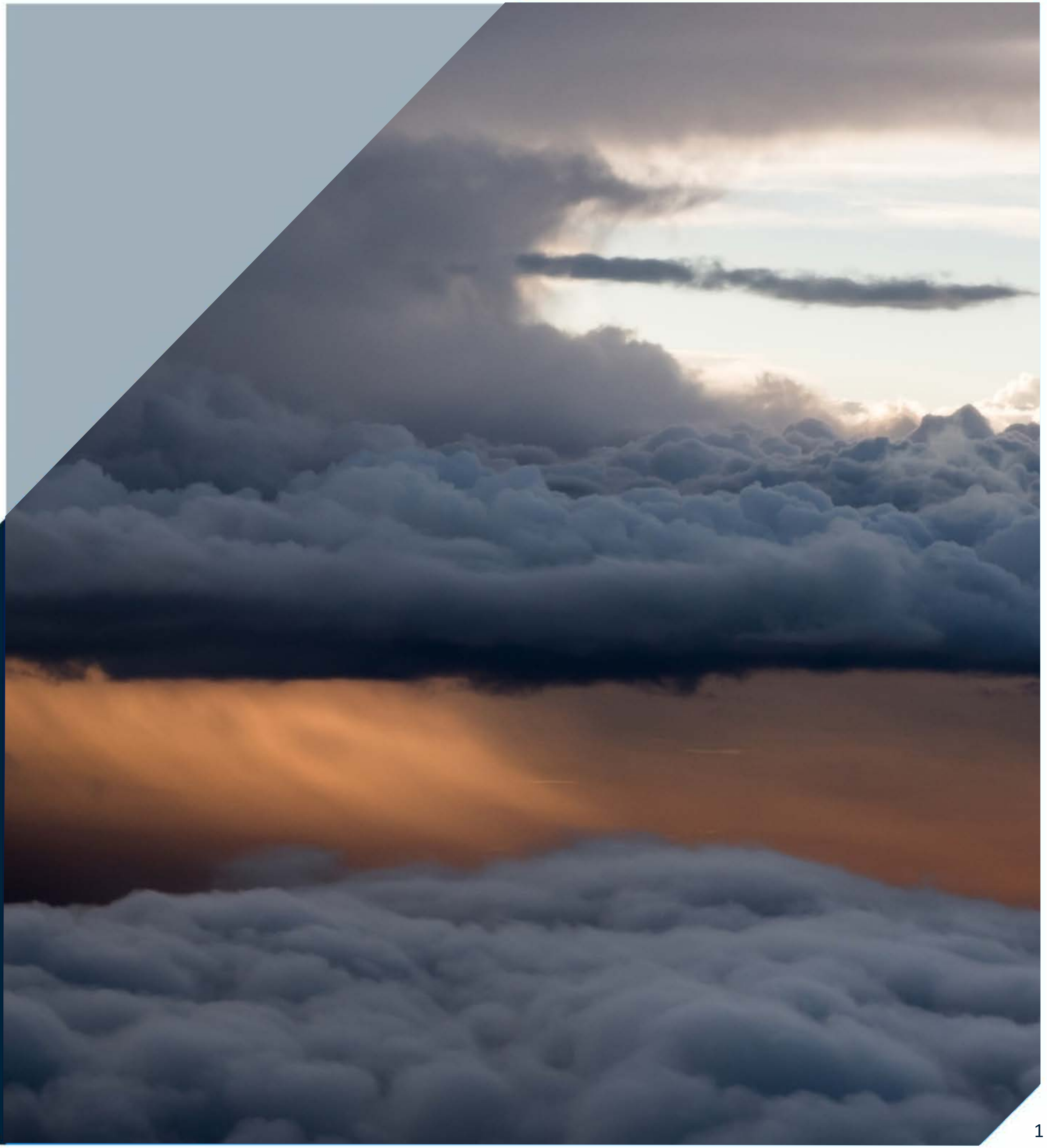
Machine learning the moist physics of a GSRM using coarse-graining

Vulcan Climate Modeling ML group
Seattle WA USA
ChrisBr@vulcan.com

May 28, 2020



VULCAN

GFDL

Goal: Improving a climate model to improve rainfall predictions using machine learning (ML)

- A global storm-resolving model (GSRM) with a finer grid of 1-3 km may (with work) better simulate individual storm clouds and mountains than a conventional 25-200 km grid GCM

  ....but is too computationally intense for ensembles of multidecadal integrations.

Goal:

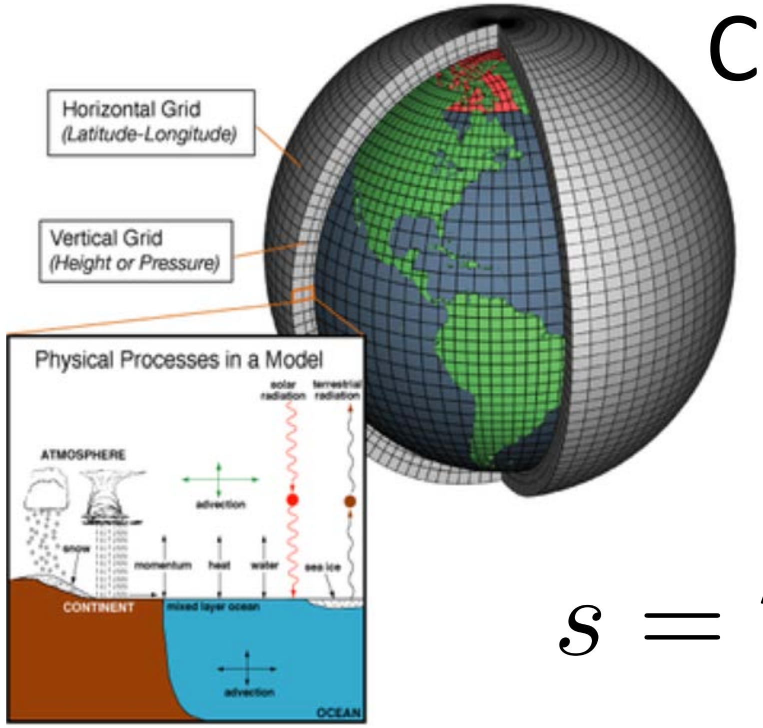
Use a realistic GSRM for training a skillful machine-learning based parameterization of subgrid clouds and precipitation for a coarser-grid global climate model.








Coarse-resolution dynamics and parameterized physics



$$s = T + \frac{g}{c_p} z$$

$$q = \frac{\text{Mass water vapor}}{\text{Mass dry air}}$$

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

Apparent heating (K/day)

