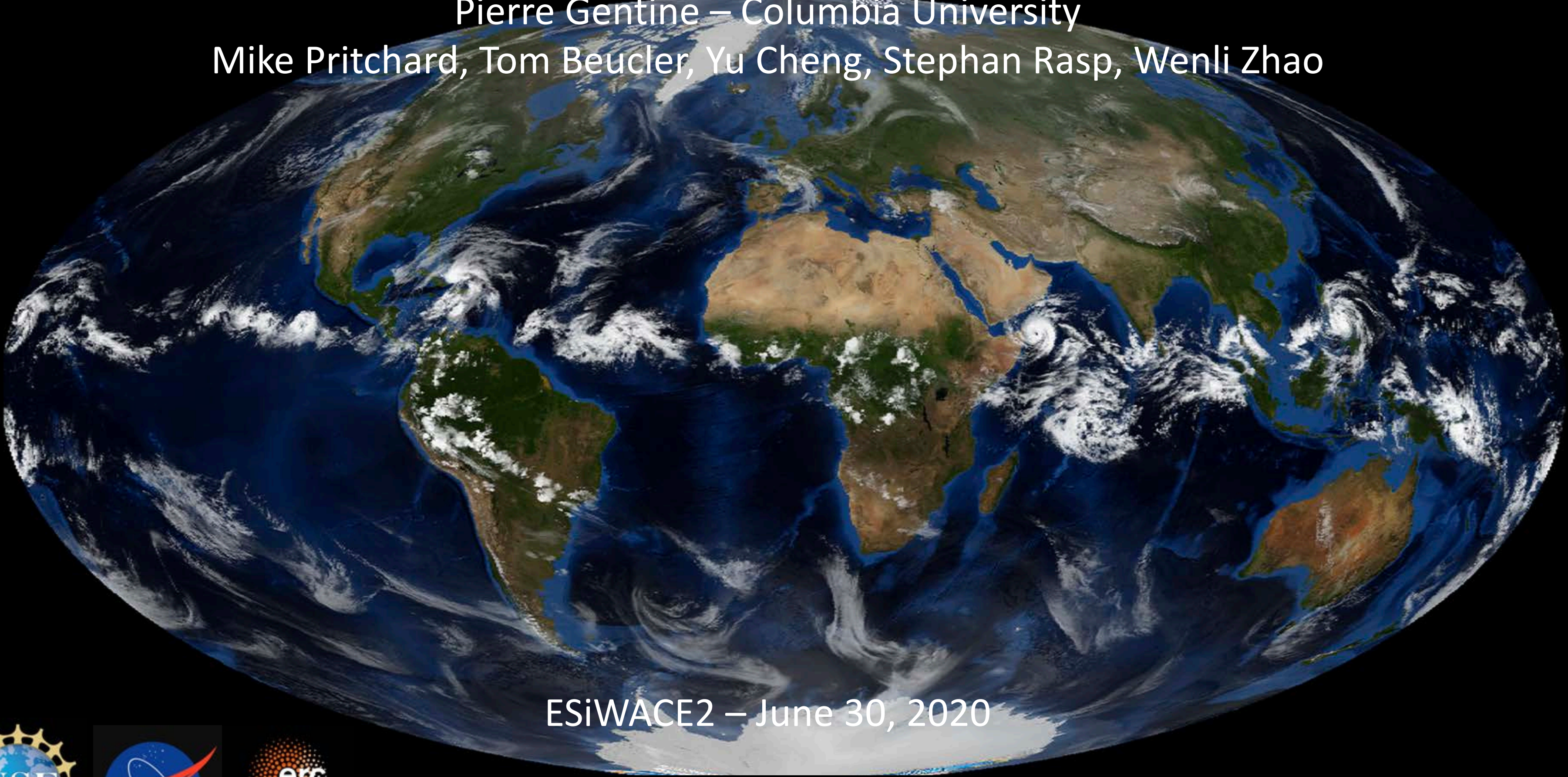
Physics-Guided Machine Learning

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



ESiWACE2 – June 30, 2020

European Research Council
Established by the European Commission

TRANSCENDING DISCIPLINES, TRANSFORMING LIVES

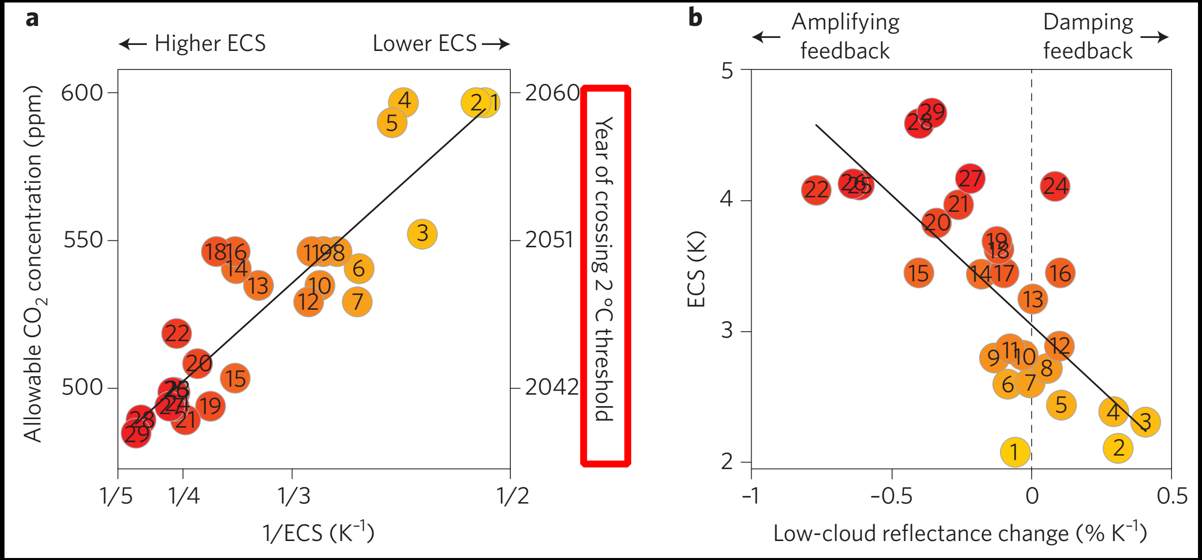
COLUMBIA | ENGINEERING
The Fu Foundation School of Engineering and Applied Science

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 **representation of clouds**



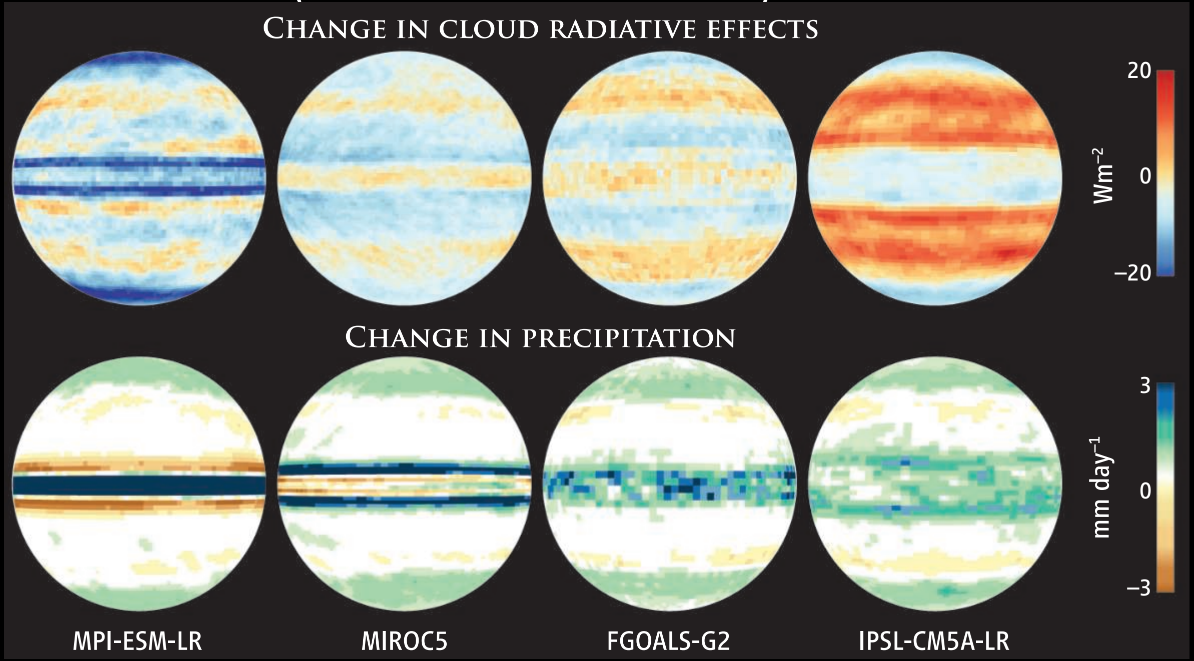
ECS = Equilibrium climate sensitivity (T response do CO_2 doubling)

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

Regional climate sensitivity

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

Aquaplanet
+4K
(no SST
feedback!)



