Correcting Coarse-Resolution Weather and Climate Models by Machine Learning from Global Storm-Resolving Simulations

Christopher Bretherton[1], Brian Henn[1], Anna Kwa[1], Noah Brenowitz[1], Jeremy McGibbon[1], Spencer Clark[1], Andre Perkins[1], and Lucas Harris[2]

[1]AI2
[2]Geophysical Fluid Dynamics Laboratory, NOAA

November 22, 2022

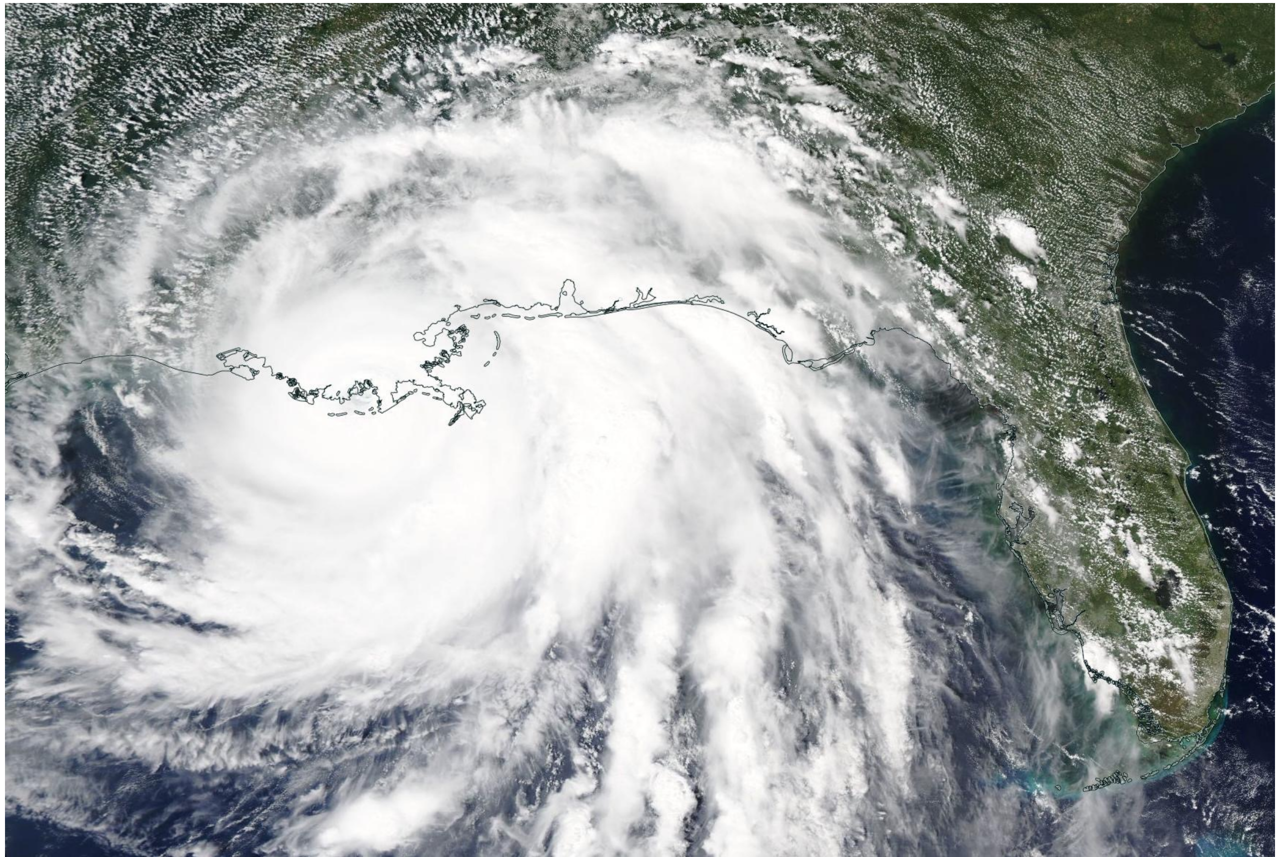
Abstract

Global atmospheric 'storm-resolving' models with horizontal grid spacing of less than 5˜km resolve deep cumulus convection and flow in complex terrain. They promise to be reference models that could be used to improve computationally affordable coarse-grid global climate models across a range of climates, reducing uncertainties in regional precipitation and temperature trends. Here, machine learning of nudging tendencies as functions of column state is used to correct the physical parameterization tendencies of temperature, humidity, and optionally winds, in a real-geography coarse-grid model (FV3GFS with a 200 km grid) to be closer to those of a 40-day reference simulation using X-SHiELD, a modified version of FV3GFS with a 3 km grid. Both simulations specify the same historical sea-surface temperature fields. This methodology builds on a prior study using a global observational analysis as the reference. The coarse-grid model without machine learning corrections has too little cloud, causing too much daytime heating of land surfaces that creates excessive surface latent heat flux and rainfall. This bias is avoided by learning downwelling radiative flux from the fine-grid model. The best configuration uses learned nudging tendencies for temperature and humidity