SW+ LW radiation, latent heating, etc

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

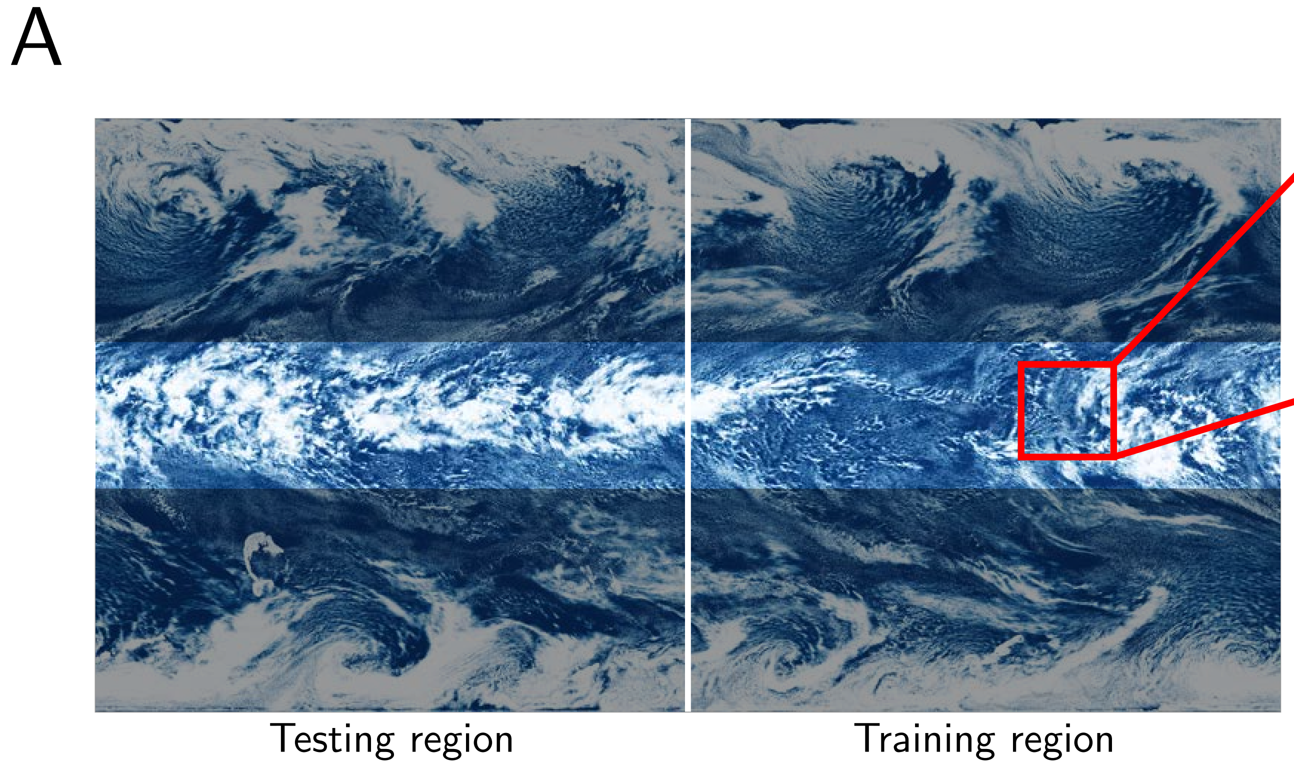
Apparent moistening (g/kg/day)

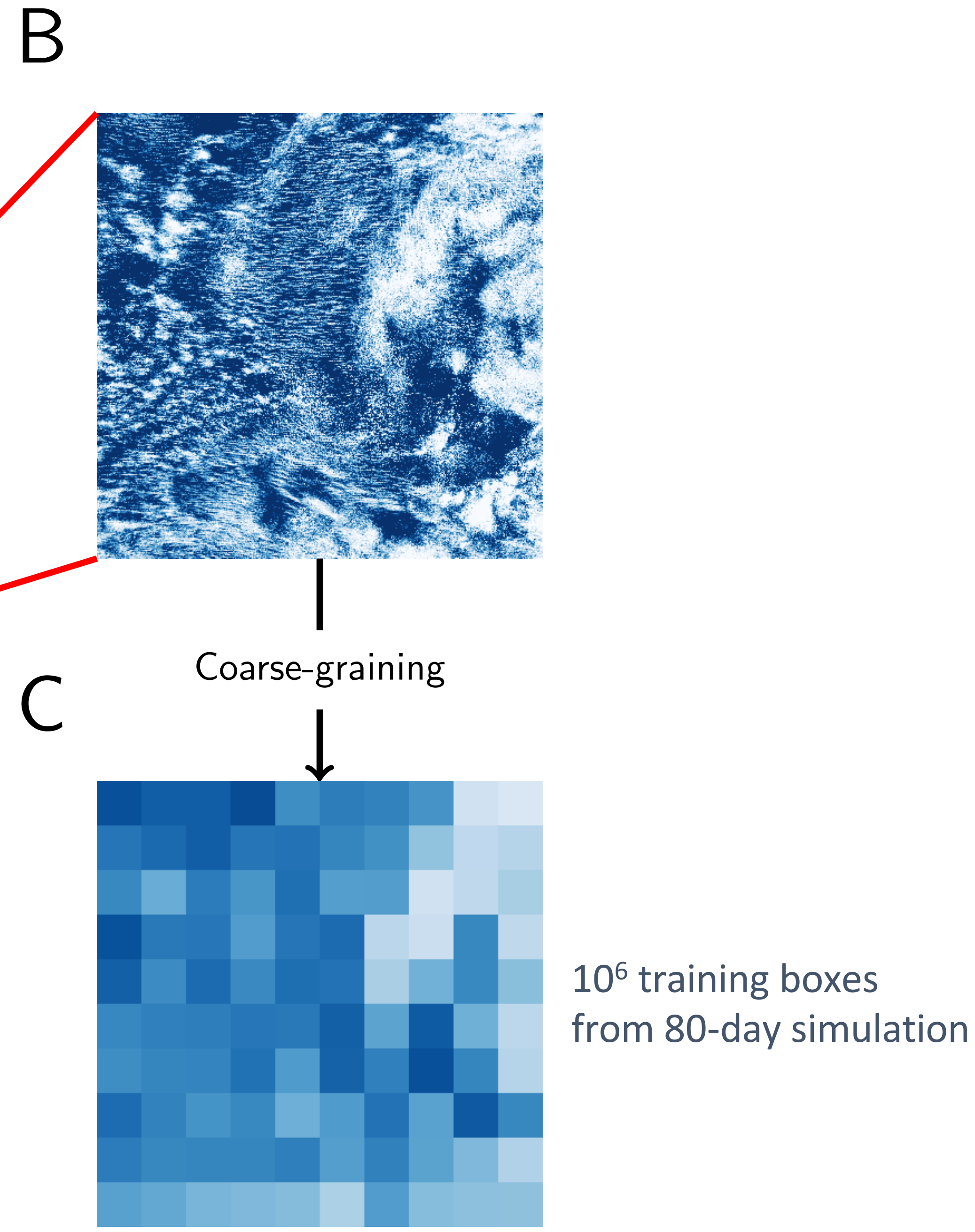
$$\frac{\partial \overline{\mathbf{u}}}{\partial t} + \overline{\mathbf{v}} \cdot \overline{\nabla} \overline{\mathbf{u}} + \mathbf{f} \times \overline{\mathbf{u}} - \frac{1}{\rho} \nabla \overline{p} = Q_{u,v}$$

Apparent momentum source
(for now rely on coarse model
parameterizations of PBL, GWD, etc.)

Past work: Training ML with a coarse-grained tropical channel simulation

- Use 80-day 4 km aquaplanet run as ‘truth’ to machine-learn moist physics parameterization for the low-res model.
- Goal: forecast with low-res dycore + ML param should match hi-res run.