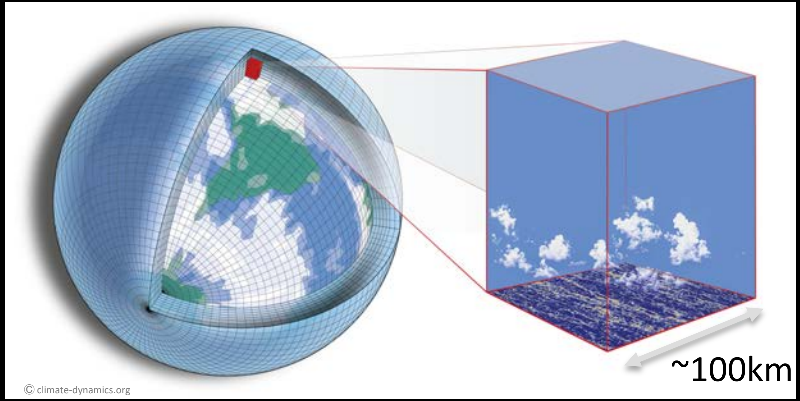
Regional climate projection is too uncertain

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

Using ML for climate

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 projection**

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

Using ML for climate

Parameterization: Difficulty

- Many orders of magnitude in scales: mm to 10^4 km

mm — m — km — 10^4 km

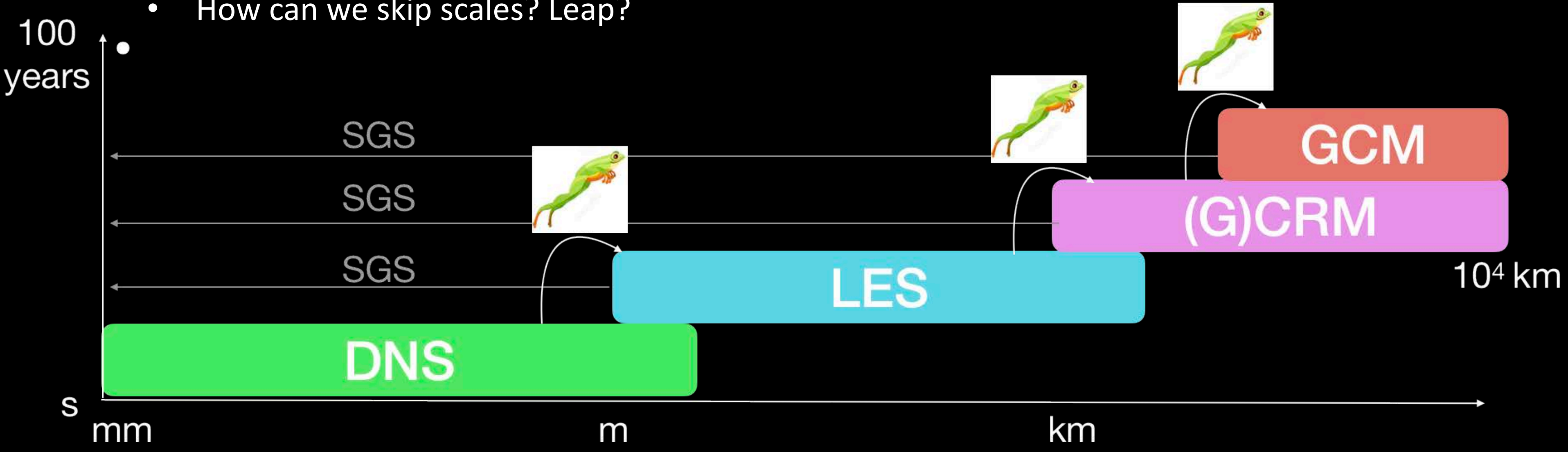
- Major numerical challenge for a long time to come
  (not just cloud resolving)

- How can we **buy us time? and (hopefully) learn on the way?**

Using ML for climate

Resolving scales in the atmosphere

- We can now resolve many processes
- Limited time and domain size + need subgrid scale (SGS) model
- How can we skip scales? Leap?

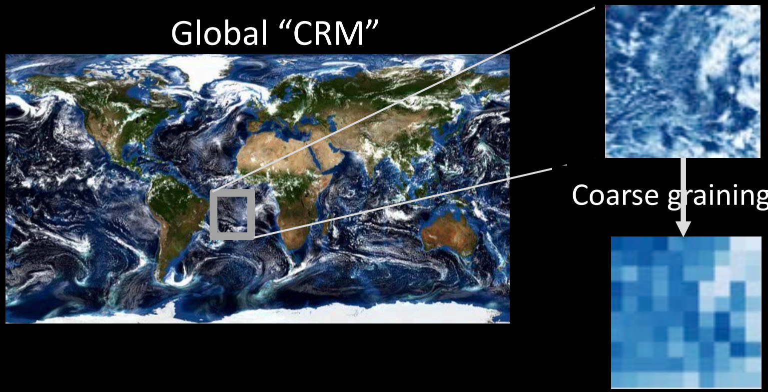


Using ML for climate: (deep) clouds

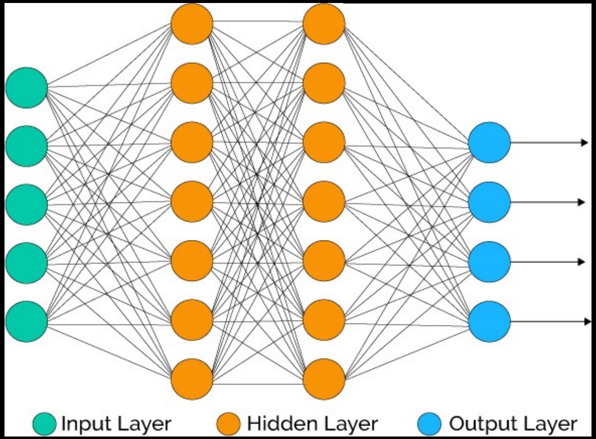
How can we solve this issue?
Take advantage of **cloud-resolving simulations**
(~1km, **alleviate most biases** but very expensive)

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

Global “CRM”