but not winds. Neural nets slightly outperform random forests. Forecasts of 850 hPa temperature gain 18 hours of skill at 3-7 day leads and time-mean precipitation patterns are improved 30% by applying the ML correction. Adding machine-learned wind tendencies improves 500 hPa height skill for the first five days of forecasts but degrades time-mean upper tropospheric temperature and zonal wind patterns thereafter. The figure shows maps of 30-day time-mean precipitation pattern difference from the fine-grid reference for prognostic simulations: (a) 200 km baseline(no machine learning correction) (b) Using random forest correction and (c) neural net correction for temperature. humidity and surface radiation corrections. RMSE is the root mean squared precipitation difference from the reference, which is 30% less for the two machine-learning corrected simulations compared to the baseline. (d) Bar charts of the land-mean, ocean-mean and global-mean precipitation biases for these three configurations, showing the machine-learning corrected simulations remove a high bias of land surface precipitation in the baseline simulation.

A14C-08: Correcting Coarse-Resolution Weather and Climate Models by Machine Learning from Global Storm-Resolving Simulations

Chris Bretherton

Senior Director of Climate Modeling,
Allen Institute for Artificial Intelligence (AI2), Seattle
Emeritus Professor of Atmos. Sci. and Applied Math, U. Washington



Hurricane Ida
29 Aug. 2021
NASA Worldview

With AI2 team members:

Noah Brenowitz
Spencer Clark
Brian Henn
Anna Kwa
Jeremy McGibbon
Andre Perkins
Oli Watt-Meyer

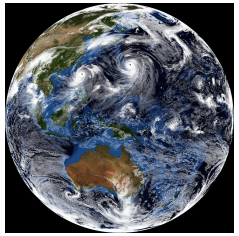
and GFDL collaborator:

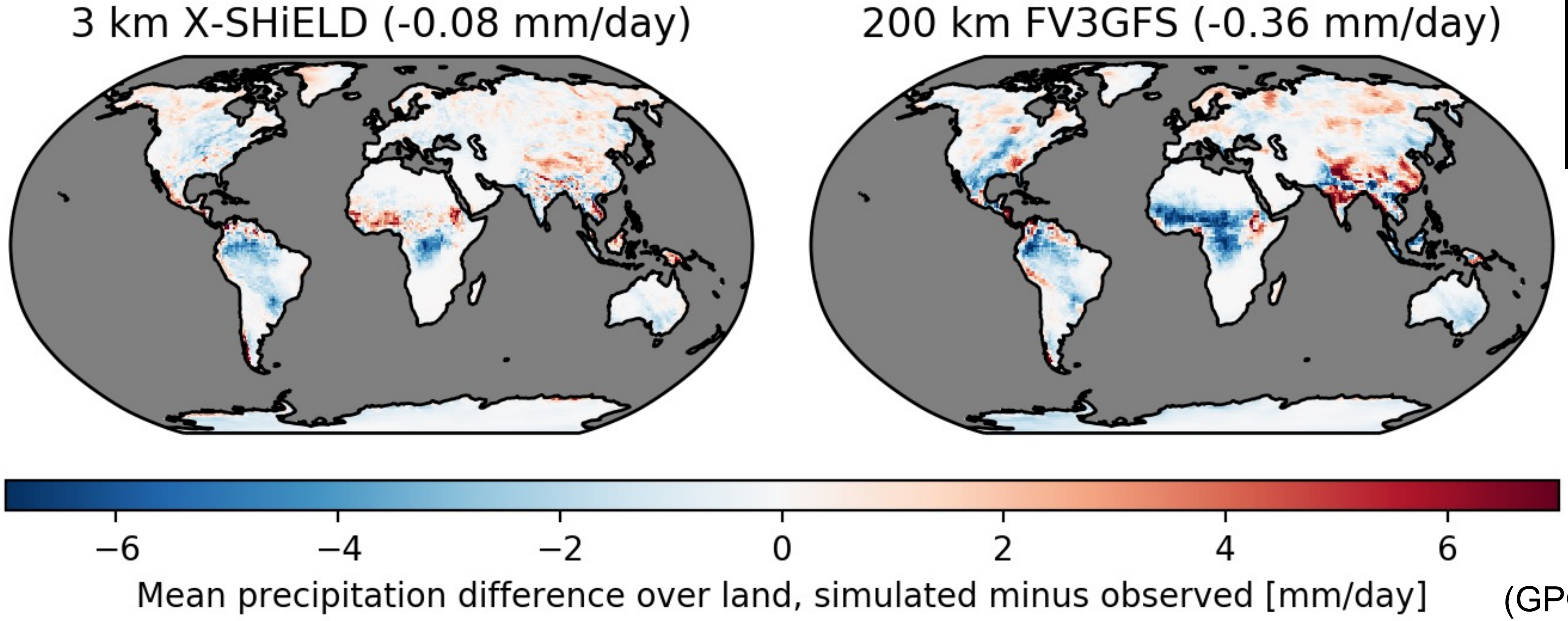
Lucas Harris

Climate models are our window into our warming world

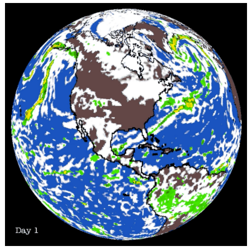
- Regional hydrological trends and extremes remain an important modeling uncertainty
- Global storm-resolving models (GSRMs, Δx = 2-5 km, 50-150 levels) give more realistic-looking simulations than CMIP6 GCMs with less subgrid modeling assumptions, but are computationally expensive.
- Could GSRMs + machine learning (ML) improve or replace the physical parameterization suites used in conventional GCMs?
- Could this enable more trustworthy projection of hydrological cycle?

3 km grid gives a better rainfall simulation than 200 km





(GPCP)



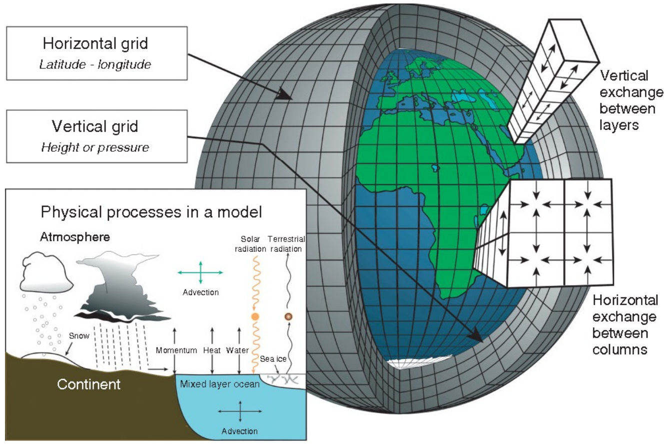
Enabled by explicit simulation of cumulonimbus cloud systems & well-resolved mountains

The 3 km model resolves variability that requires complex subgrid parameterization in GCMs

AI2-CM ML approach to climate modeling

- A global atmosphere model simulates weather and its interaction with land, ocean and ice surfaces for a long time
- Equations for temperature, moisture and winds on a 3D grid:

  Rate of change = grid-resolved air flow + other processes
- Usual approach: Expert-designed parameterizations represent other processes: clouds/rain, turbulence, radiation, etc.
- These processes are complex, but mostly work column-wise
- AI2-CM approach: Correct or replace parameterizations using column-wise ML trained on a reference data set.

