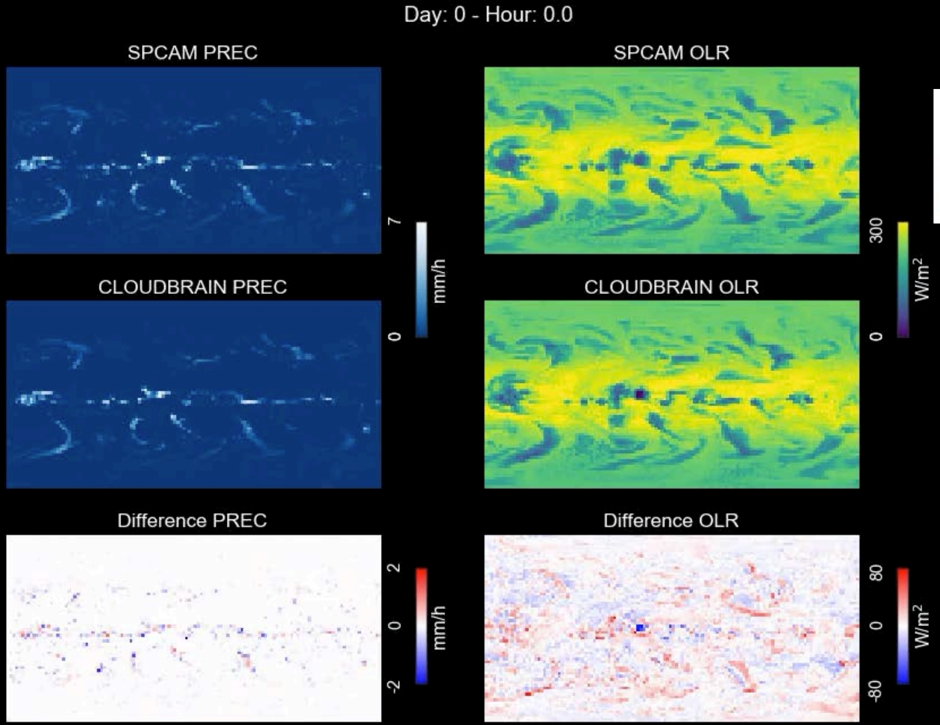
Cost function: misfit to coarse-grained high-res. model

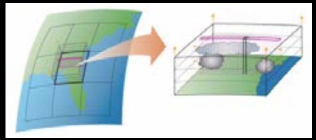
Gentine P., Pritchard M., Rasp S., Reinaudi G., *GRL, 2018*
Rasp, Pritchard and Gentine, *PNAS* 2018
Brenowitz and Bretherton, *GRL*, 2018

1. Using ML for climate: (deep) clouds

Coarse-grained Cloud-resolving Model (superparameterization)