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



$\frac{\partial \overline{T}}{\partial t}\Big|_{\text{convection}}$

$\frac{\partial \overline{q}}{\partial t}\Big|_{\text{convection}}$

Precipitation

Deep Neural Net or Convolutional NN

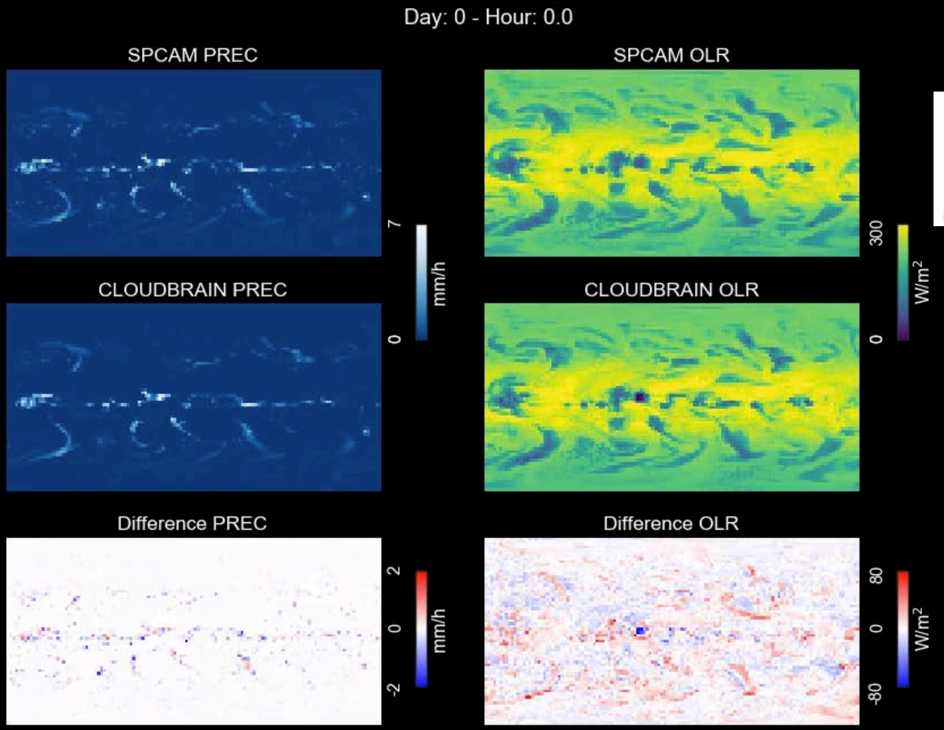
Cost function: misfit to coarse-grained high-res. model

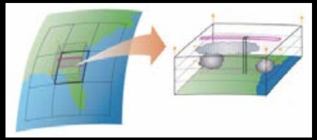
Gentine P., Pritchard M., Rasp S., Reinaudi G., *GRL, 2018*
Rasp, Pritchard and Gentine, *PNAS* 2018
Brenowitz and Bretherton, *GRL*, 2018

Using ML for climate: (deep) clouds

Coarse-grained
Cloud-resolving
Model
(superparameterization)

Machine
learning
Coarse-resolution
model



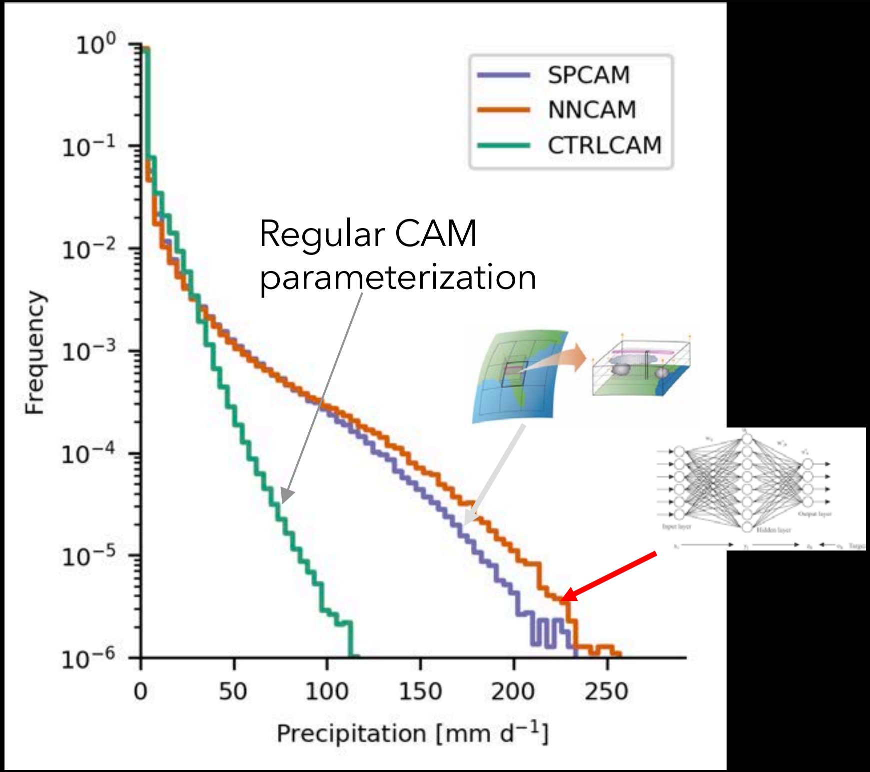


10 times cheaper than original coarse model, 1000 less expensive than high-res model

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

Using ML for climate: (deep) clouds



Good hydrologic cycle

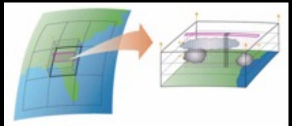
Interactive model

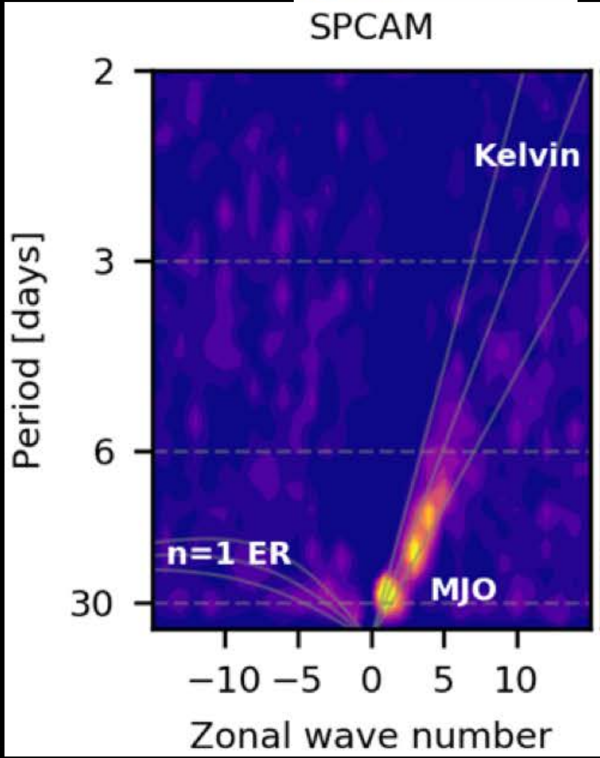
Now we have good boundary condition to study hydrology ☺

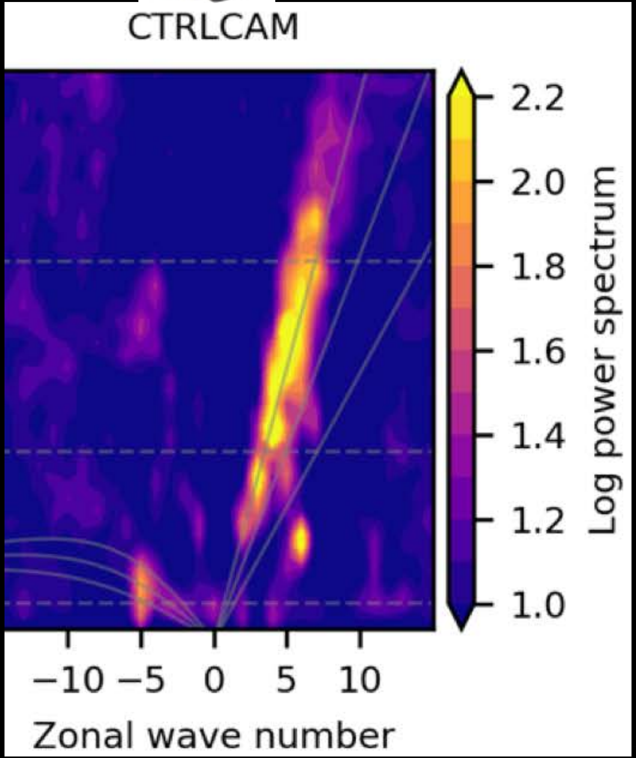
Gentine P., Pritchard M., Rasp S., Reinaudi G., *GRL, 2018*
Rasp, Pritchard and Gentine, *PNAS* 2018

