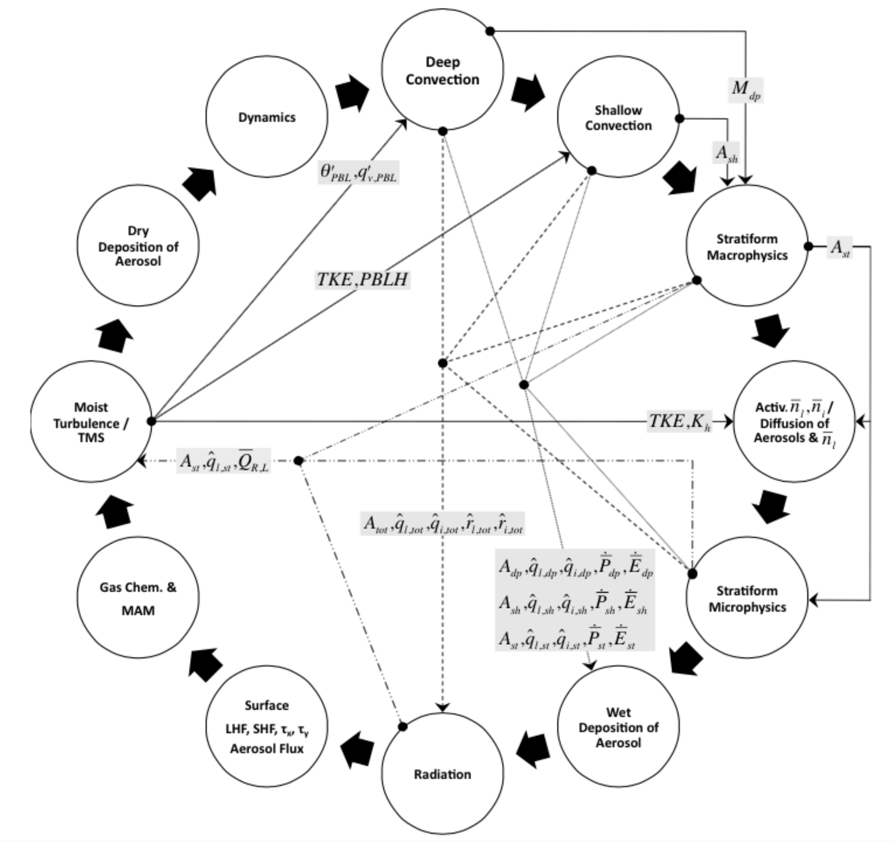
Machine Learning for Moist Physics Parameterization

Christopher S. Bretherton
Departments of Atmospheric Science and Applied Mathematics,
University of Washington

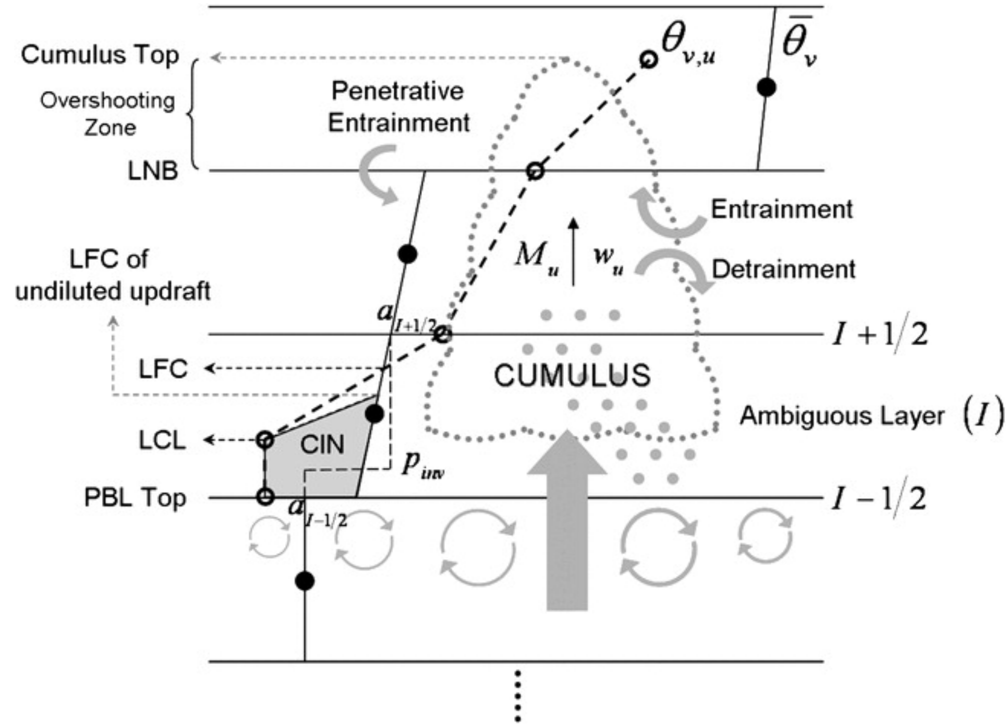
Atmospheric Sciences
University of Washington

A single grid column of a global climate model (100 x 100 km)