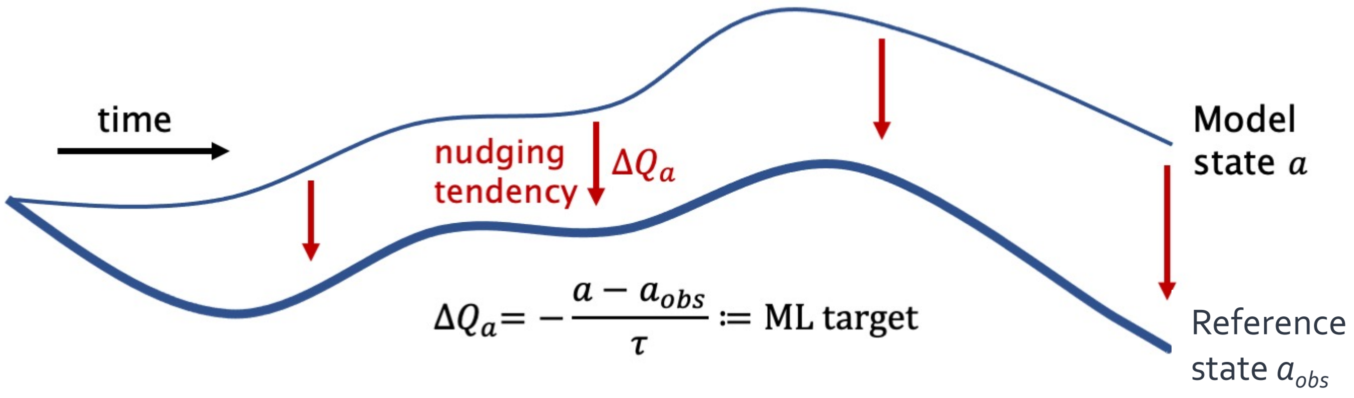


ML goal: Make coarse-grid model evolve like reference

Reference datasets:

- Reanalysis (present-day climate, data-based) - *Watt-Meyer et al. 2021, GRL, doi:10.1029/2021GL092555*
- **Fine-grid model** (range of climates, uncertain biases) – *Bretherton et al. 2021, JAMES, submitted (ESSOAr)*

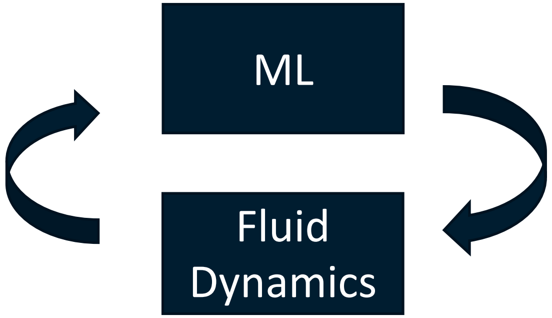
Corrective ML Method: 'Nudge' coarse model state to the reference state on a 3-6 hour timescale and machine-learn the 'nudging tendencies' that do this.



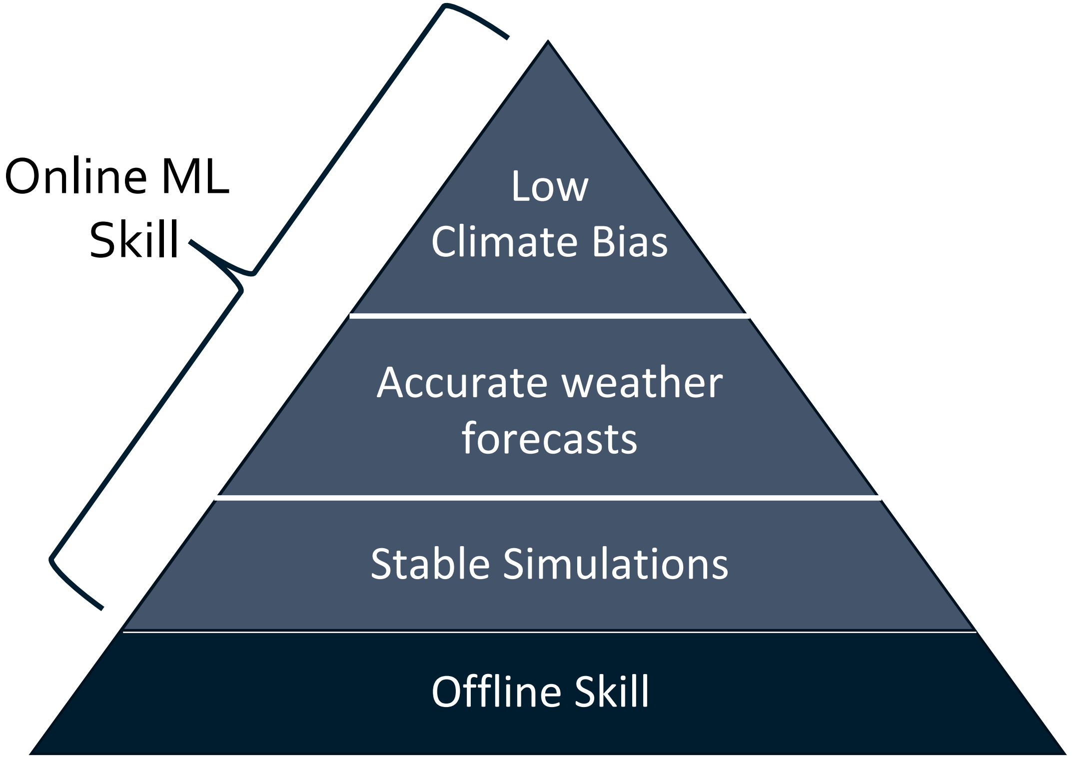
- This methodology can naturally transfer to any coarse-grid target model and fine-grid reference