Using ML for climate: (deep) clouds

Spectra



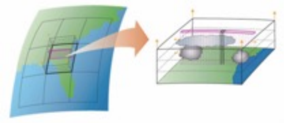


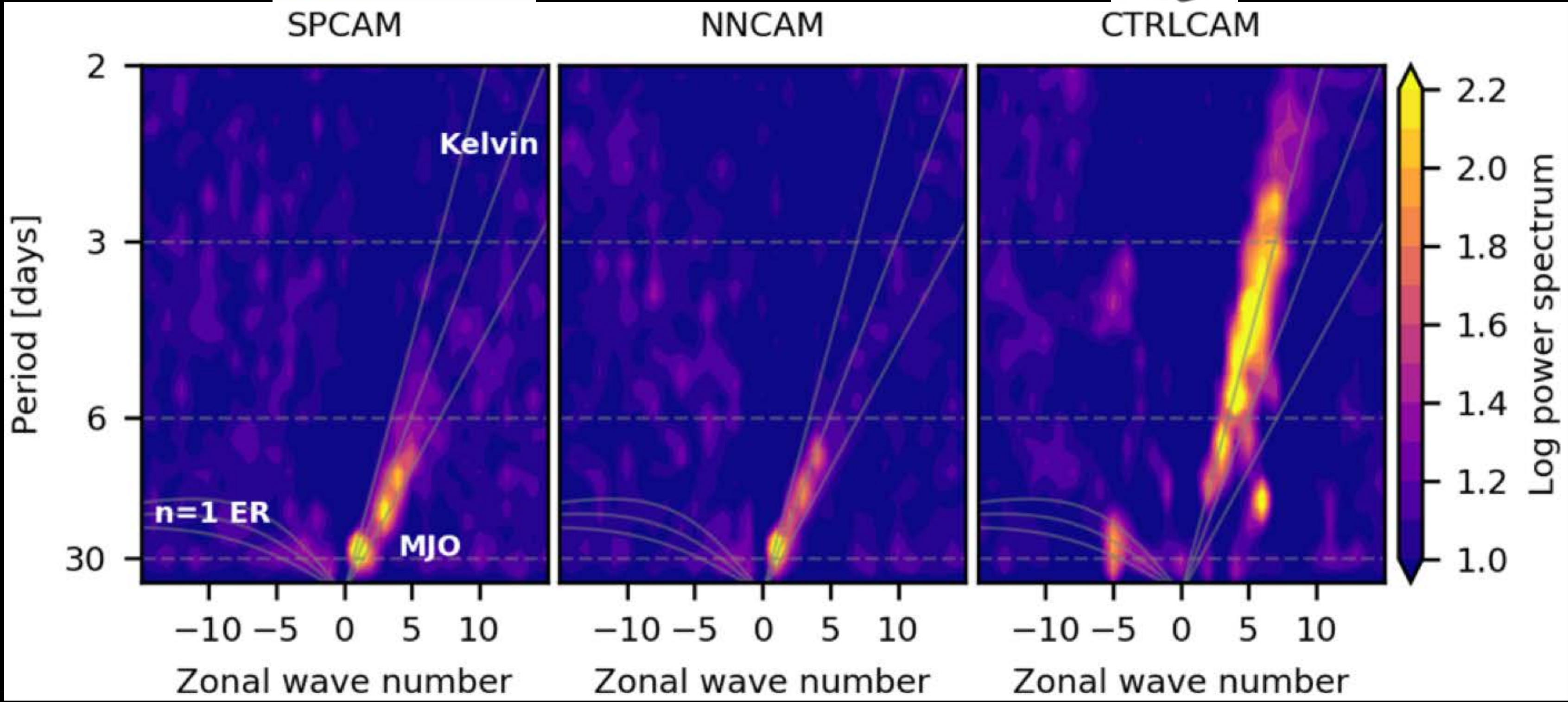


Gentine P., Pritchard M., Rasp S., Reinaudi G., *GRL, 2018*
Rasp, Pritchard and Gentine, *PNAS* 2018

Using ML for climate: (deep) clouds

Spectra





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

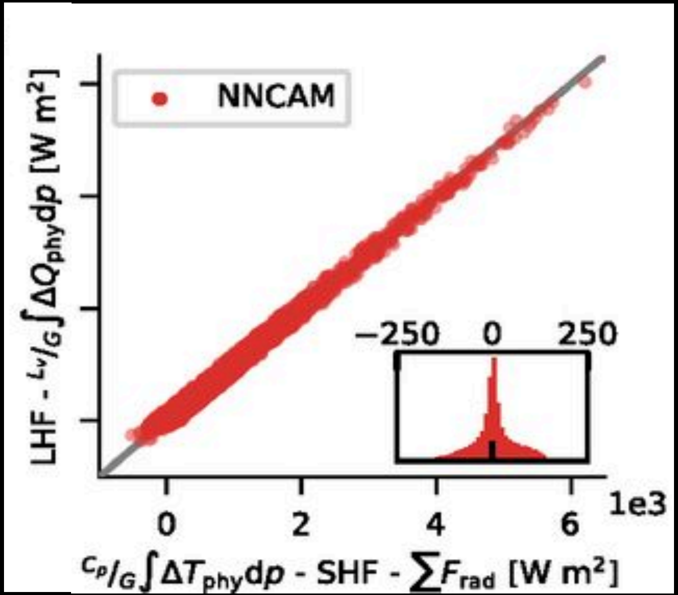
Issues

1. Physical Constraints

Energy conservation

Mass conservation

Only approximate with ML

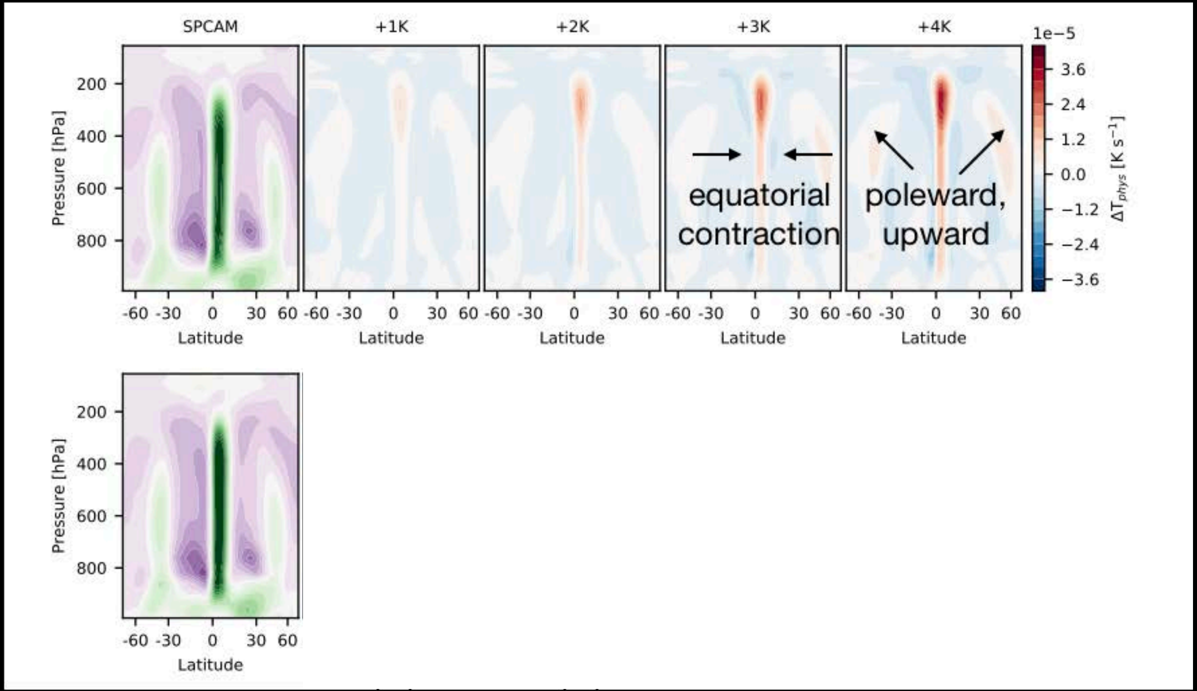


Gentine P., Pritchard M., Rasp S., Reinaudi G., *GRL, 2018*
Rasp, Pritchard and Gentine, *PNAS* 2018