Representing such unresolved variability is challenging



Park et al. 2014

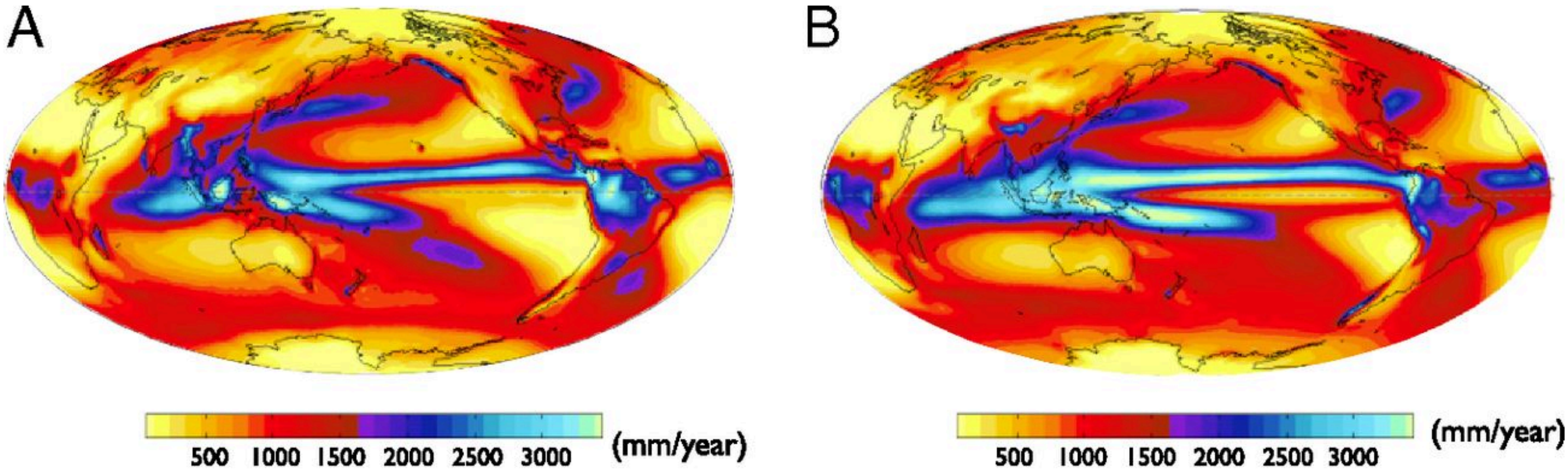


Park and Bretherton 2009

···so ‘parameterizations’ of moist physics are difficult to develop, subjective and work imperfectly.

Feeds into weather/climate model biases & uncertainties

Example: Double-ITCZ rainfall bias in climate models



GPCP
(observations)

CMIP5
(models)

Subgrid parameterization improvements have come more through ‘targeted tweaking’ than from fundamental advances.

Cloud processes are better simulated on fine grids

Giga-LES (*Khairoutdinov et al. 2009*)
100x100 km with 100 m grid
Figure courtesy Ian Glenn U. Utah

1 km global simulations are (briefly) possible

MIYAMOTO ET AL.: CONVECTION IN A SUB-KM GLOBAL SIMULATION

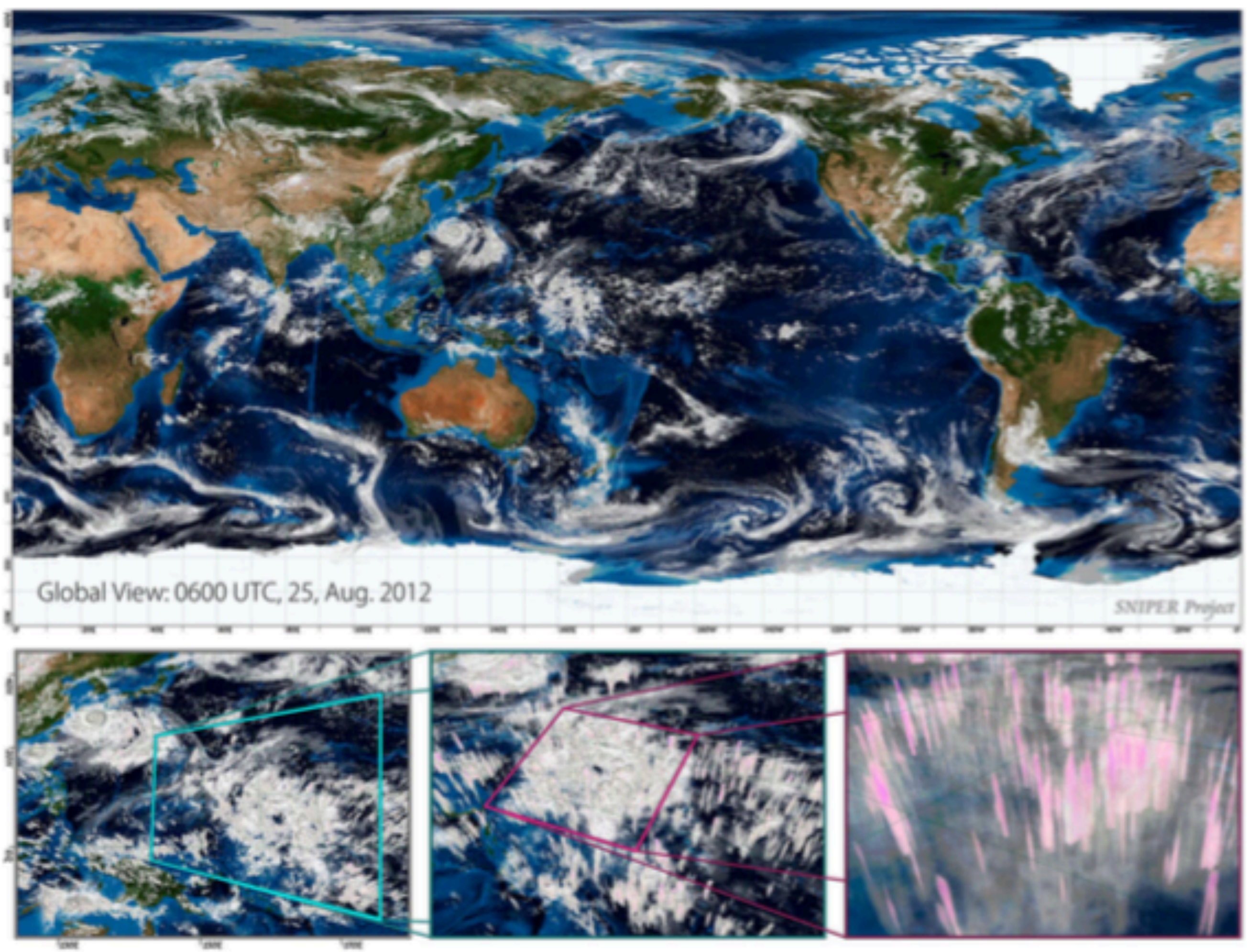


Figure 1. (top) Horizontal view of the total mixing ratio of condensed water contents in Δ0.87, (bottom left) close-up view of the northwestern Pacific, (bottom middle) a further close-up view for a cloud cluster, and (bottom right) an extreme close-up of an active convection region. The pink color indicates the hydrometeor density larger than $2\,g\,kg^{-1}$. Topography and bathymetry are Blue Marble (August) by Reto Stöckli, NASA Earth Observatory.