Challenge of ‘hybrid’ ML coupled to other components

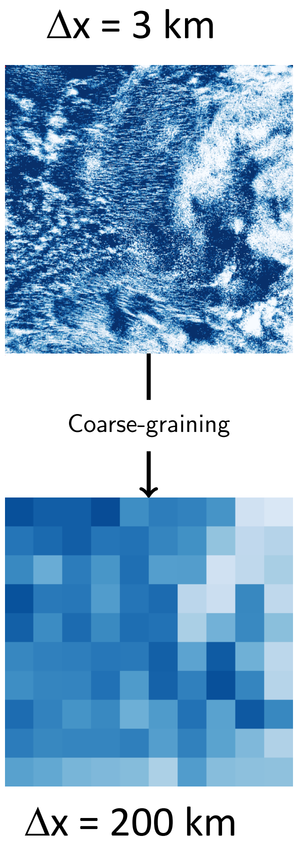
Coupled to fluid dynamics and parameterized physics



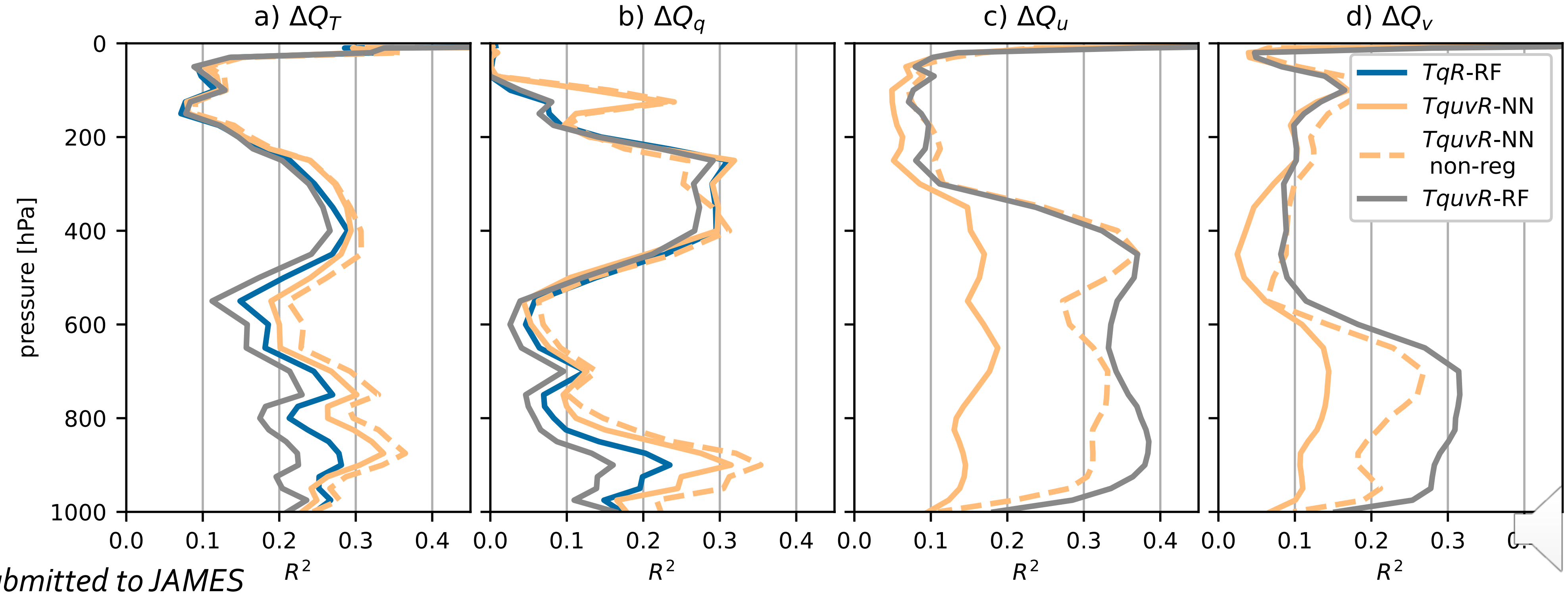
Training $\neq$ Testing
(offline) (online)