Issues

2. **Generalization**

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

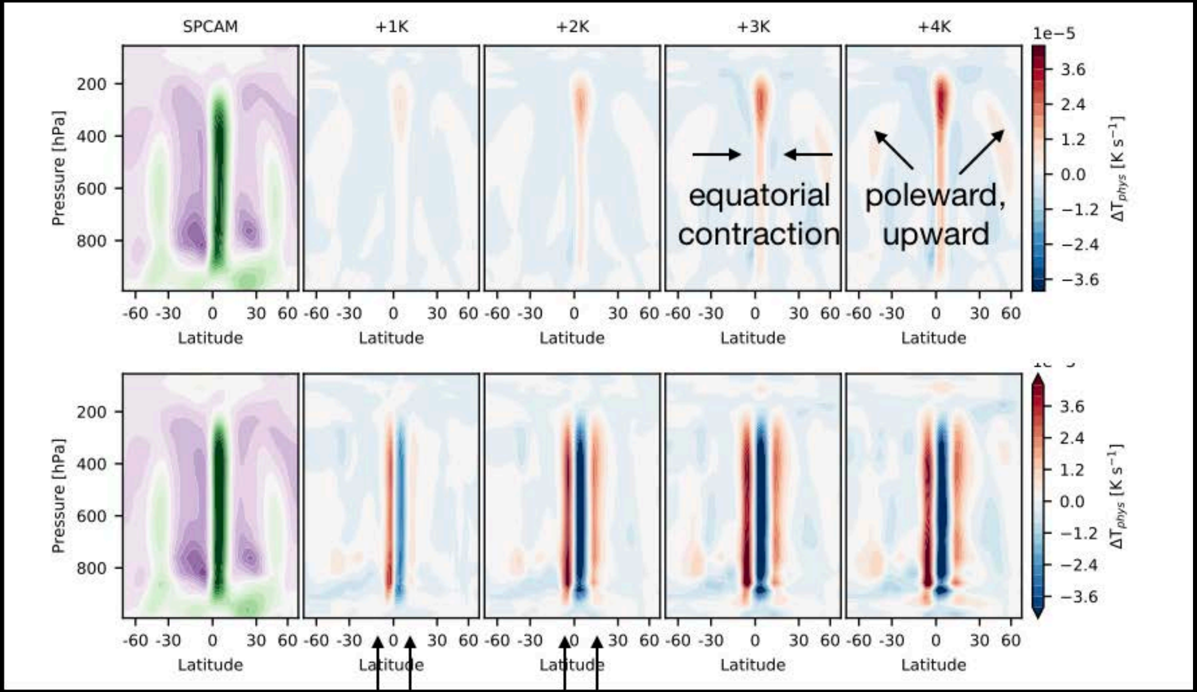


Gentine P., Pritchard M., Rasp S., Reinaudi G., *GRL, 2018*
Rasp, Pritchard and Gentine, *PNAS* 2018
Brenowitz and Bretherton, *GRL*, 2018

Issues

2. **Generalization**

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



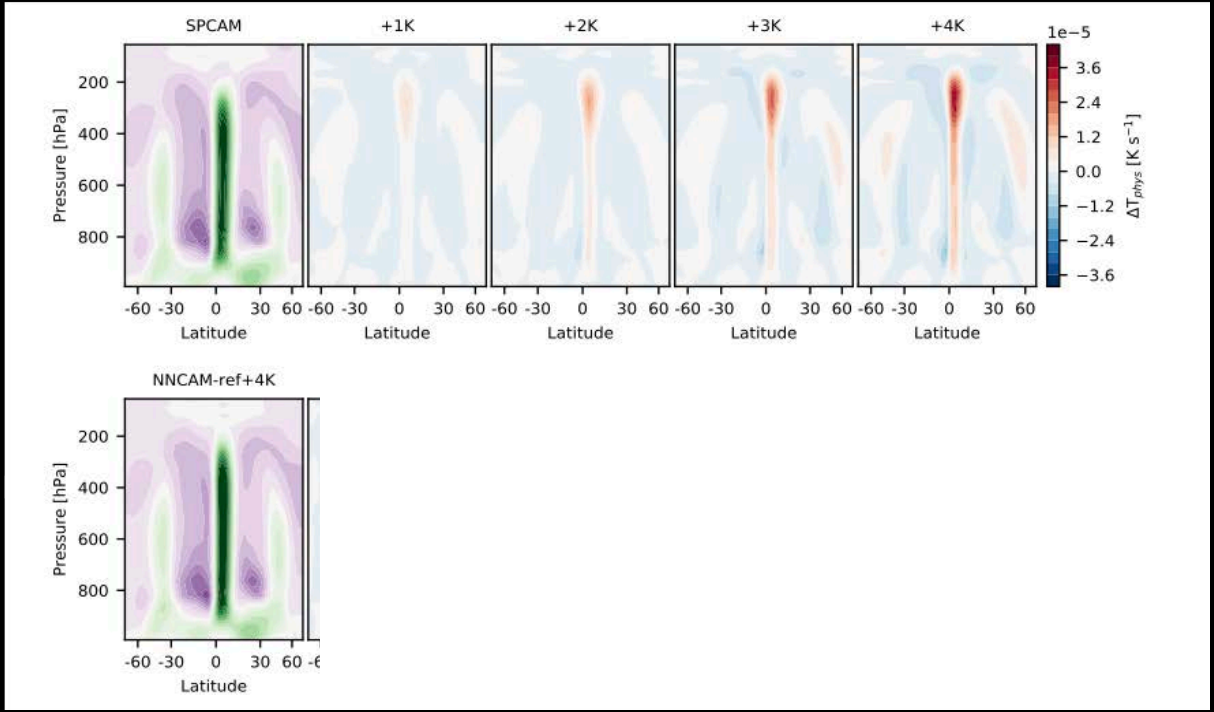
Gentine P., Pritchard M., Rasp S., Reinaudi G., *GRL, 2018*
Rasp, Pritchard and Gentine, *PNAS* 2018
Brenowitz and Bretherton, *GRL*, 2018

Issues

2. **Generalization**

Interpolating works

Using both 0K and +4K



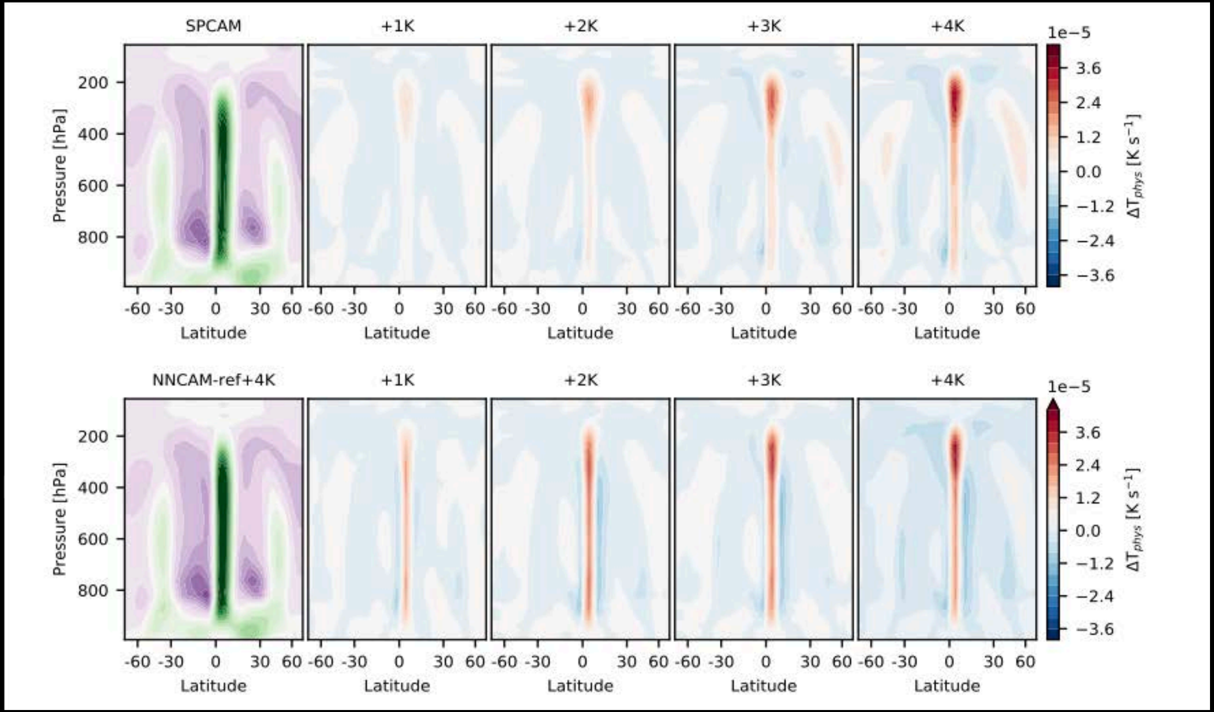
Gentine P., Pritchard M., Rasp S., Reinaudi G., *GRL, 2018*
Rasp, Pritchard and Gentine, *PNAS* 2018
Brenowitz and Bretherton, *GRL*, 2018