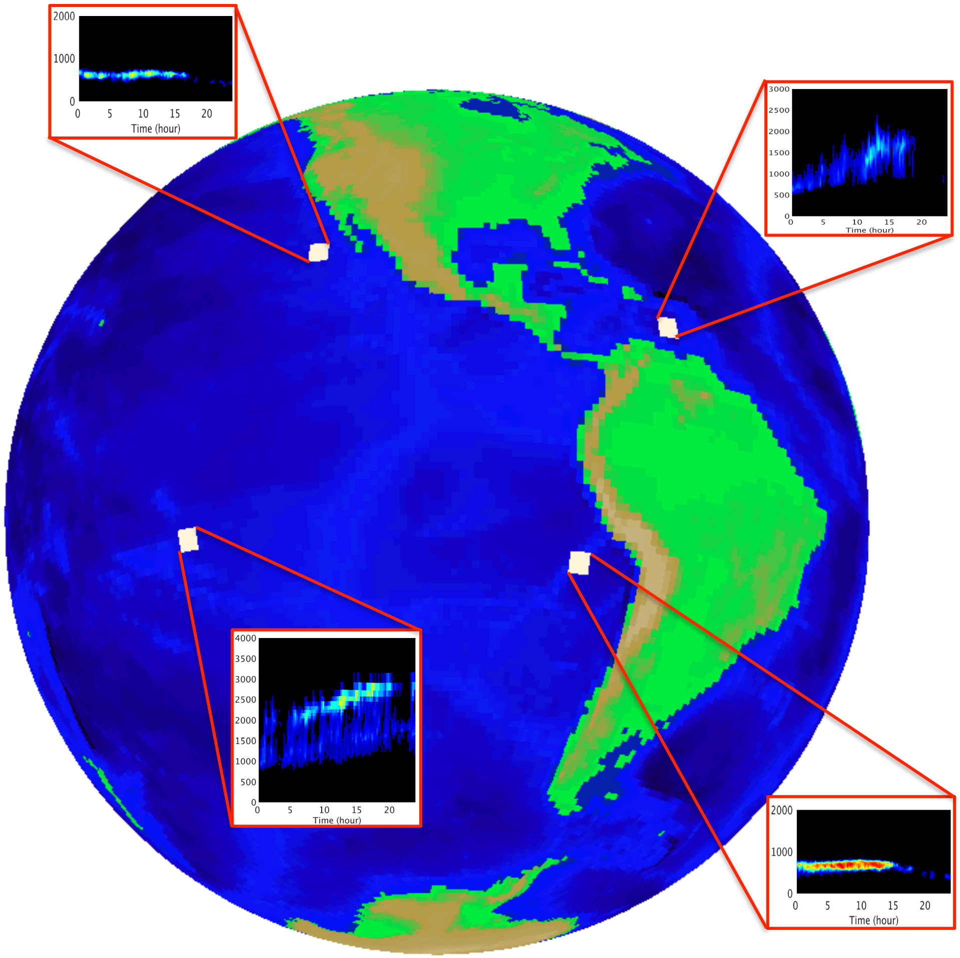
NICAM

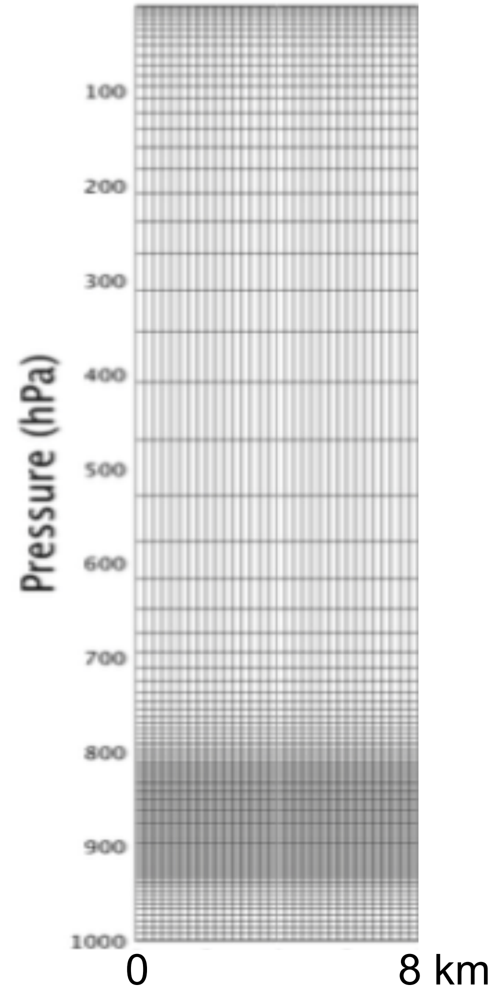
3.5-14 km (mo-yr)
Tomita et al. 2005

0.9 km (for 24 hr):
Miyamoto et al. 2013

Multiyear simulations with 'ultraparameterization'

Ultrafine-grid 2D cloud-resolving model in each GCM grid column (Δx = 250 m, Δz = 20 m for z=0.5-2 km)





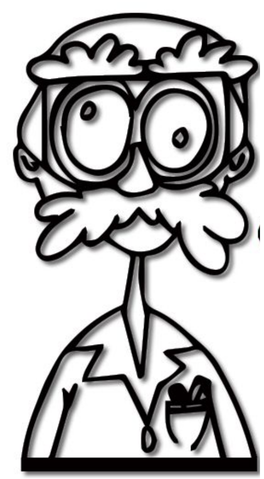
Parishani et al. 2017

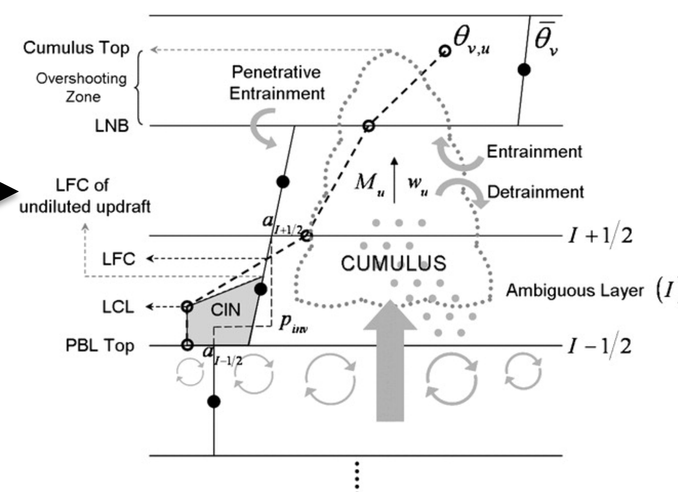
Can hi-res simulations help subgrid parameterizations?

- High-resolution models have progressed faster than moist physics parameterizations in GCMs.
- They are conceptually simpler because cloud properties and air velocity vary less within a grid cell.
- Must still parameterize smaller-scale processes (microphysics, turbulence, aerosols).
- Like other models, they are works in progress needing constant testing vs. observations.
- Still, high-resolution simulations provide realistic reference datasets for parameterizing subgrid cloud process variability
- Since 1992, GCSS program has pushed this approach (Randall et al. 2003).