Machine learning Coarse-resolution model



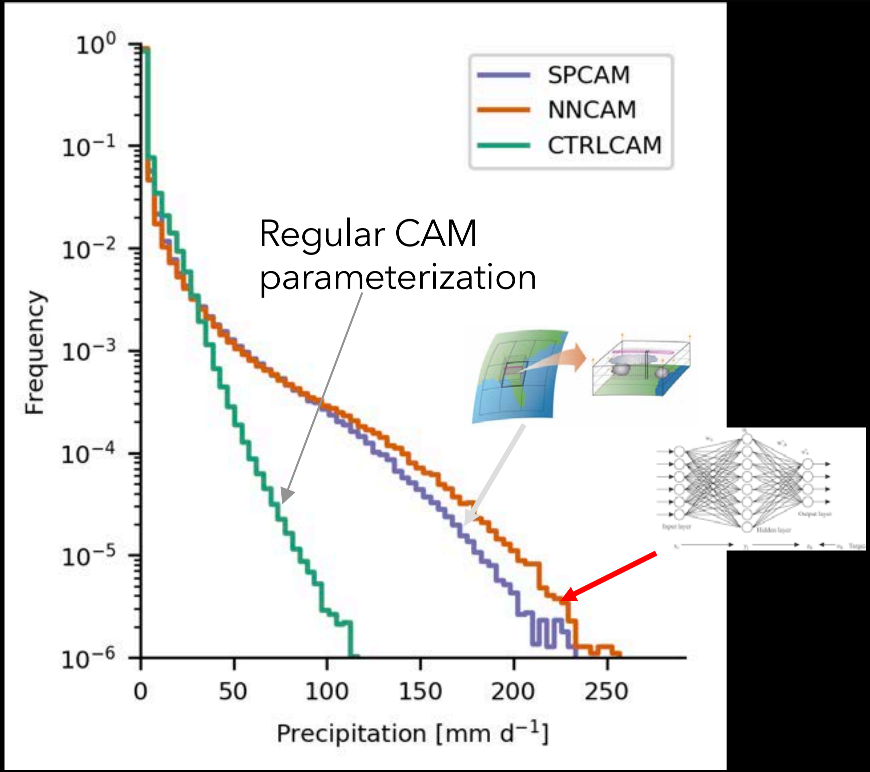


10 times cheaper than original coarse model, 1000 less expensive than high-res model

Question: generalization to unforeseen conditions? Climate change

Gentine P., Pritchard M., Rasp S., Reinaudi G., *GRL, 2018*
Rasp, Pritchard and Gentine, *PNAS* 2018
Brenowitz and Bretherton, *GRL*, 2018

1. Using ML for climate: (deep) clouds



Interactive model

Now we have good boundary condition to study hydrology ☺

10 times cheaper than original coarse model, 1000 less expensive than high-res model

Question: generalization to unforeseen conditions? Climate change: Poor!

Gentine P., Pritchard M., Rasp S., Reinaudi G., *GRL, 2018*
Rasp, Pritchard and Gentine, *PNAS* 2018