Issues

2. **Generalization**

Interpolating works

Using both 0K and +4K



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

Summary of issues with brute force ML

1. **Do not respect physical laws**

e.g. conservation of energy and mass

→ strict requirement

2. **Issue with out-of-sample generalization**

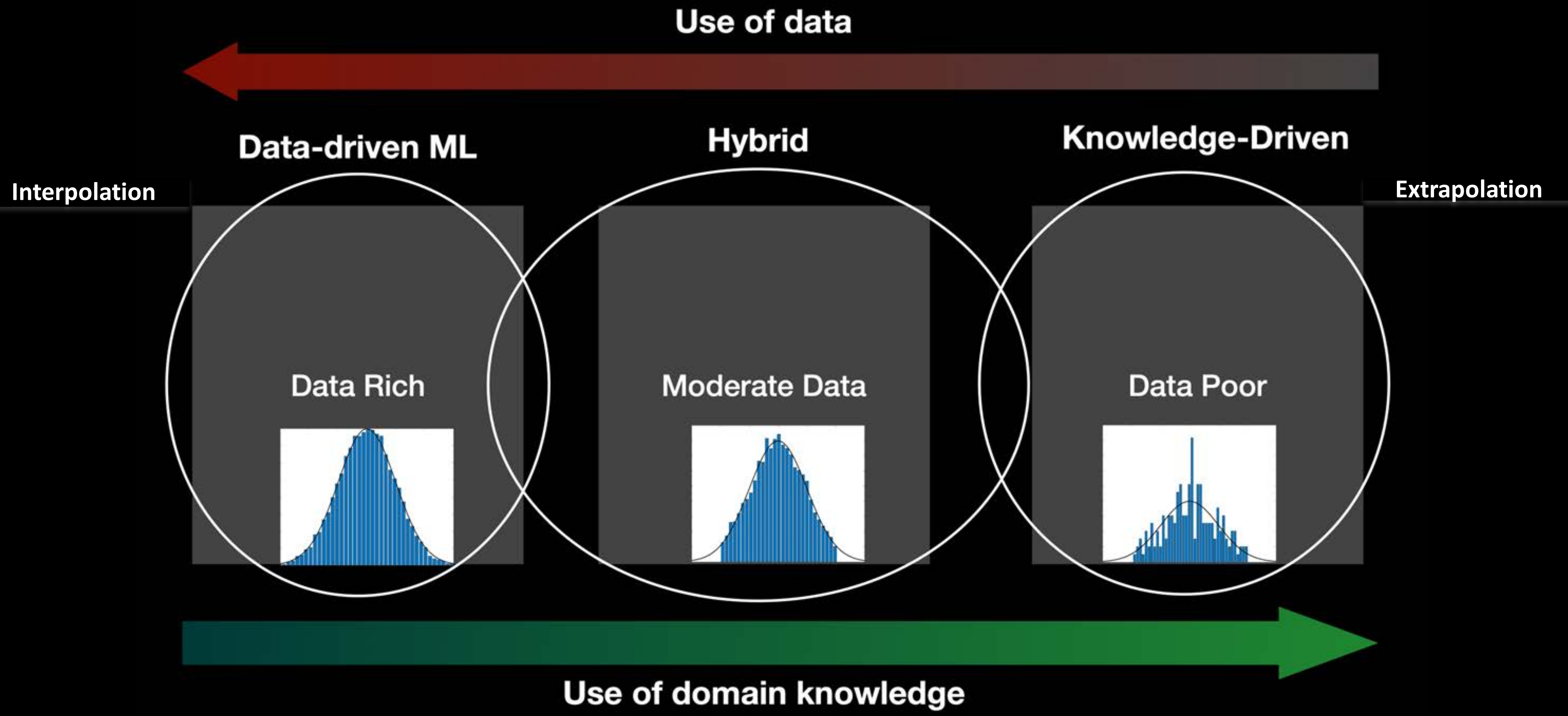
Important for many climate applications

e.g. extremes, climate change

input → **ML algorithm** → Within range prediction

input → **ML algorithm** → Out-of-sample prediction

Potential Overcoming Strategies



For knowledge-driven see Yang… Gentine 2019 *ERL*, or Jia,…, Kumar 2018 *ArXiv*

Hybrid approaches

Constraining physics within ML

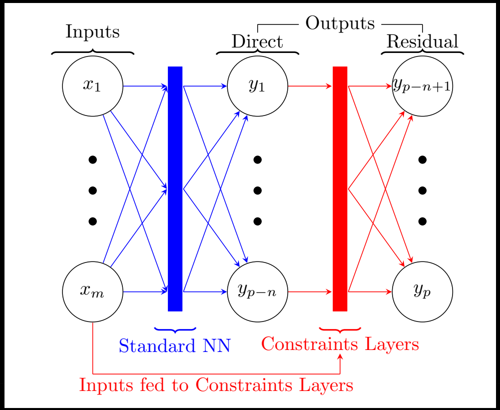
1. Convection

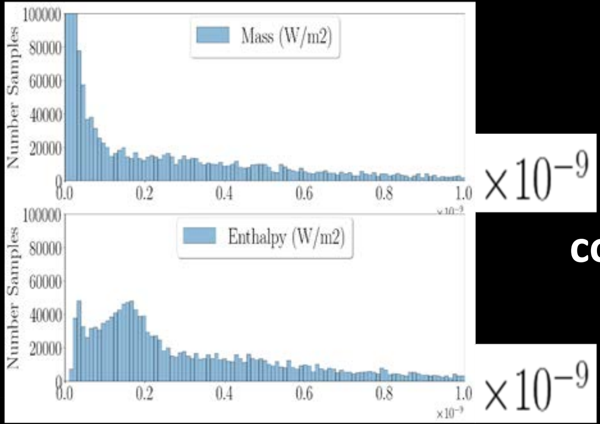