The human bottleneck

- Improved parameterizations of cumulus convection, turbulence, and cloud microphysics have been implemented in leading weather and climate models
- But slow uptake of new insights from hi-res modelling and observations because humans concoct parameterizations.





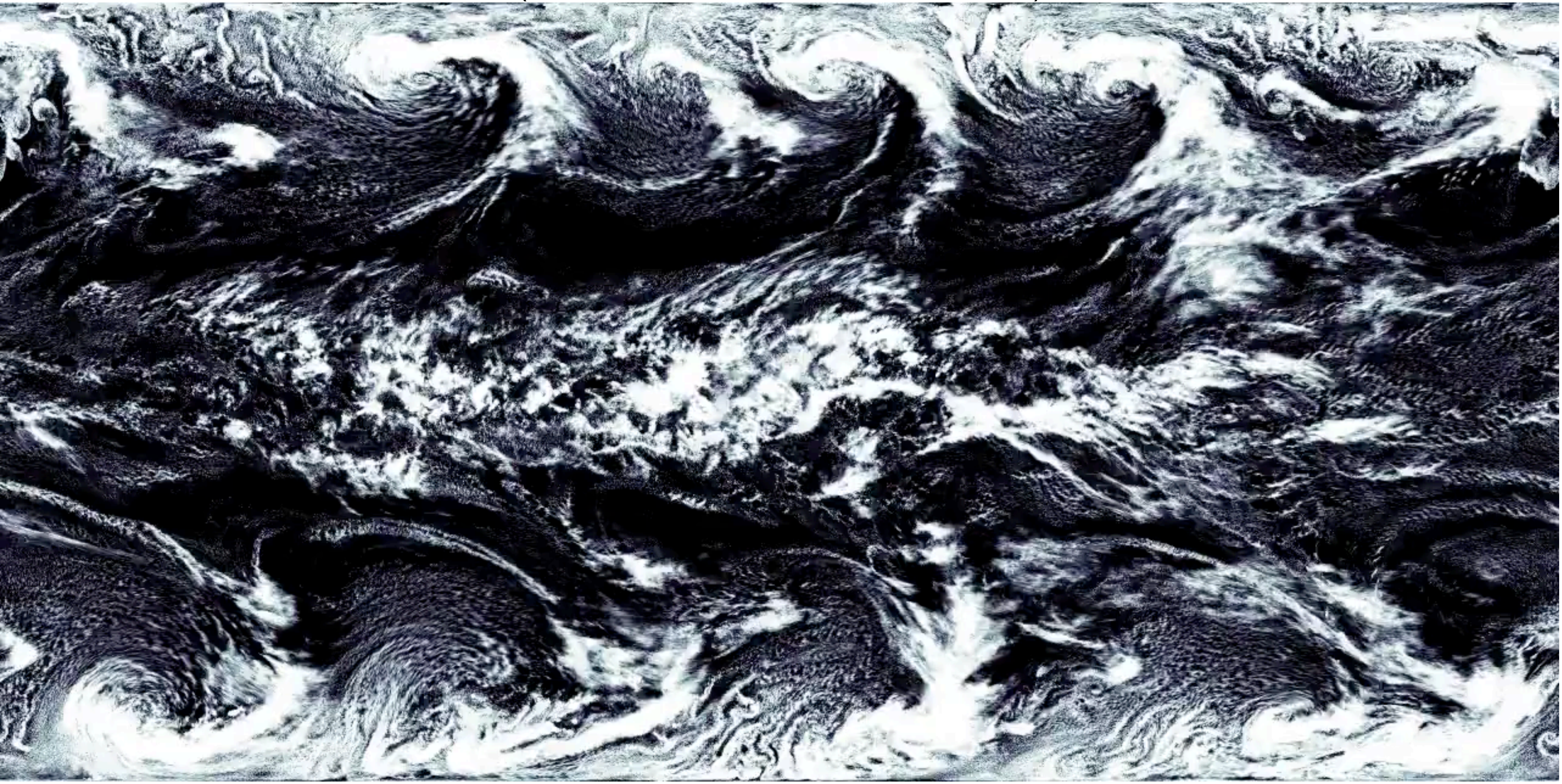
So, how about machine learning (ML)?

- Parameter optimization in existing schemes
  - e. g. Ollinaho et al. (2013)
- Acceleration/regularization of existing schemes
  - e. g. Krasnopolsky (2005), O'Gorman and Dwyer (2018)
- Coarse-graining: Train coarse-grid parameterization from fine-grid simulations (and observations)
  - Brenowitz and Bretherton (2018, submitted to *GRL*), inspired by Krasnopolsky et al. (2013)