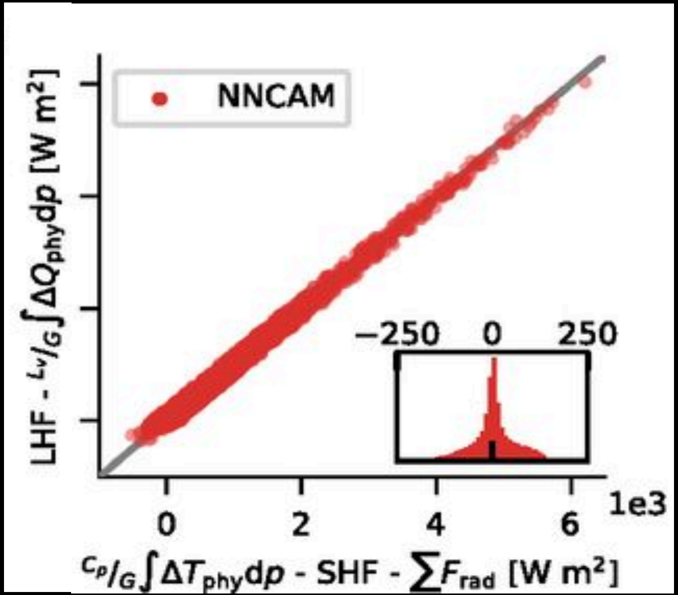
Issues

1. Physical Constraints

Energy conservation

Mass conservation

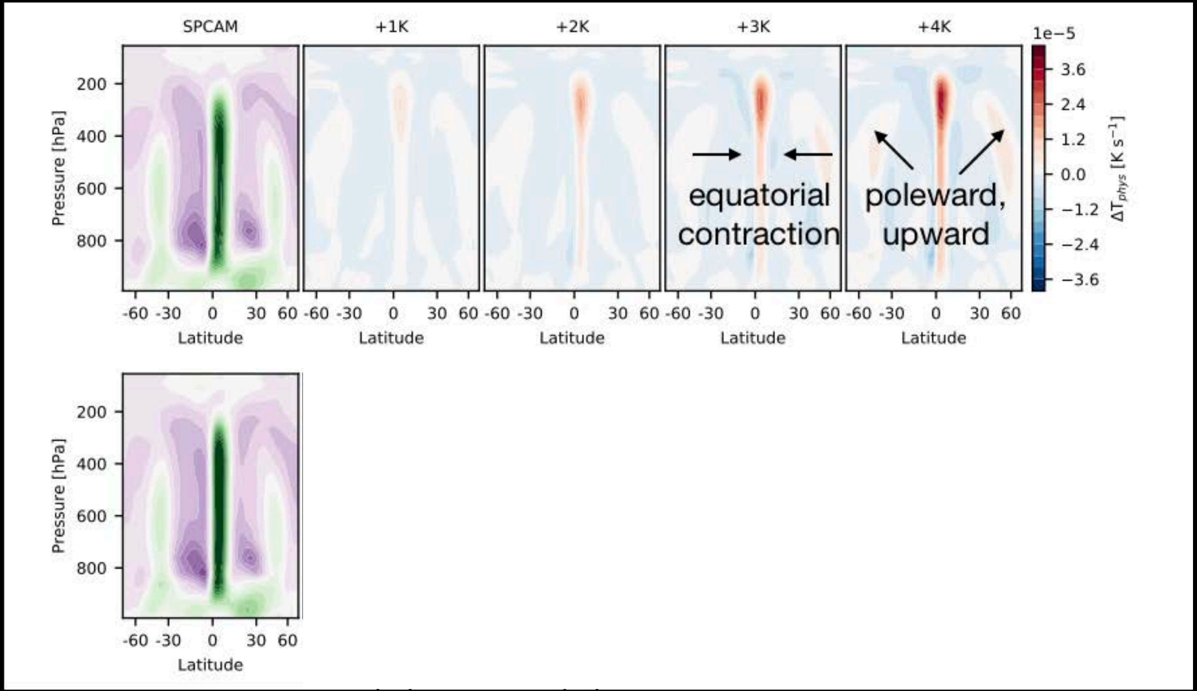
Only **approximate with ML**



Gentine P., Pritchard M., Rasp S., Reinaudi G., *GRL, 2018*
Rasp, Pritchard and Gentine, *PNAS* 2018

Issues

2. **Extrapolation**
ML has mostly been about interpolations using lots of data, poor extrapolation



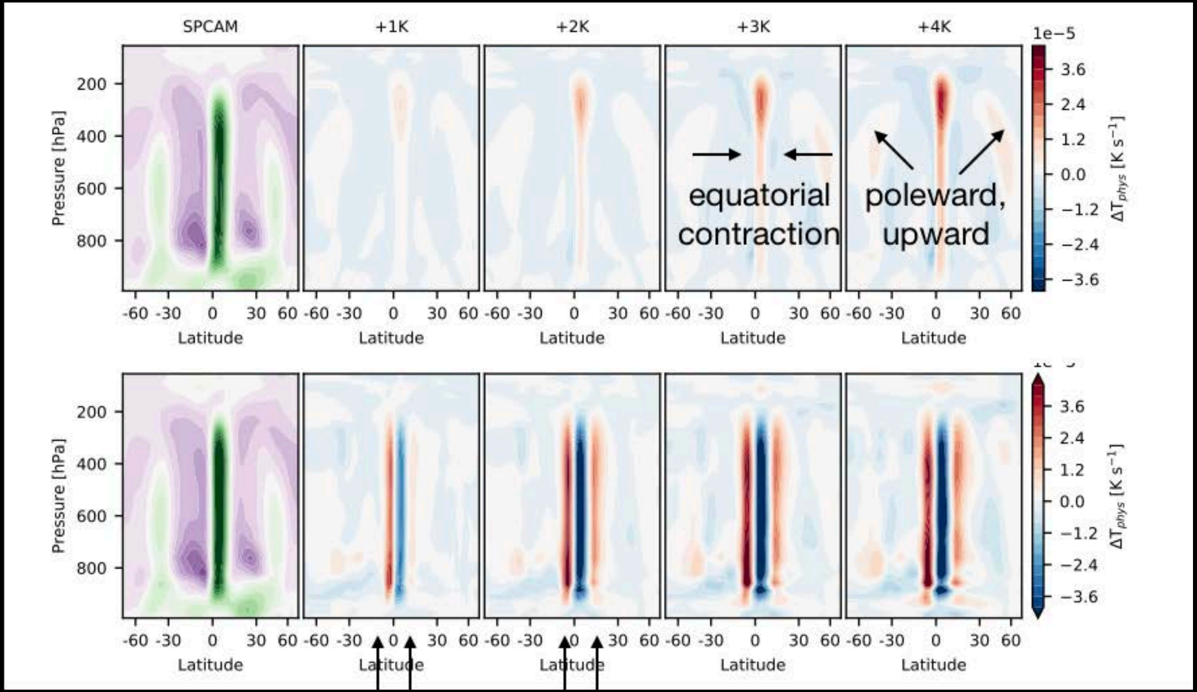
Gentine P., Pritchard M., Rasp S., Reinaudi G., *GRL, 2018*
Rasp, Pritchard and Gentine, *PNAS* 2018
Brenowitz and Bretherton, *GRL*, 2018