Energy and mass conservations

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

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

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



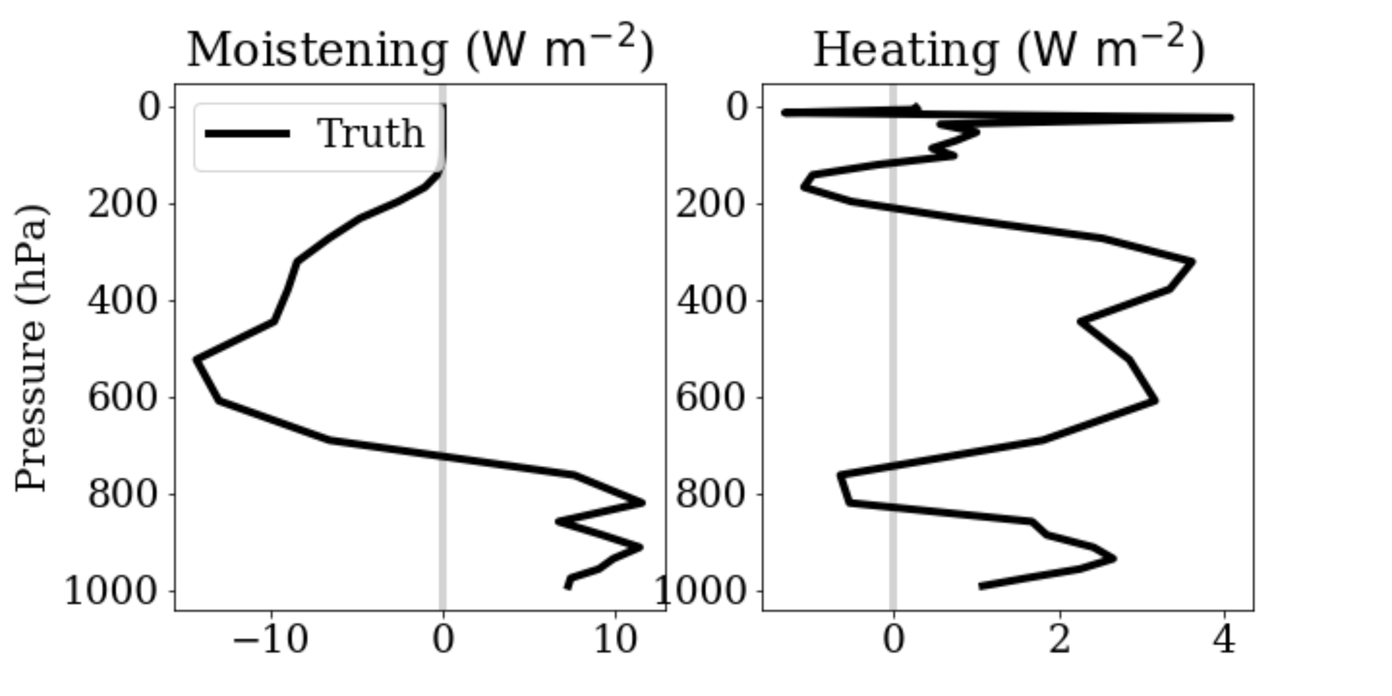


Exact conservations

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

Generalization

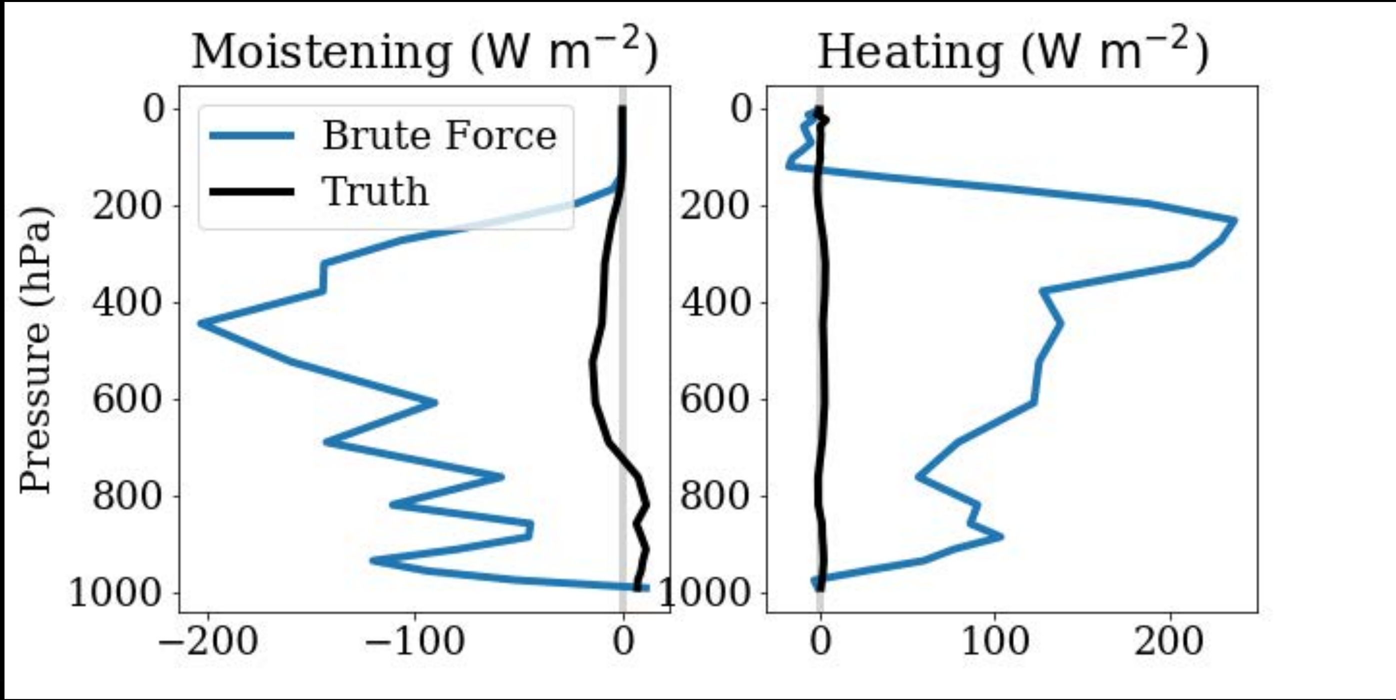
Warm climate +8K generalization experiment



Beucler, Pritchard, Rasp, Gentine, *in preparation*

Generalization

Warm climate +8K generalization experiment
Pure ML (deep NN)

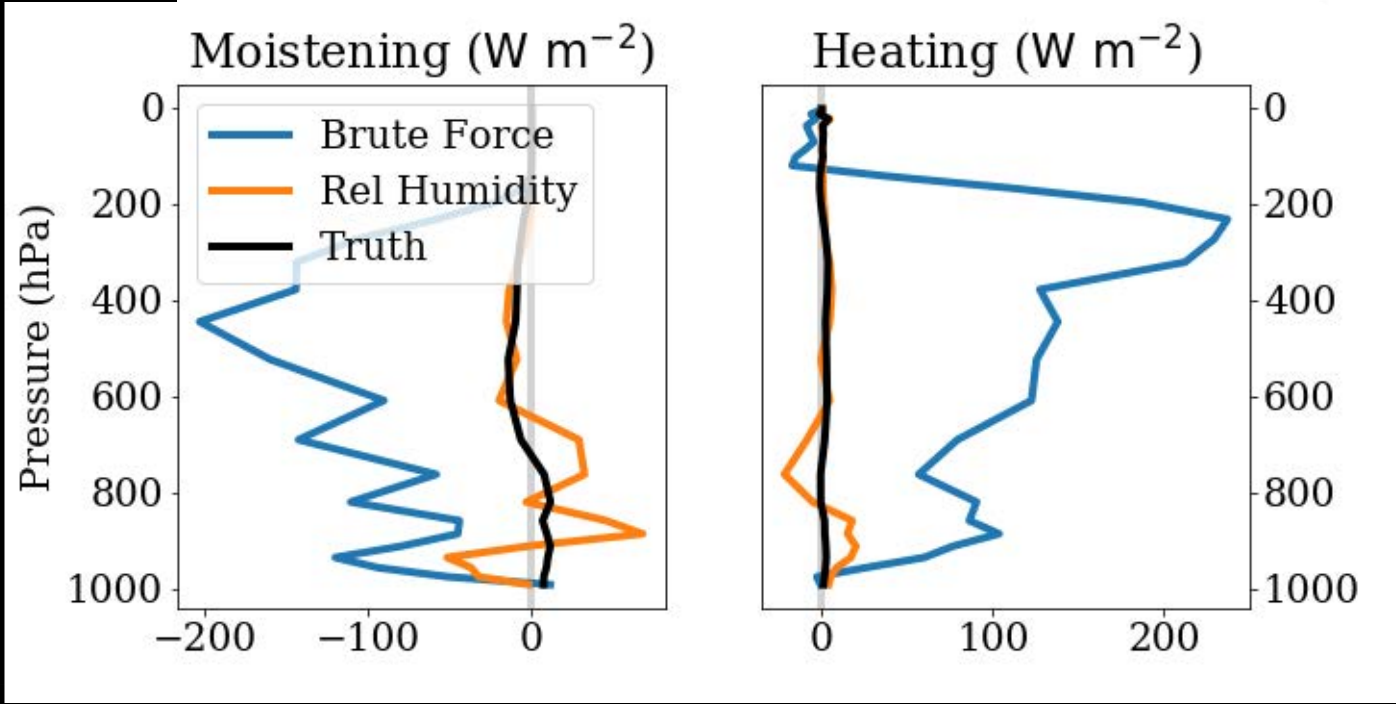


Beucler, Pritchard, Rasp, Gentine, *in preparation*

Generalization

Use physical knowledge

$$\text{Relative humidity}\,(z) = \frac{\text{Partial water vapor pressure}\,(z)}{\boxed{\text{Saturation water vapor pressure}\,(T, p)}}$$