Schneider et al. (2017): Grand ML-based vision for ESM2.0

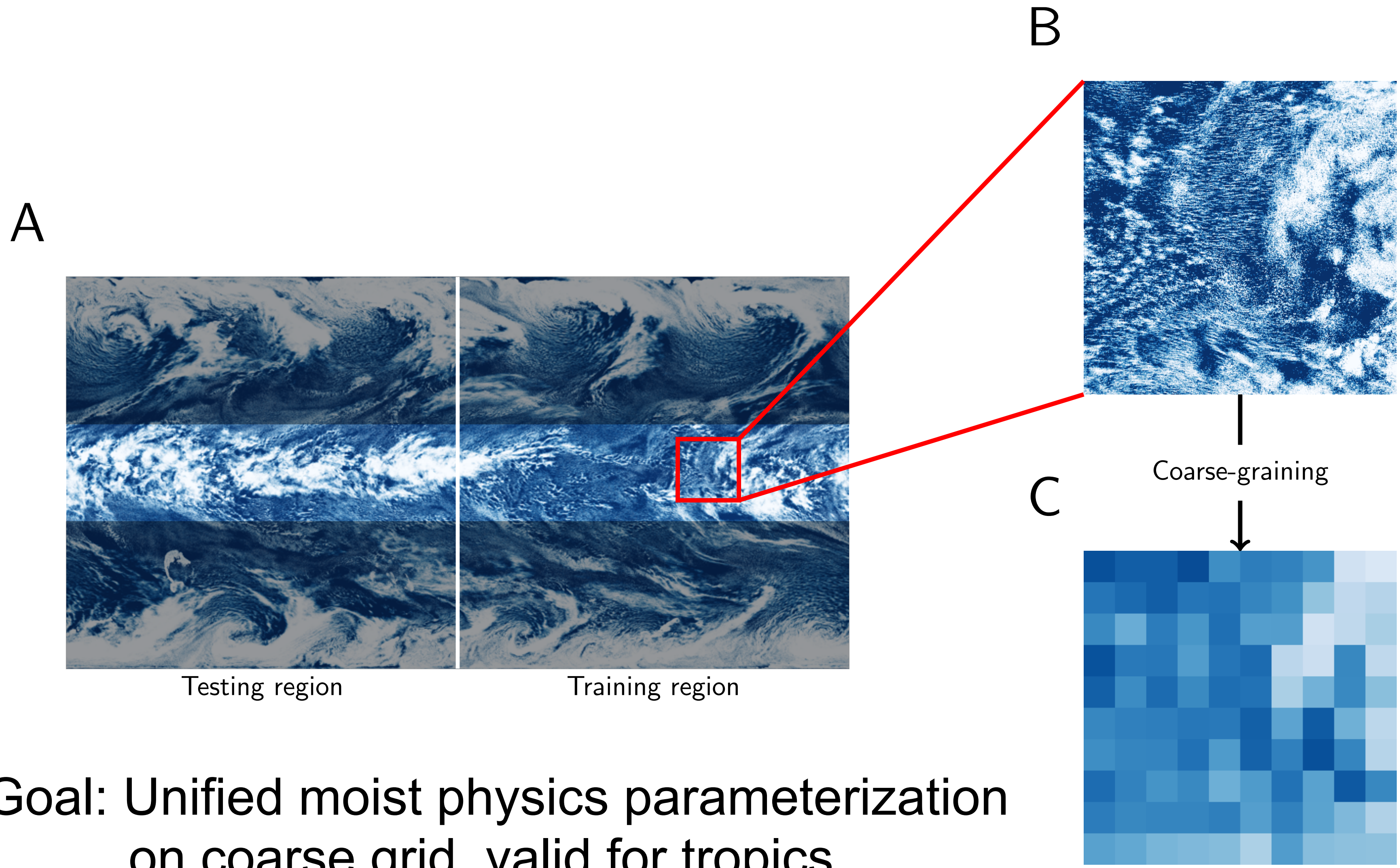
BB18: Show potential of ML-based moist physics parameterization trained on a near-global cloud-resolving model in a non-trivial but simplified context.

Near-global aqua-planet simulation generated by the System for Atmospheric Modeling ($\Delta x = 4$ km)
(Bretherton and Khairoutdinov 2015)



80 days of 3 hourly 3D outputs (5120 x 2560 x 32 = 5×10^8 cells)

Coarse-grain data to (160 km)^2 boxes



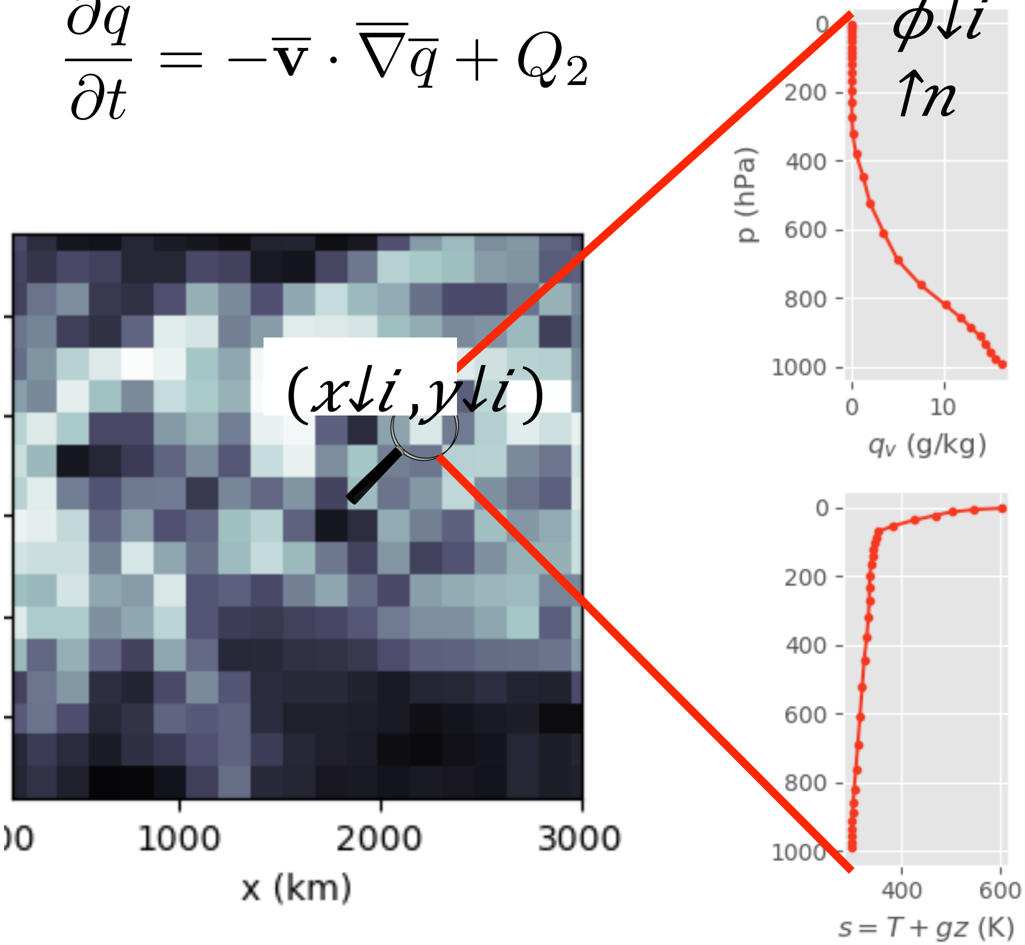
Goal: Unified moist physics parameterization on coarse grid, valid for tropics.