Corrective ML skill: Good on time-mean, modest on variability



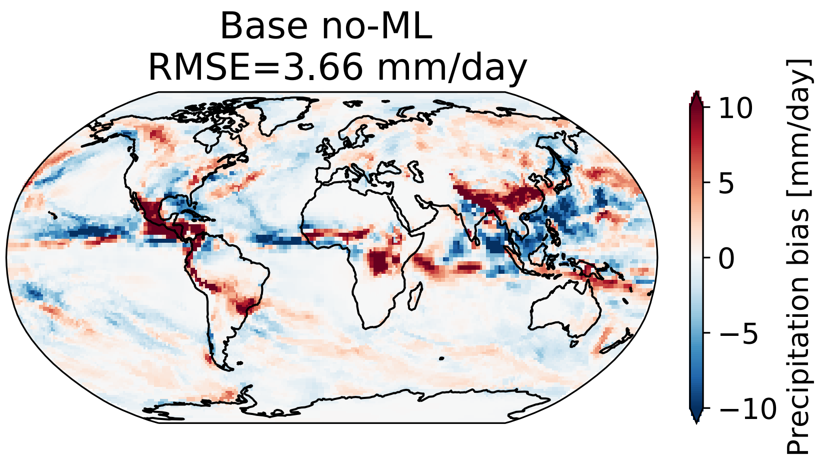
- Training run nudged to coarsened 3 km X-SHiELD, 40 days
- ML for nudging tendencies of *T*, *q*, and optionally *u*, v, and surface downwelling radiation *R*;
- ML inputs: column *T*, *q*, *u*, *v*, cos(*zenith*), z_{sfc}
- Google Cloud workflow with custom Python wrapper for FV3GFS (*McGibbon et al. 2021, GMD*)



Bretherton et al. 2021, submitted to JAMES

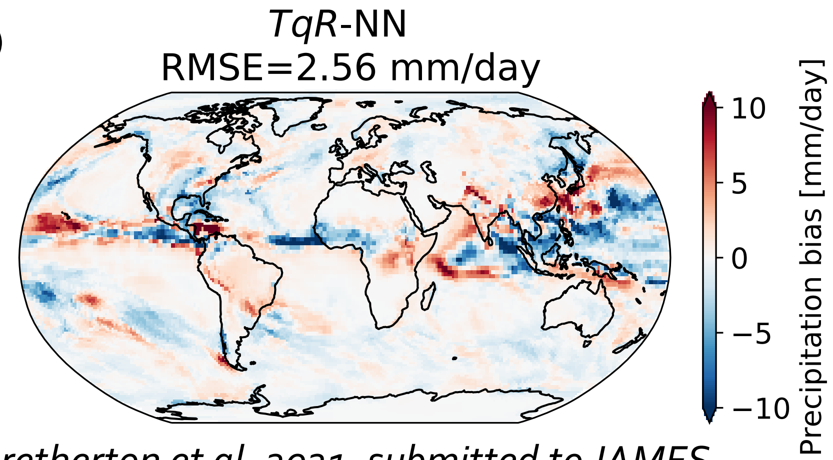
‘Nudge-to-fine’ ML reduces climate bias vs. reference

a)



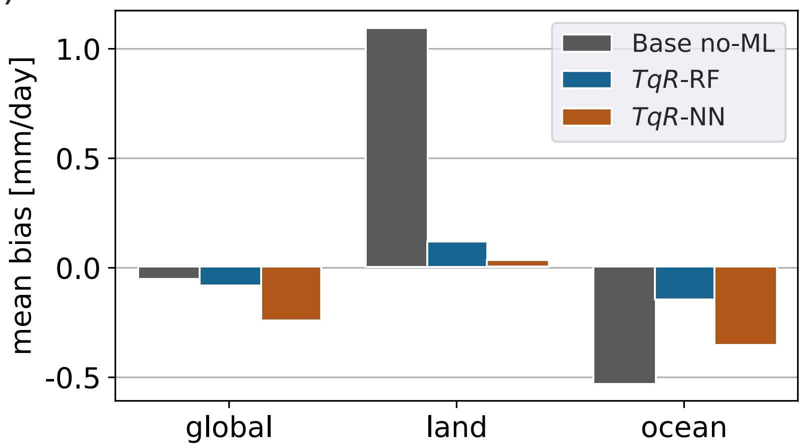
- Reference for prognostic runs (like training): 3 km X-SHiELD, 40 days
- RF or 3-layer neural net reduces time-mean precipitation error vs. reference by 30%.
- ML for surface radiation removes land precip bias

c)