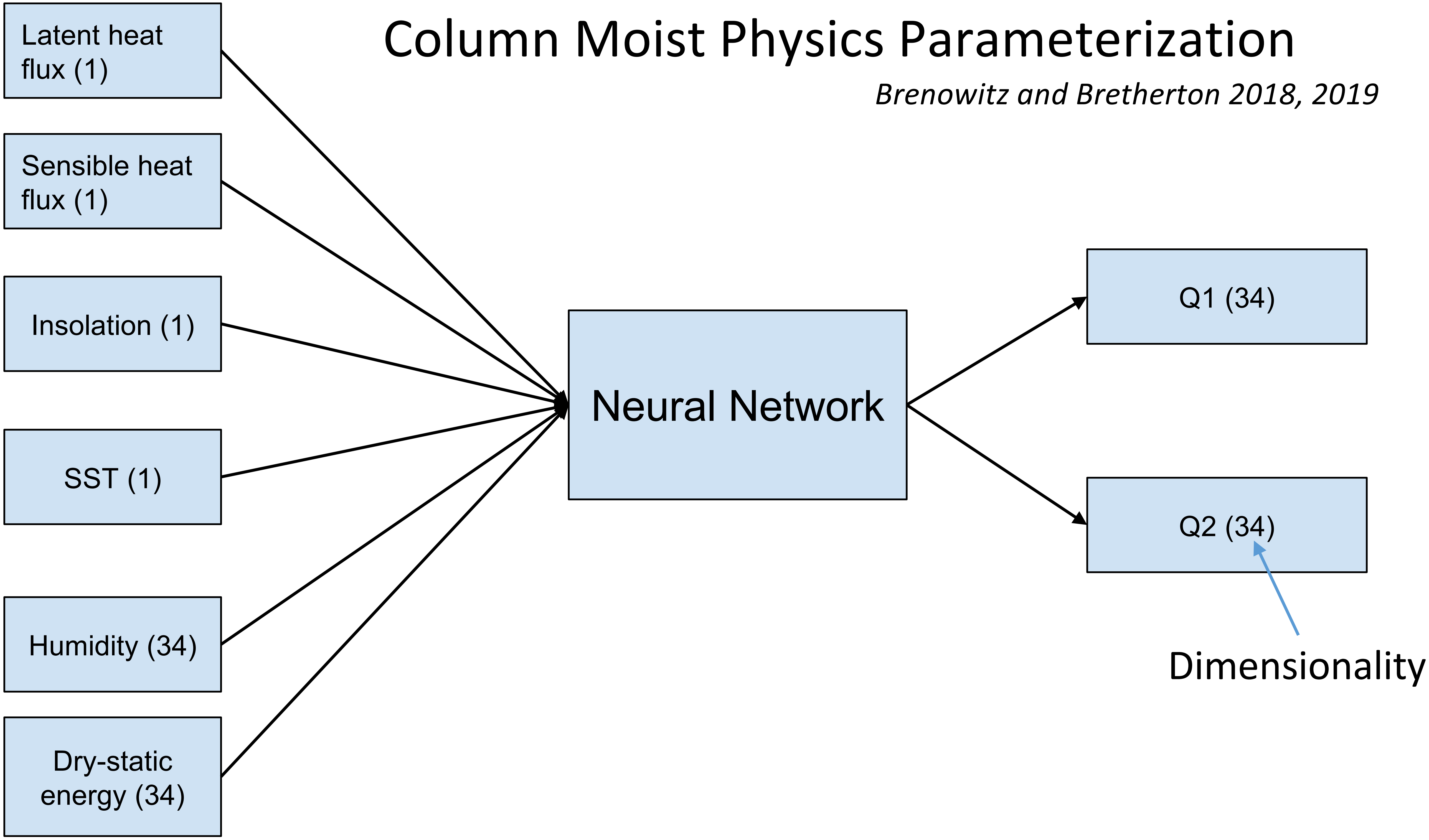


- 160 km coarse (low-res) grid
- Calculate $Q_{1,2}(\mathbf{r}, t)$ (coarse-grid ‘moist physics’ tendencies including radiation) as residuals of dynamical equations.
- Unified moist physics, turbulence and radiation parameterization: Learn $Q_{1,2}$ as functions of local column conditions using a neural net.

Brenowitz and Bretherton 2018, 2019; Rasp et al. 2018; O’Gorman and Yuval 2020

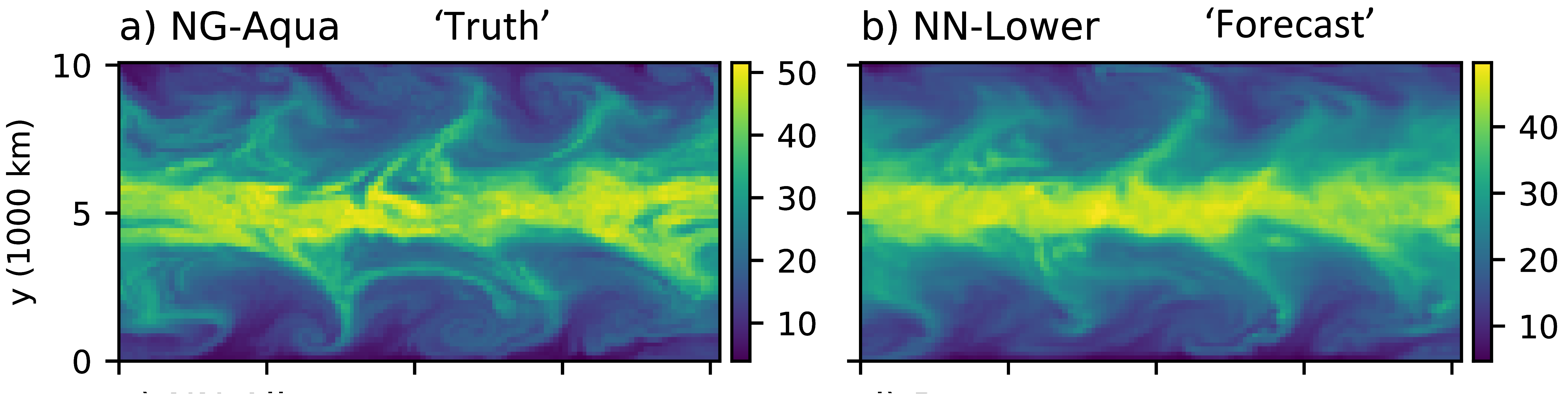
Column Moist Physics Parameterization

Brenowitz and Bretherton 2018, 2019



Couple the ANN to the flow solver on 160 km grid

If inputs and error metric are carefully designed to prevent rapid model blow-up, hi-res model is skillfully forecast by low-res model with NN parameterization



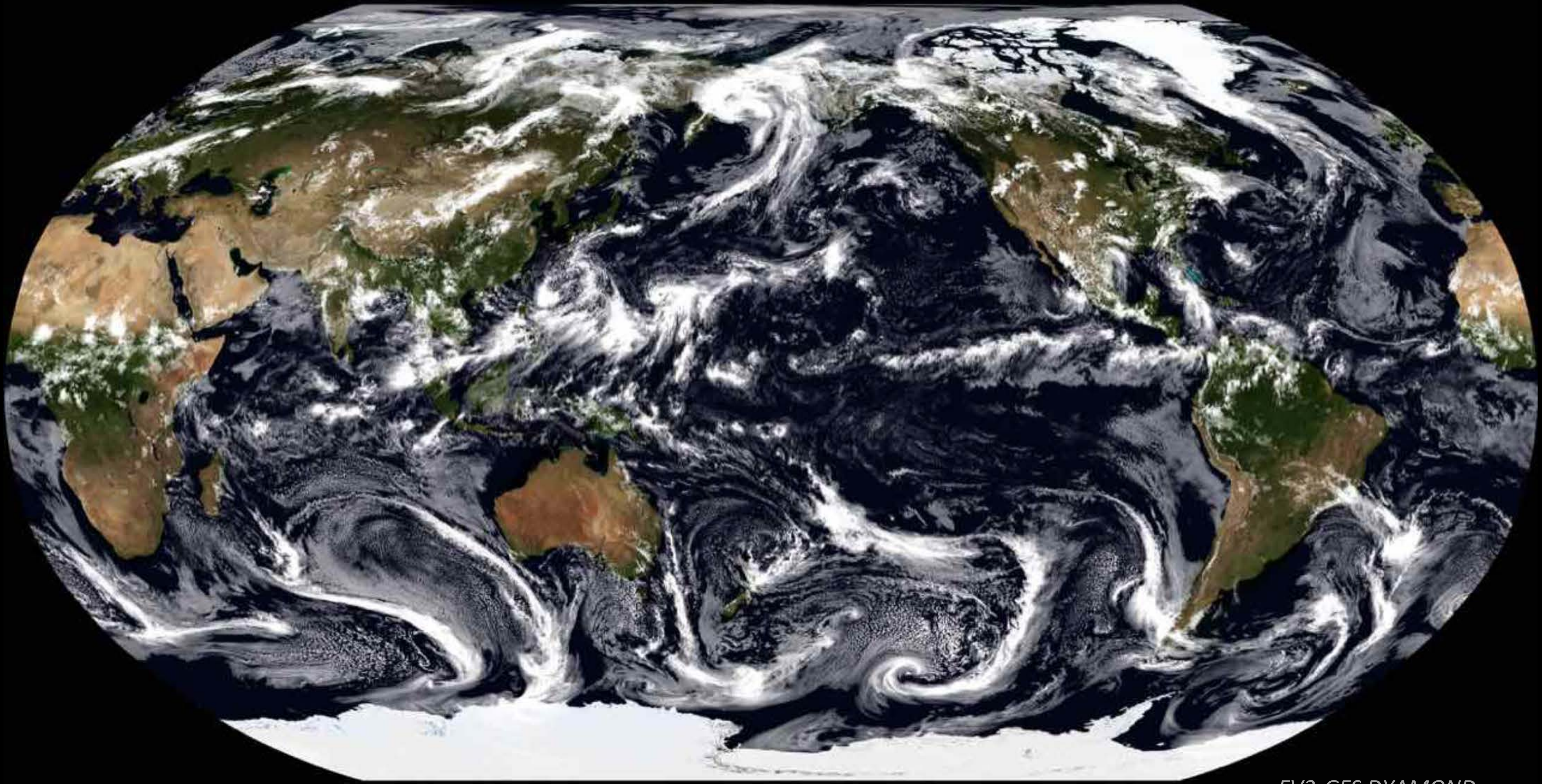
Precipitable water after 5 days

Brenowitz and Bretherton (2019)

…but the 'climate' slowly drifts after 10 days toward a weaker ITCZ

See Rasp et al. (2018, GRL) and O'Gorman and Yuval (2020, arXiv) for other aquaplanet successes with similar methods applied to related models.