Machine learning inputs and outputs

$$\frac{\partial \overline{s}}{\partial t} = -\overline{\mathbf{v}} \cdot \overline{\nabla s} + Q_1$$

$$\frac{\partial \overline{q}}{\partial t} = -\overline{\mathbf{v}} \cdot \overline{\nabla q} + Q_2$$

Inputs: ϕ↓i ↑n and α↓i↑n =[■SOLIN@LHF@



Output: [■Q↓1,i↑n @Q↓2,i

ML model: [■Q↓1 @Q↓2]=f(ϕ,α) + r

$$\frac{d\phi}{dt} = \underbrace{f(\phi;\alpha)}_{\text{Model}} + \underbrace{g(t)}_{\text{Prescribed Forcing}}$$

Discretization

$$\frac{d\phi}{dt} = \underset{\text{Model}}{f(\phi;\alpha)} + \underset{\text{Prescribed Forcing}}{g(t)}$$

$$\phi^* = \phi^n + \frac{\Delta t}{2}\left(g(t^n) + g(t^{n+1})\right)$$

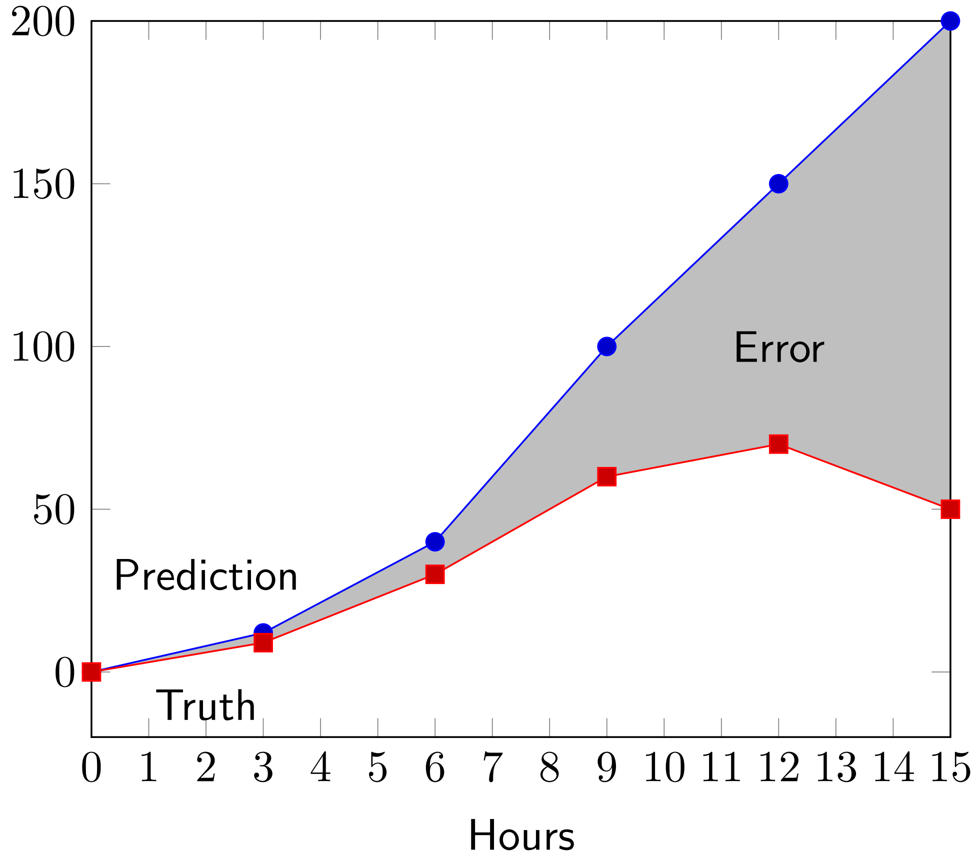
$$\tilde{\phi}^{n+1} = \phi^* + \Delta t f(\phi^*;\alpha) \qquad (\Delta t = 3\ hrs)$$

Use single-layer neural net with 128 hidden nodes.
Train on left half of domain, test on right half.

Test using single-column mode with prescribed advective forcing time series from a random NGAqua test column .

Loss function 1: Minimize 1-timestep MAE of ϕ↑1
…typically blows up within a day (8 timesteps) ☠️

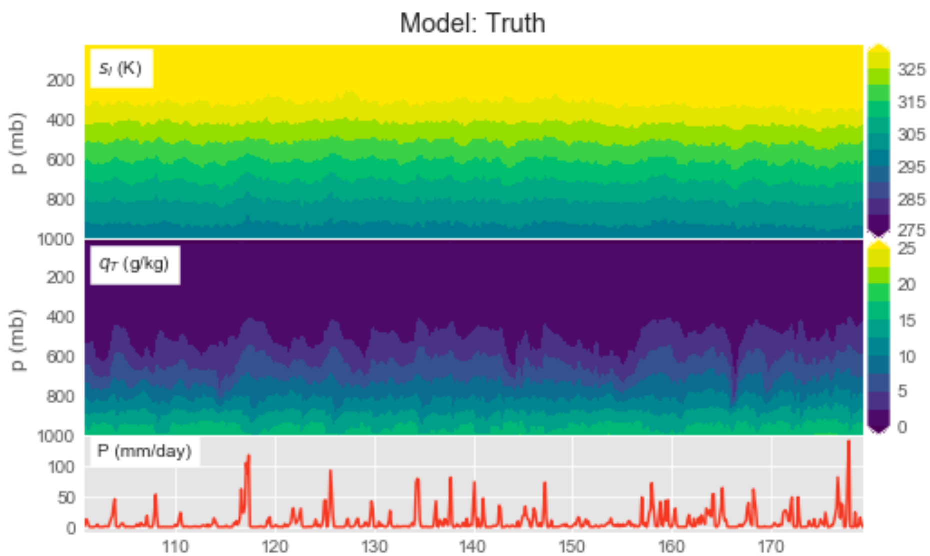
Loss function 2: Penalize error over many time steps



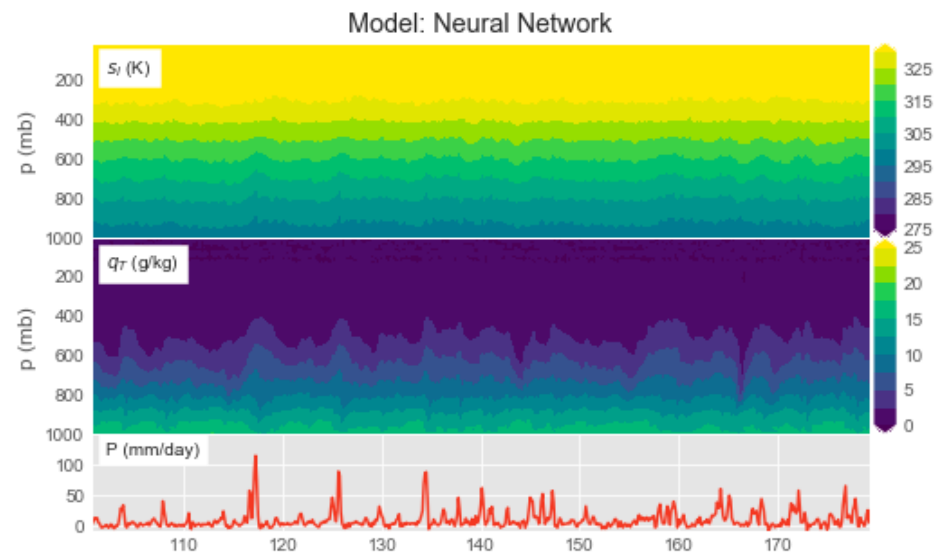
We found T = 20 timesteps = 2.5 days works well.