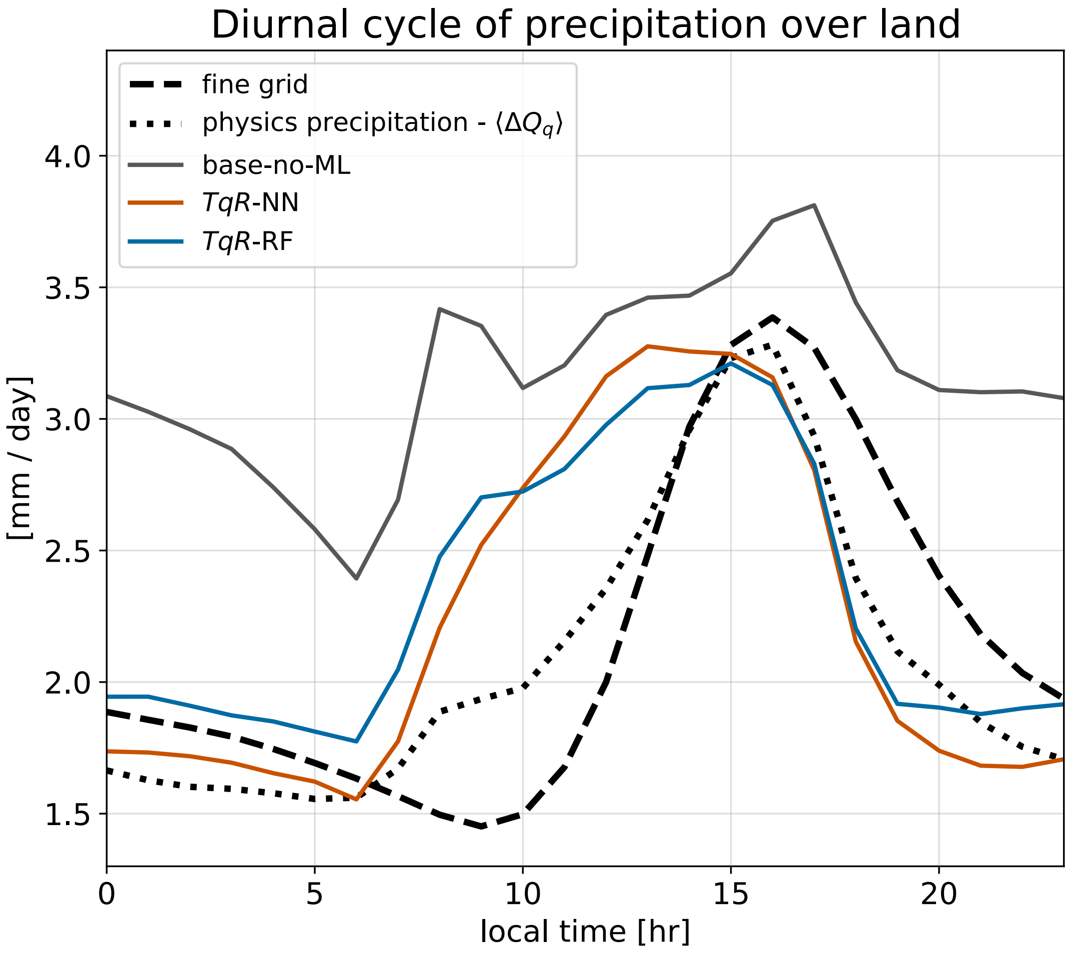


Bretherton et al. 2021, submitted to JAMES

d)