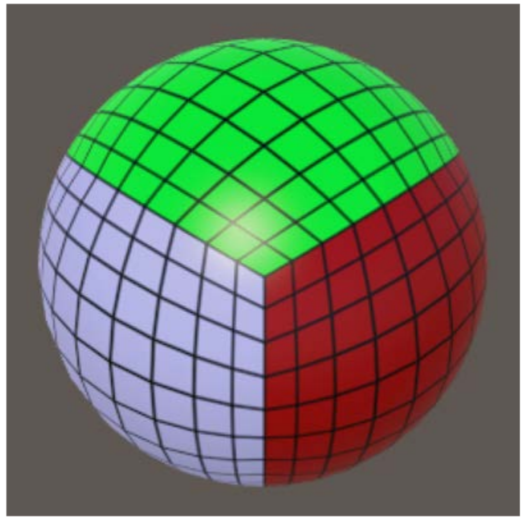
Can we apply same ML approach to GFDL's 3 km FV3-GFS global atmospheric model?



FV3-GFS DYAMOND run
S.-J. Lin and Xi Chen, GFDL

FV3GFS and SHiELD[1] global weather/climate models

- FV3GFS: Open-source global atmosphere model used by NOAA for operational weather forecasts
- FV3 dycore – Customized D-grid finite volume method on cubed sphere.
- Nonhydrostatic by default, 80 vertical levels used here.
- Specified time-varying sea-surface temperature used here
- **Horizontal grid resolutions:**
  - 3 km (C3072) No deep cumulus parameterization or gravity-wave drag
  - 13 km Used for NCEP's current operational global weather forecasts
  - 25 km Finest grid currently practical for climate simulations of many decades
  - 200 km (C48) Typical coarse climate model grid – good for prototyping or millennial runs.
- **Physical parameterizations:**
  - Land surface and surface fluxes (NOAH)
  - Radiation (RRTMG)
  - Gravity-wave drag
  - Boundary-layer (including shallow clouds) and shallow Cu (Han-Bretherton, Han-Pan)
  - Cloud microphysics and subgrid variability (GFDL one-moment)
  - Deep cumulus convection (SAS)



[1] GFDL's SHiELD is FV3GFS with modest changes to cloud physics and advection and is not open-source.

Tendency-difference method for coarse-graining

- $a_f(t, x, y, \sigma)$: space-time field (e.g. humidity) at fine resolution. $a_c(t, x, y, \sigma)$ is coarse-res field.
- Coarse-graining operator: $\bar{\cdot}$ (some form of horizontal averaging from fine to coarse grid)
- Coarse model (200 km FV3GFS) should match fine model (3 km SHiELD) starting at $a_c = \overline{a_f}$:

$$\frac{\partial a_c}{\partial t} \approx \frac{\partial \overline{a_f}}{\partial t}$$

- Uncorrected coarse model:

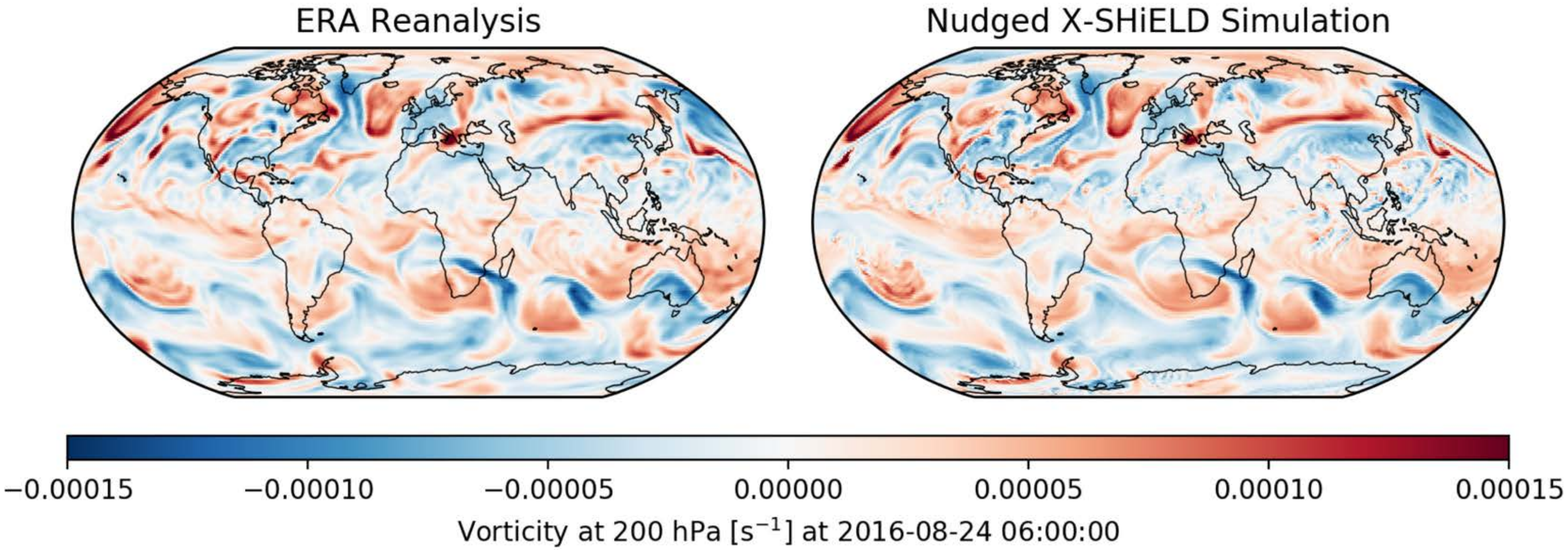
$$\left(\frac{\partial a_c}{\partial t}\right)_0 = A_c + Q_a^p, \qquad A_c = -\mathbf{u}_c \cdot \nabla a_c$$

- Coarse model can include no physics ($Q_a^p = 0$) or a subset of parameterized physical processes that help track the fine-grid model (e. g. turbulence, radiation, clouds, Cu parameterization).
- Machine-learn a state-dependent corrective source ΔQ_a for the coarse model:

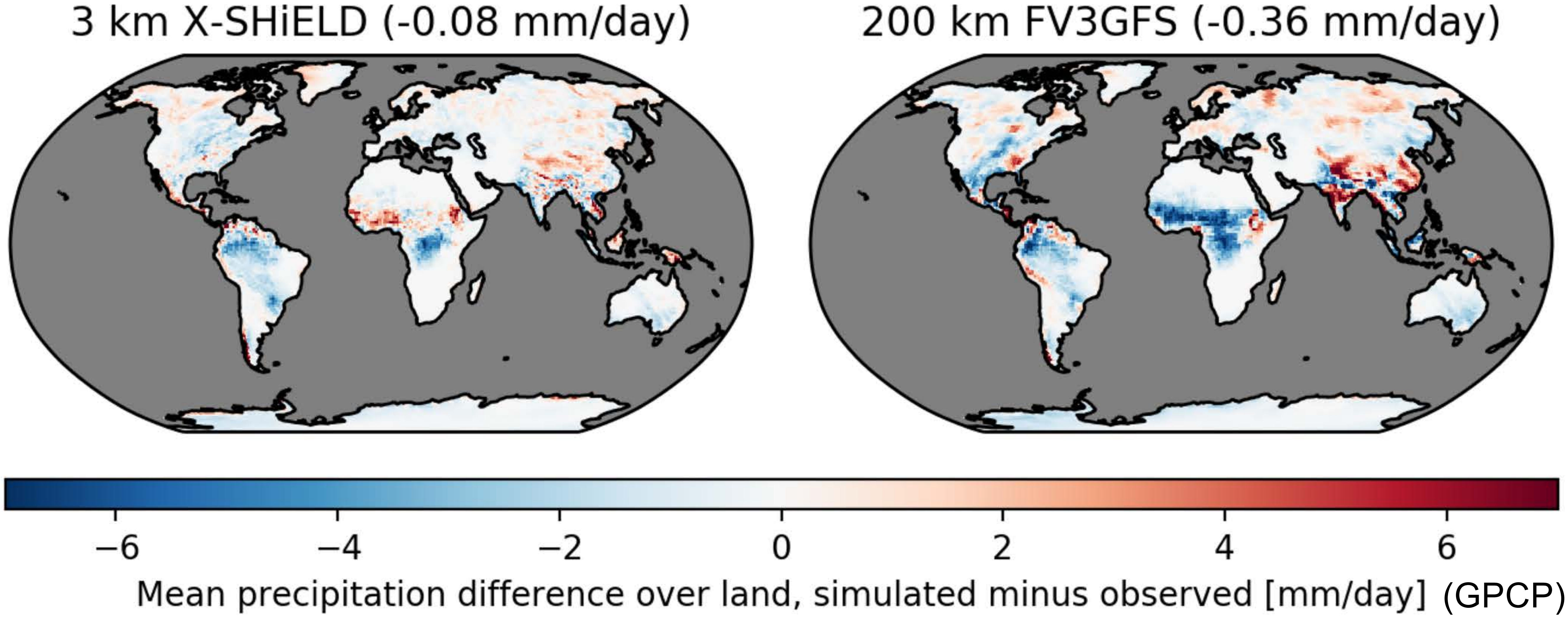
$$\Delta Q_a = \frac{\partial \overline{a_f}}{\partial t} - \left(\frac{\partial a_c}{\partial t}\right)_0 \qquad \rightarrow \qquad \left(\frac{da}{dt}\right)_c = \left(\frac{da_c}{dt}\right)_0 + \Delta Q_a^{ML}$$

Training dataset: nudged 3 km SHiELD (modified FV3-GFS)

- Training dataset: 40 d 'nudged DYAMOND' simulation on GAEA (1 Aug to 9 Sep 2016):
  - Observed SSTs
  - Light nudging ($\boldsymbol{\tau}$ = 1 day) of 3 km T/u/v/p_s to ERA5 reanalysis keeps meteorology 'data-aware'. Nudging tendencies are considered to be part of the learned physics
  - Store atmospheric and land-surface restart fields coarse-grained to 25 km every 15 min



40 d mean precipitation bias over land: 3 km SHiELD vs. 200 km FV3GFS

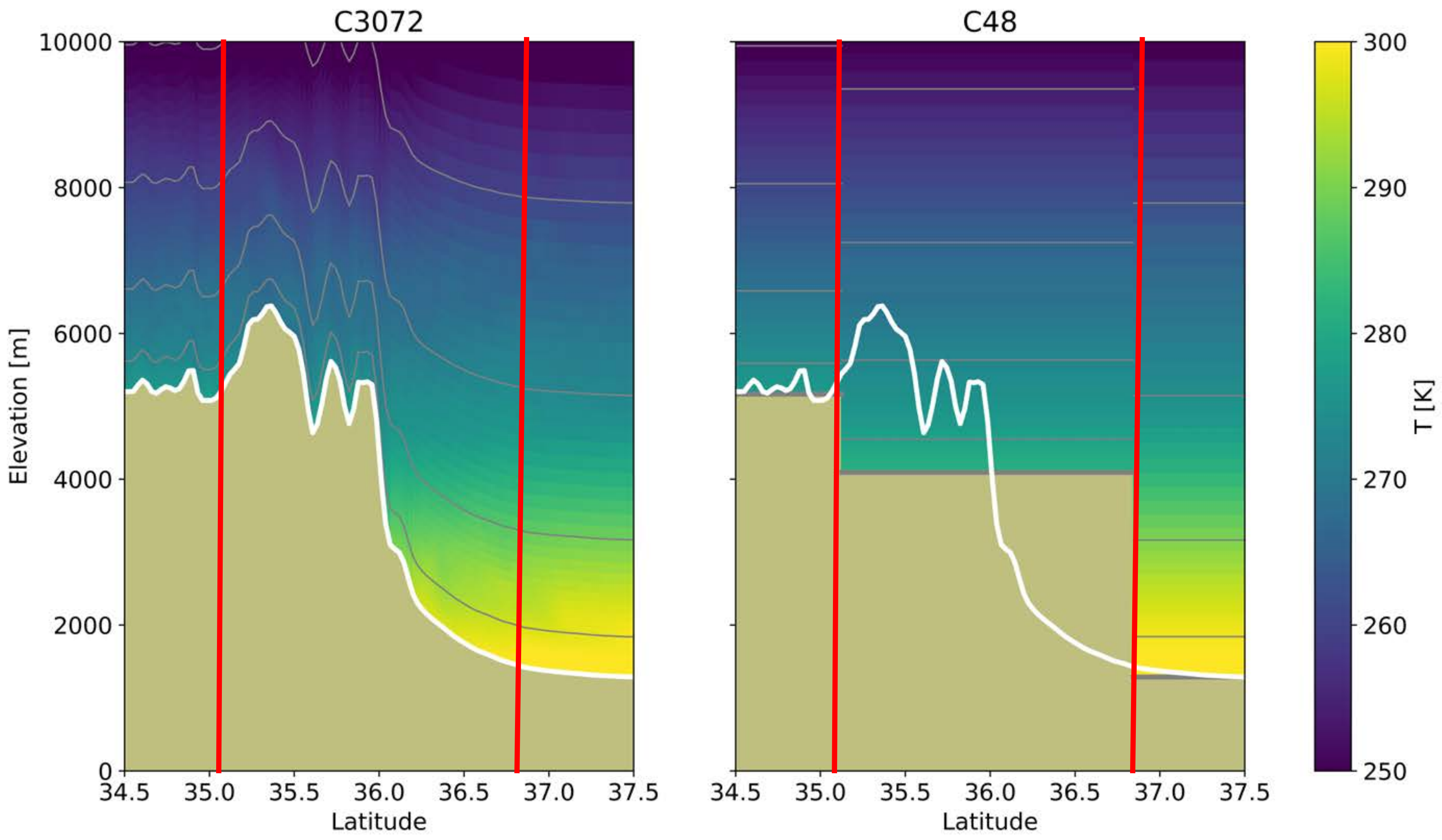


3 km rainfall bias much smaller over sub-Saharan Africa and Himalayas

Diurnal cycle of precipitation over land is also greatly improved in SHiELD

Conceptual issues over topography

- Consider 3 km -> 200 km coarse-graining over the Himalayas

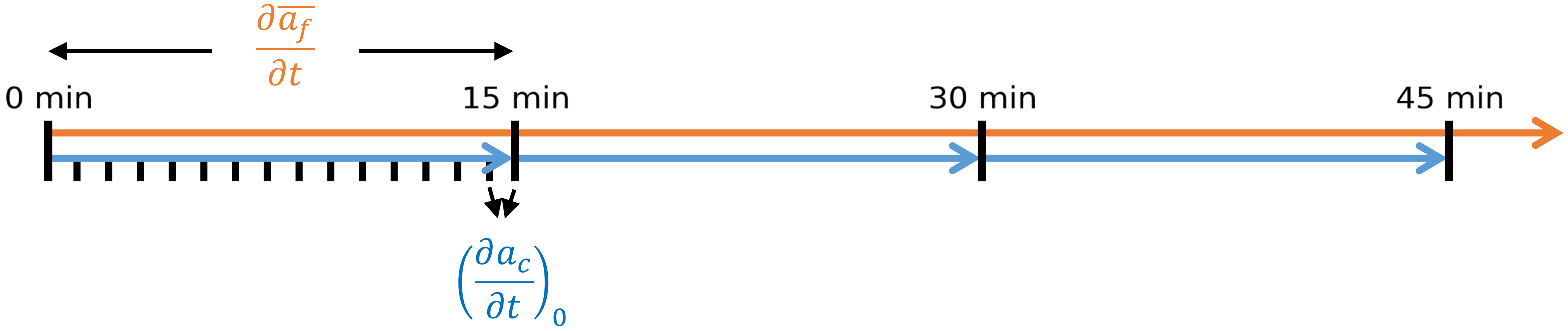


- We coarse-grain to obtain vertical **profiles** and **apparent sources** of T, q, etc.
- 5 km relief within a coarse cell
- Most fields are much more constant along a pressure surface than along a terrain-following model surface