Columbia Engineering
The Fu Foundation School of Engineering and Applied Science

Issues

2. **Extrapolation**

ML has mostly been about interpolations using lots of data, poor extrapolation



Gentine P., Pritchard M., Rasp S., Reinaudi G., *GRL, 2018*
Rasp, Pritchard and Gentine, *PNAS* 2018
Brenowitz and Bretherton, *GRL*, 2018

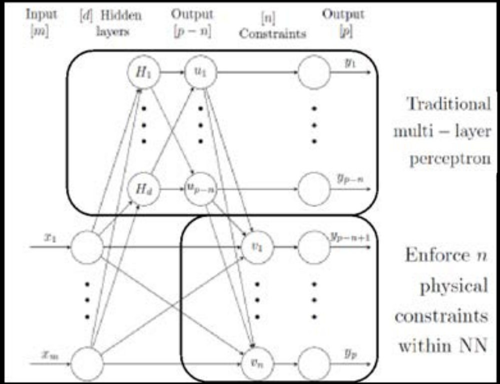
Hybrid approaches

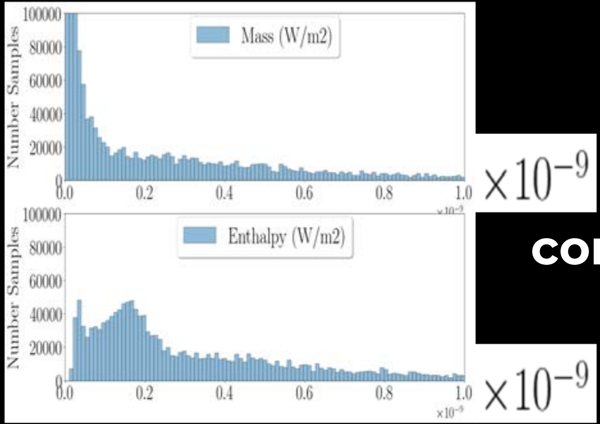
Constraining physics within ML
1. Convection
Energy and mass conservations

Impose them within NN as function of inputs (x) and outputs (y):