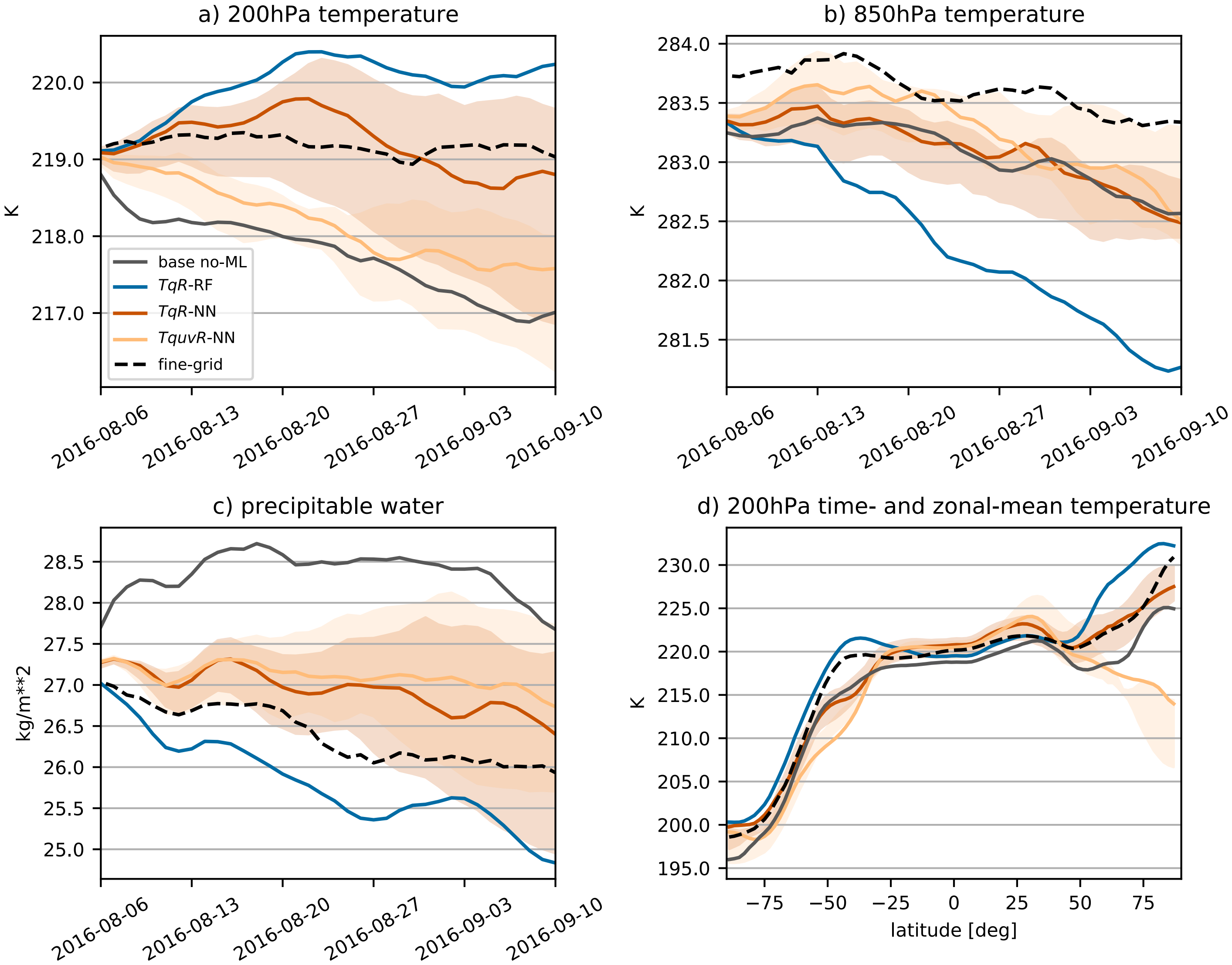
ML improves land precipitation diurnal cycle amplitude



Bretherton et al. 2021, submitted to JAMES

- Baseline simulation: too weak
- ML-corrected amplitude is good but phase is 3 hrs early.
- Diurnal cycle error is:
  - Half from nudging-based training
  - Half from ML

Mean-state drifts



- Drifts of *TqR*-NN are less than for baseline model, esp. in lower troposphere (T850, PW).
- Drifts are sensitive to initial random seed for NN (range of dark brown shading)
- Wind nudging (*TquvR*-NN) induces upper-tropospheric temperature drift

Bretherton et al. 2021, submitted to JAMES

Ongoing work

- Corrective ML trained and run for multiple climates (ΔSST = -4K, 0, 4K, 8K)
  - Reference fine-grid model: 25 km FV3GFS (runs fast), same physics as 200 km target
  - ML-corrected 5-year run reduces land surface T and precip biases vs. 200 km baseline across all climates, but stability and performance sensitive to random seed
  - See our poster A15E-1683 by Clark et al. for details
- Corrective ML trained on a year-long X-SHiELD 3 km training run from GFDL.
  - Prognostic 200 km runs corrected with some ML configurations can run for 2 years
  - Double-ITCZ bias and upper-tropospheric temperature drifts are still problem areas
- Prognostic simulations with ‘fine-only’ ML of full fine-grid physics: Fast PW drifts

Conclusions

- ‘Nudge-to-fine’ corrective ML trained with nudging of a coarse-resolution global atmosphere model to a fine-grid reference can improve its weather and climate skill
- In our example, time-mean precipitation distribution was improved 30%.
- The nudge-to-fine method generalizes easily to any global model.
- Two keys to its success:
    1. The coarse model physical parameterizations help maintain out-of-sample stability of the ML-corrected model
    2. The nudging framework avoids jolting the coarse model during training
- Controlling prognostic stability and climate drift remain challenging.