→ Coarse-grain on pressure levels, not model levels

Our implementation of tendency difference method

Coarsened state of fine-resolution model saved every 15 minutes.
Fine-res tendencies computed from these snapshots.



Coarse-resolution model initialized from each coarsened high-resolution snapshot and run forward for 15 minutes, with a 1-minute timestep.

Low-res tendencies computed from final minute.

Apparent source:

$$\Delta Q_a = \frac{\partial \overline{a_f}}{\partial t} - \left(\frac{\partial a_c}{\partial t}\right)_0$$

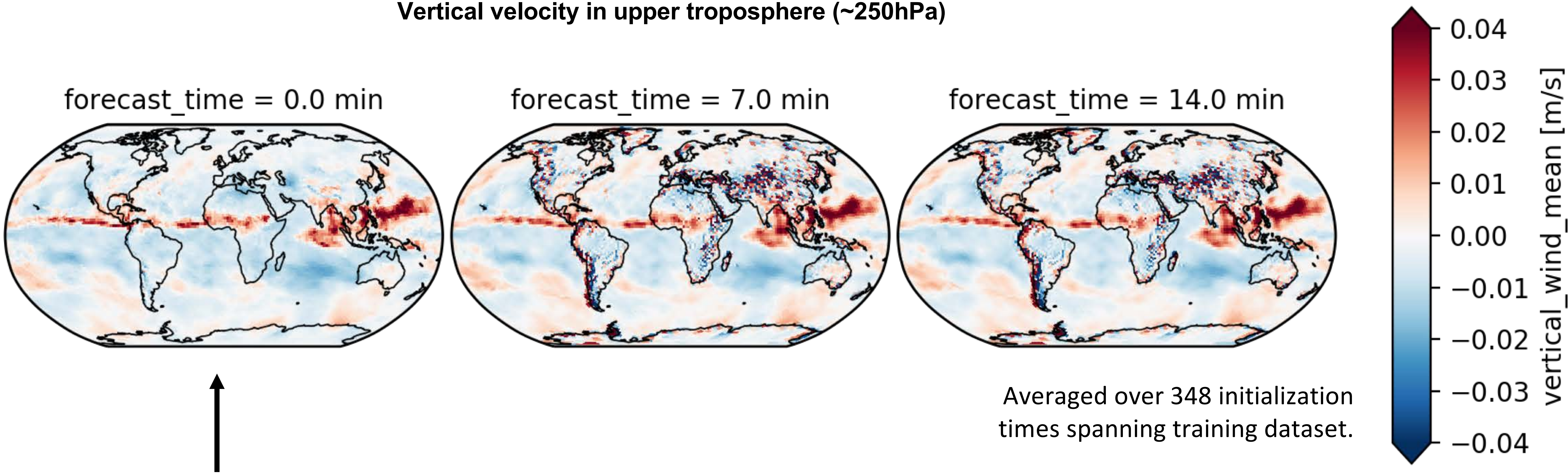
Coarse model physics

We run ML on top of four configurations of the coarse-resolution model:

1. **physics-on**
   - All physical parameterizations on
     (land surface, boundary layer, convection, radiation, microphysics, gravity wave drag)
2. **deep-off**
   - Turn off deep convection scheme
3. **clouds-off**
   - Deep and shallow convection schemes off
   - No microphysics
   - Use clear-sky radiation only
4. **physics-off**
   - Run only dynamical core

Despite careful efforts of pressure-level coarse-graining, vertical velocity noise remains over topography

Vertical velocity in upper troposphere (~250hPa)



Averaged over 348 initialization times spanning training dataset.

Fine resolution model coarsened to 200km resolution

These results are from clouds-off, but all physics configurations give comparable results

Machine learning: model training

Training set = 1.7M samples (130 initializations x 13824 grid points)
Test set = 660K samples (48 initializations x 13824 grid points)
Train/test data separated by split date to minimize correlated data across sets

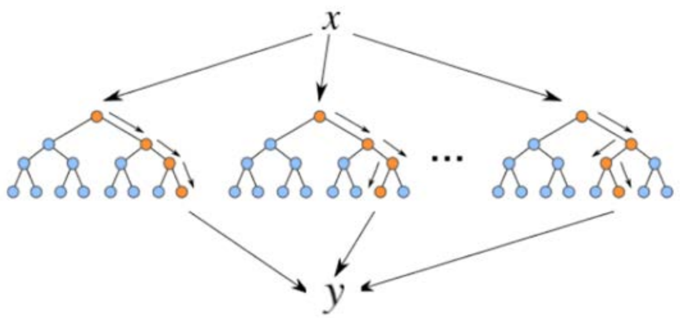
Features

air temp vertical profile

specific humidity vertical profile

additional potential features, ex. insolation, surface temp, surface geopotential, ...

Random forest model



- Ensemble of 13 decision tree estimators, each with max depth of 13
- Easier to run with stability in prognostic simulations (relative to neural nets)