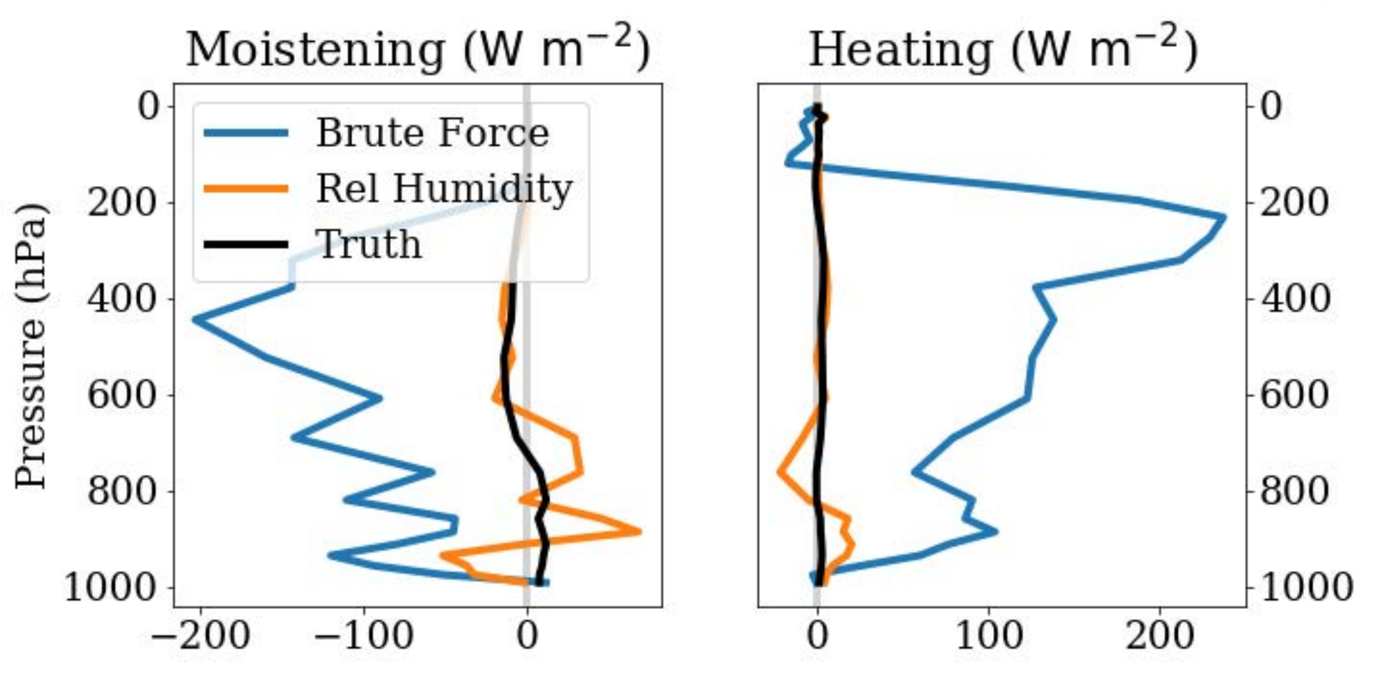


Beucler, Pritchard, Rasp, Gentine, *in preparation*

Generalization

Use physical knowledge

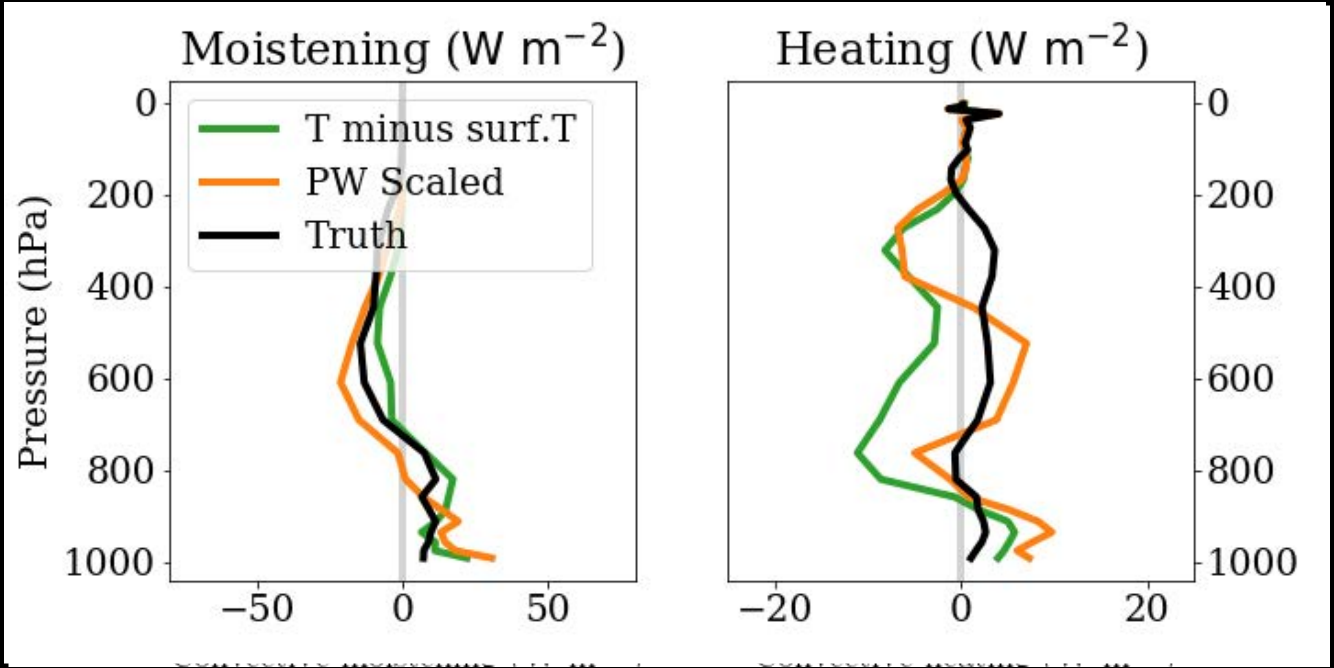
$$\text{Relative humidity}\,(z) = \frac{\text{Partial water vapor pressure}\,(z)}{\boxed{\text{Saturation water vapor pressure}\,(T,p)}}$$



Beucler, Pritchard, Rasp, Gentine, *in preparation*

Hybrid approaches

Using physical knowledge – … – output flux rescaling
Further improvements



Constrained physics + improved generalization ☺

Beucler, Pritchard, Rasp, Gentine, *in preparation*

Conclusions

Machine learning is an appealing approach for subgrid parameterizations

Working example
Deep clouds (convection)

Issues:

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

Solution:
Hybrid physical+ML approaches appear as powerful tool to tackle this

THANK YOU

Questions?

pg2328@columbia.edu

@PierreGentine

COLUMBIA | ENGINEERING
The Fu Foundation School of Engineering and Applied Science

A hope