Scheme is now stable and accurate for 80 days

Observed time series at x=78km, y=5198 km

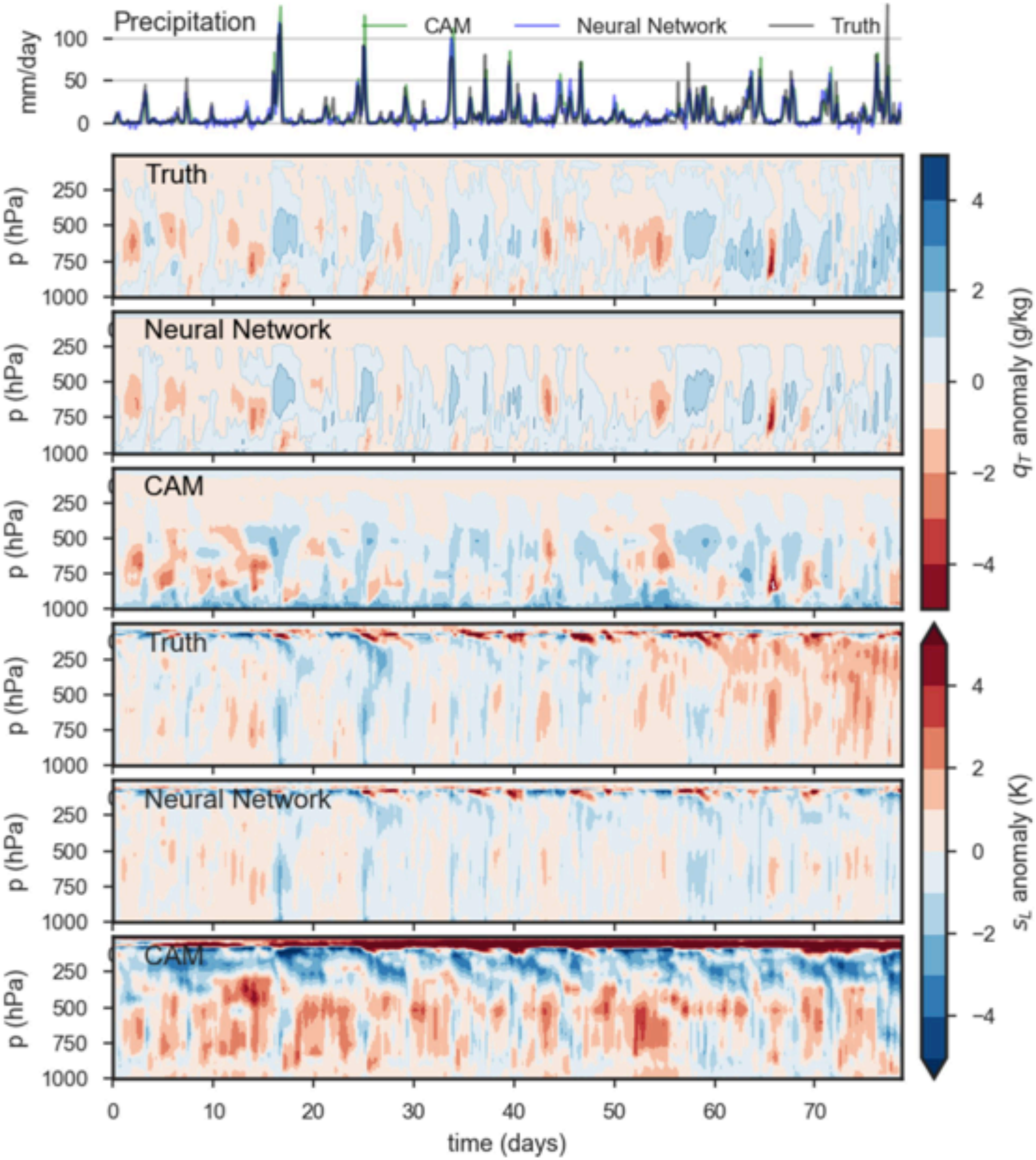


Forced initial value experiment with Neural network

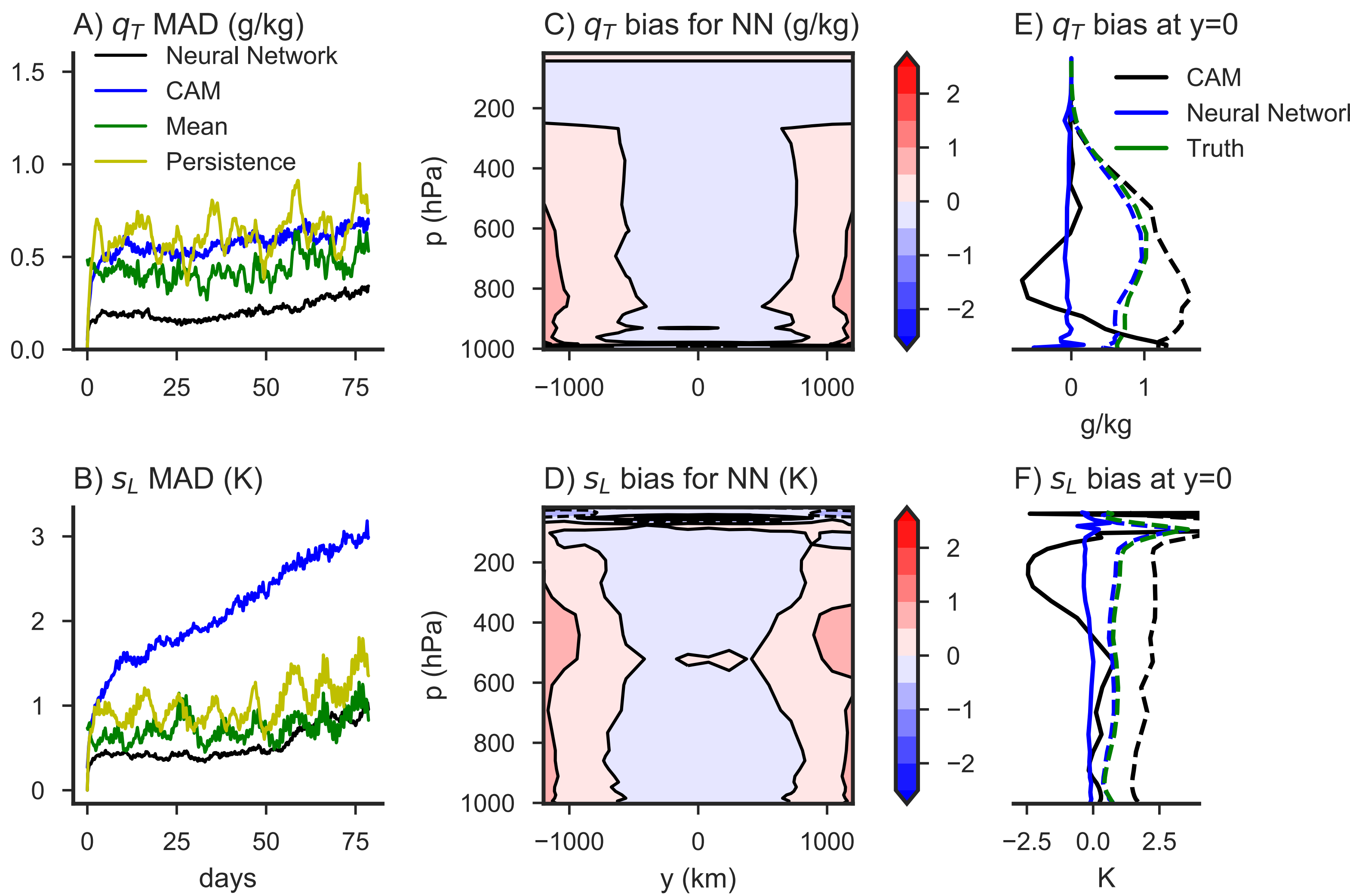


- 5 hidden neurons minimize 1-step error and explain 60-70% of $Q_{1,2}$ variance
- But 128 are required to also minimize 64-step (long-term) error

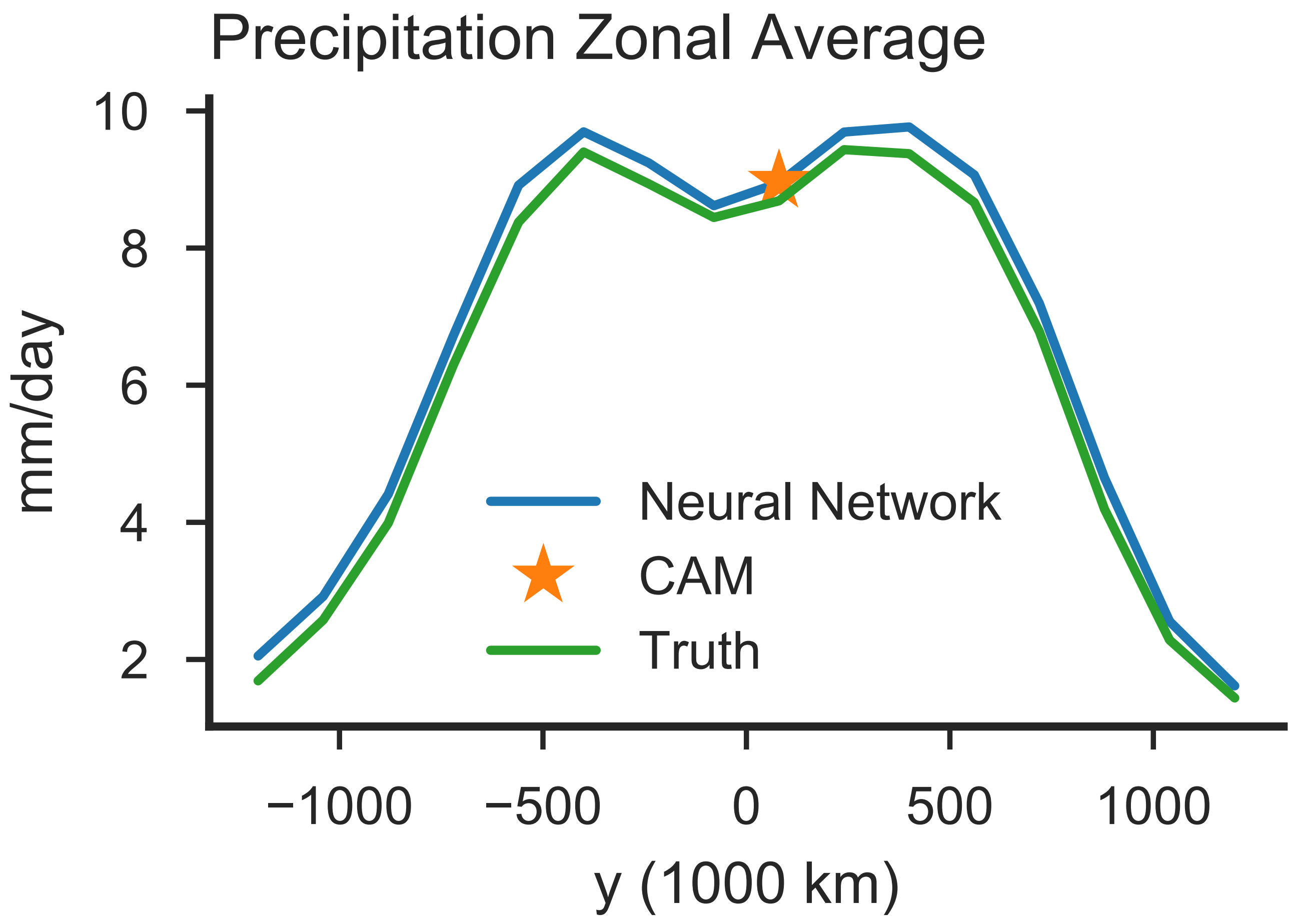
Comparison with single-column CAM5



Outperforms SCAM5 statistically as well



Zonal-mean precipitation also good



Ongoing and future work

- Enable other time steps
- Extend NN into subtropics and midlatitudes
- Test in 3D prognostic mode
- Stochastic parameterization
- Land

More thoughts about ML

- ML holds great promise for improving weather and climate models, esp. for parameterizing complex subgrid variability.
- ML still requires human expertise to succeed.
- ML parameterizations can and will be integrated into existing weather and climate models in the next five years.
- Marrying ML and data assimilation will be a growth industry
- Brightest near-term application may be in sub-seasonal to seasonal-range prediction where global cloud and ocean eddy-resolving models are still too expensive but hi-res models and observations provide a good training data set.