Predictions

dQ1 vertical profile

dQ2 vertical profile

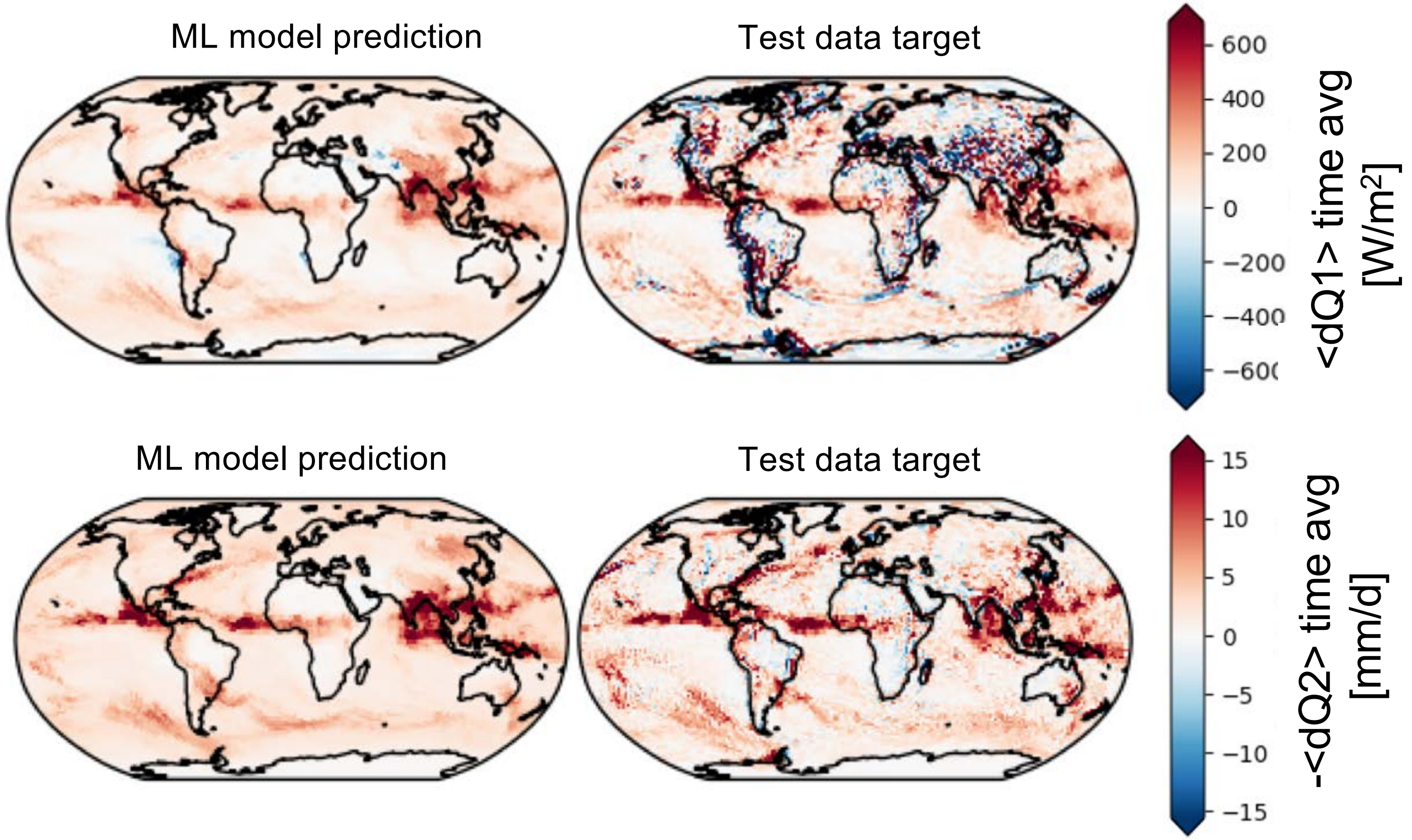
$$\Delta Q_a = \left(\frac{d\bar{a}}{dt}\right)_f - \left(\frac{da}{dt}\right)_{c0}$$

Machine learning: diagnostic skill

Column integrals of the ML-predicted vertical profiles reproduce spatial features of net heating and precipitation, while also smoothing out noise from coarse-graining and initialization.