$$\left\{ \boldsymbol{C} \begin{bmatrix} x \\ y \end{bmatrix} = 0 \right\}$$

2 equations: reduce NN degrees of freedom to **n-2 degrees of freedom**





Exact conservations

Beucler, Pritchard, Rasp, Gentine, *PRL, submitted*

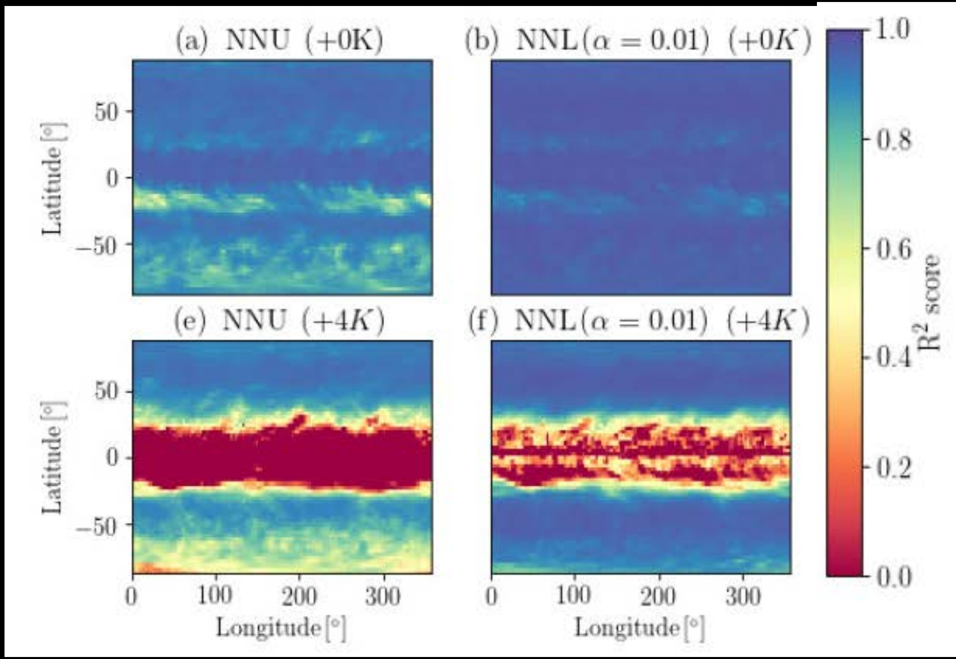
Hybrid approaches

Constraining physics within ML
1. Convection
Energy and mass conservations

Unconstrained Constrained

Current climate

Future climate

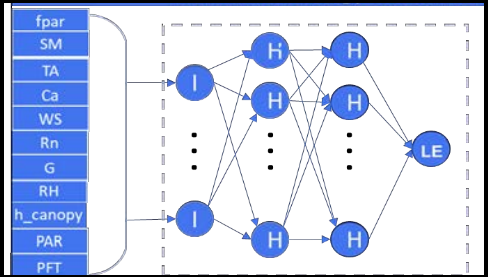


Constrained physics + improved generalization ☺

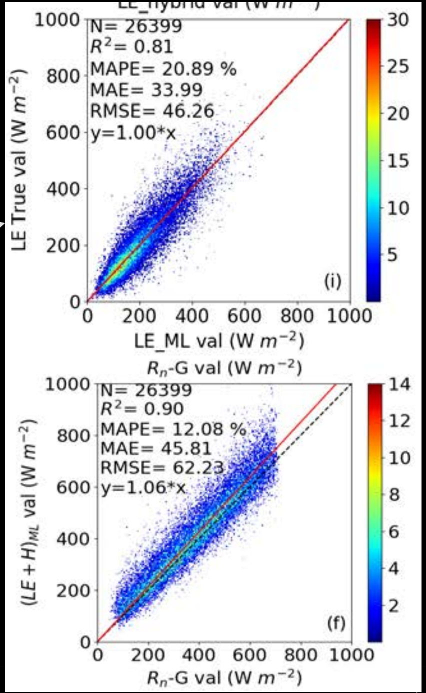
Beucler, Pritchard, Rasp, Gentine, *PRL, submitted*

Constraining physics within ML
2. Land surface latent heat flux (LE)

Objective: predict LE from environmental variables



- Pure ML (feedforward NN) performs well



- But **does not conserve surface energy budget**

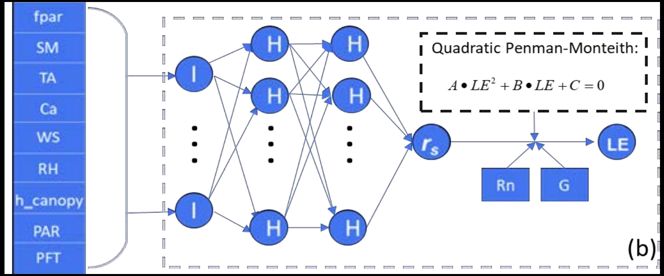
$$R_n - G \neq H + LE$$

Zhao, Gentine, Reichstein, *GRL, submitted*

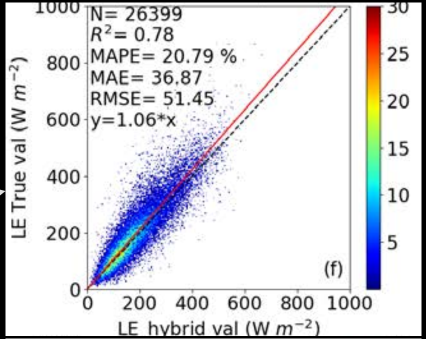
Hybrid approaches

Constraining physics within ML
2. Land surface latent heat flux (LE)