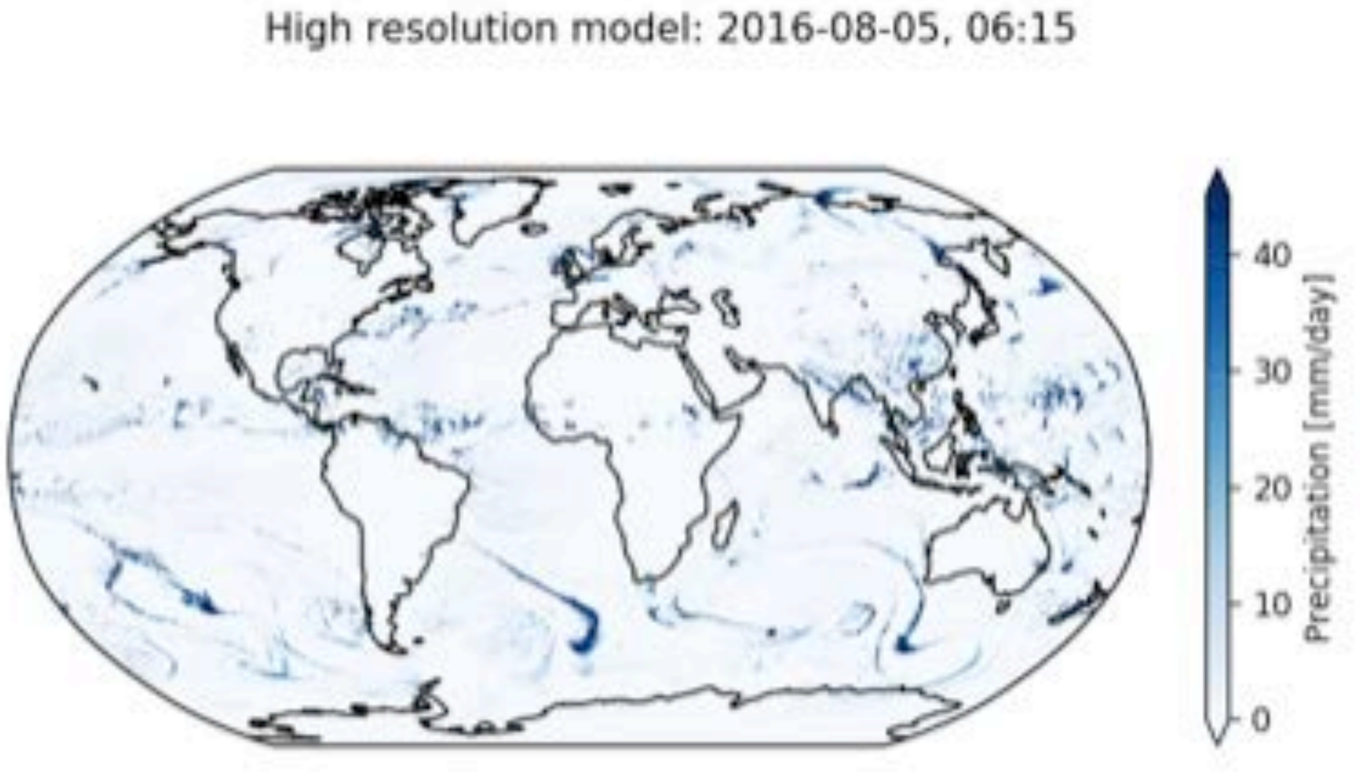
Net ML-implied heating

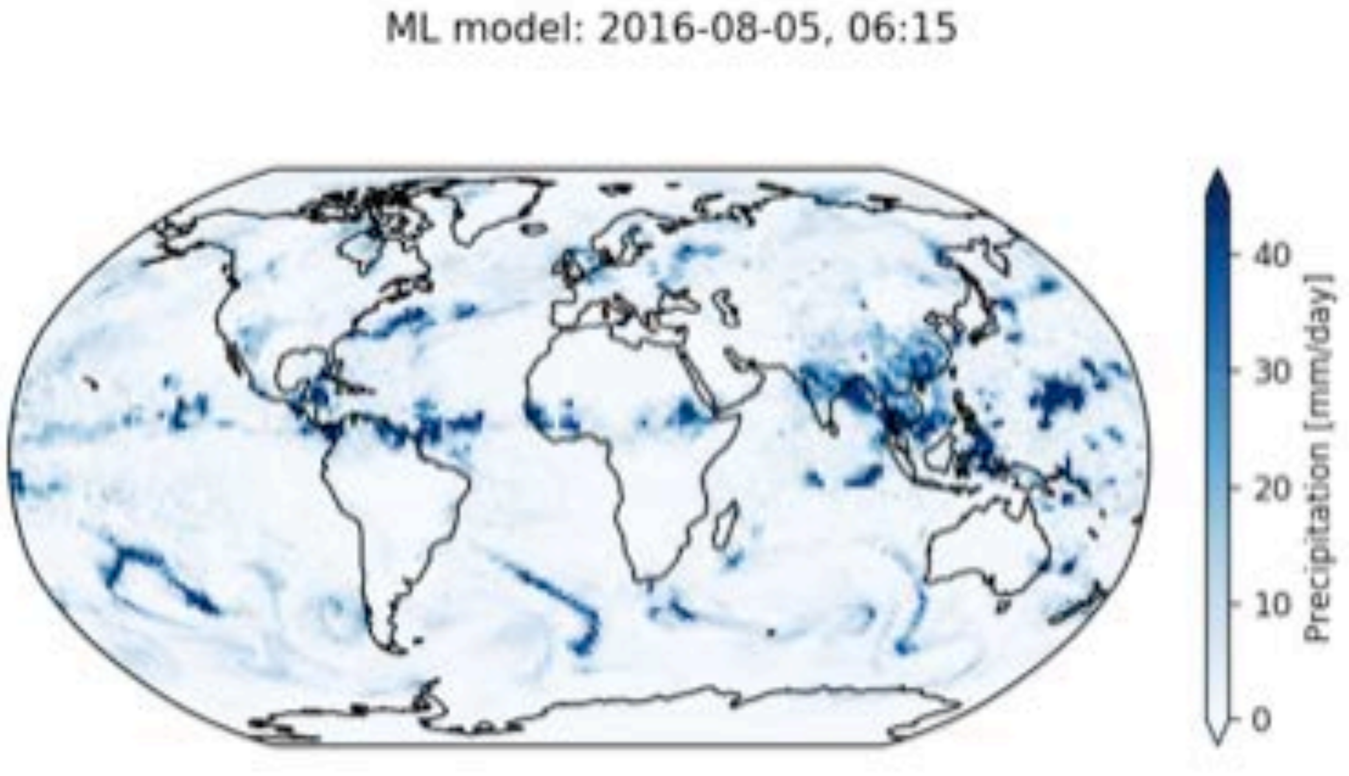
Net ML-implied precipitation



ML model trained with clouds-off configuration

Skillful 2-day prognostic forecast of precipitation





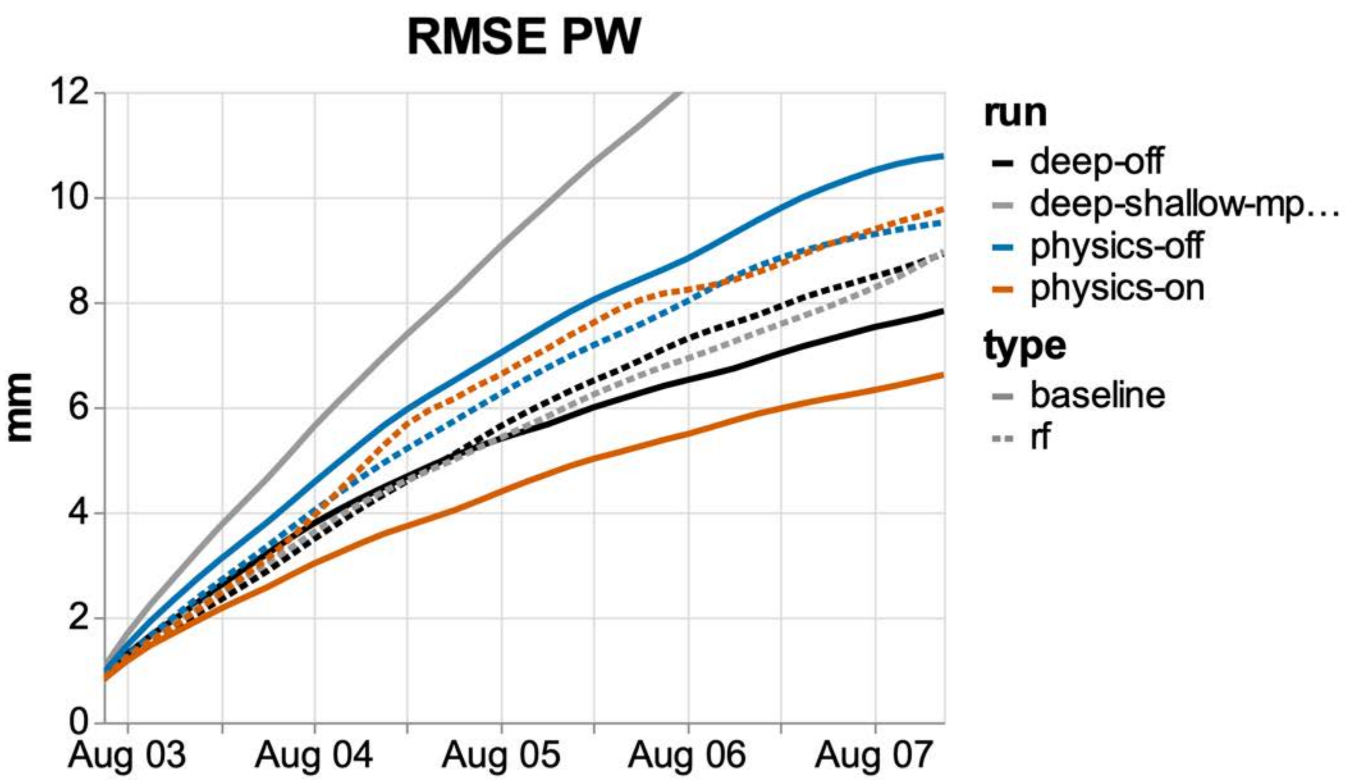
3 km simulation averaged to hourly 25 km

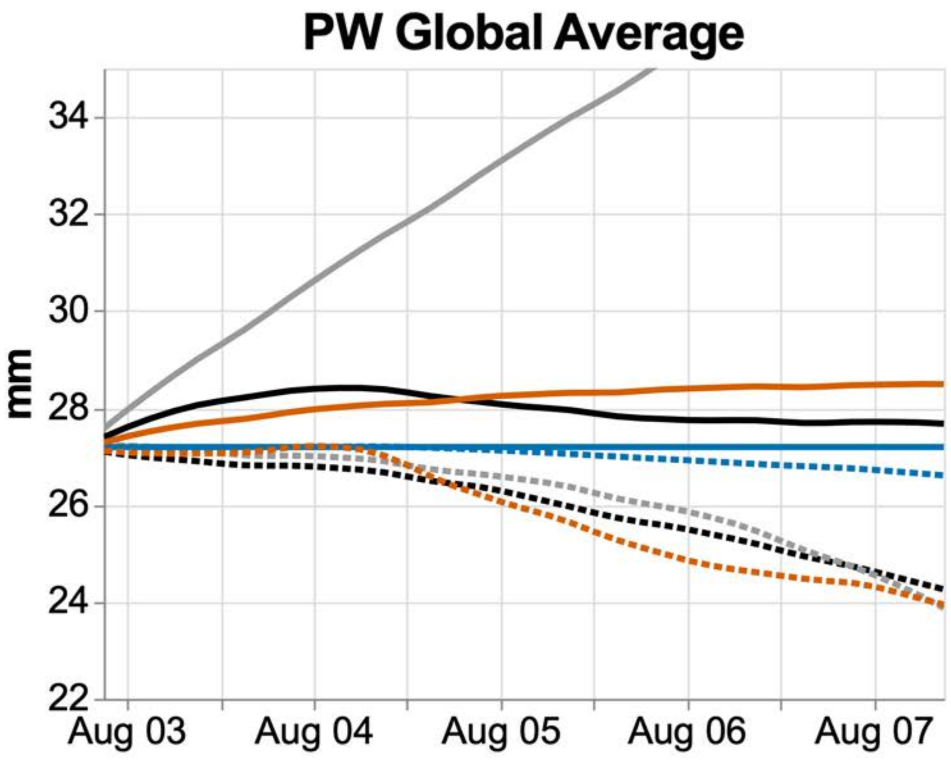
200 km FV3GFS with deep convection param replaced by ML

Weather forecast test

- 5-10 day 'weather forecasts' are an acknowledged test of global atmospheric model skill
- Goal is to match the evolution of the 3 km training model.
- Skill metric: root-mean-square error (RMSE) of map of column water vapor in 200 km model vs. coarsened 3 km model. Smaller is better.
- RMSE grows as coarse model diverges from training model.
- `Climate' skill metric: minimal global-mean drift over 5 days
- Currently, the best model configuration includes all conventional physics parameterizations and no ML.
- Most (not all) ML runs to date crash between 5 and 10 days

…but it's early days, and we are working to improve ML skill.





Conclusions and Outlook

- VCM has developed a unique cloud-based workflow for training a ML correction to a coarse-resolution climate model based on fine-resolution GSRM simulations.
- We have trained stable ML schemes that can make skillful global rainfall forecasts over land and ocean for 5 days or longer given specified SST.
- Tendency-difference method is flexible but is degraded by vertical velocity transients; we are exploring improved approaches to improve accuracy and reduce climate drift
- A dynamics-coupled machine learning scheme will ultimately be required.