Objective: predict LE + conserve energy + respect diffusion $LE = \rho \frac{e_s - e_a}{r_s + r_a}$

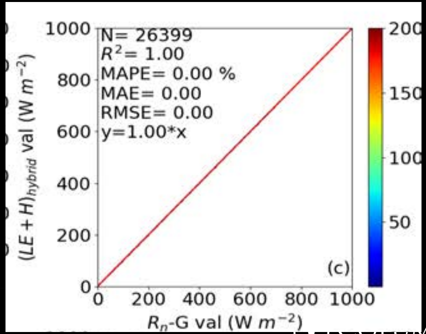


- Hybrid ML performs as well as pure ML



- **Conserves surface energy budget ☺**

$$R_n - G = H + LE$$

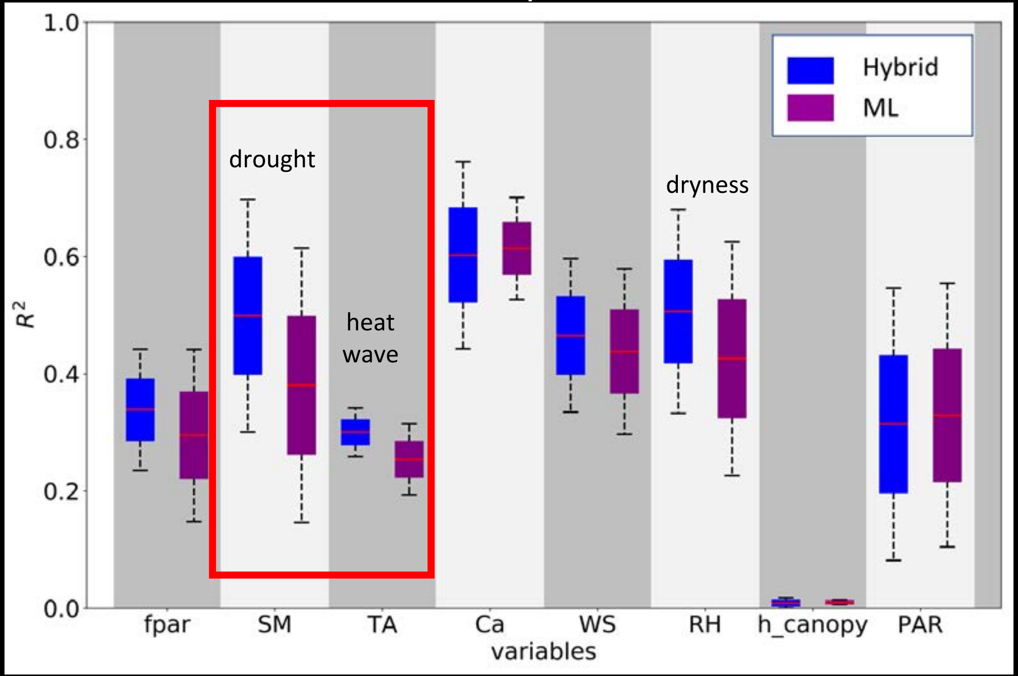


Zhao, Gentine, Reichstein, *GRL, submitted*

Hybrid approaches

Constraining physics within ML
2. Land surface latent heat flux (LE)
Out-of-sample generalization/extremes

Test 1 and 99 percentiles



Hybrid systematically outperforms pure ML for extremes ☺

Zhao, Gentine, Reichstein, *GRL, submitted*

Conclusions

Machine learning is an appealing approach for subgrid parameterizations

Two working examples

1. **Deep clouds**
2. **Land surface processes (evapotranspiration)**

Issues:

1. **Conservations, physical invariances, physical laws**
2. **Generalization**

Solution:

Hybrid physical+ML approaches appear as powerful tools to tackle this

THANK YOU

Questions?