A RENAISSANCE IN CLOUD SUPERPARAMETERIZATION

Emerging flexibilities for regionalization & machine learning opening new CESM science opportunities.

Mike Pritchard
Associate Professor
University of California, Irvine

Model biodiversity is healthy for global climate simulation

JAMES | Journal of Advances in Modeling Earth Systems

Commentary | Open Access

The Fall and Rise of the Global Climate Model

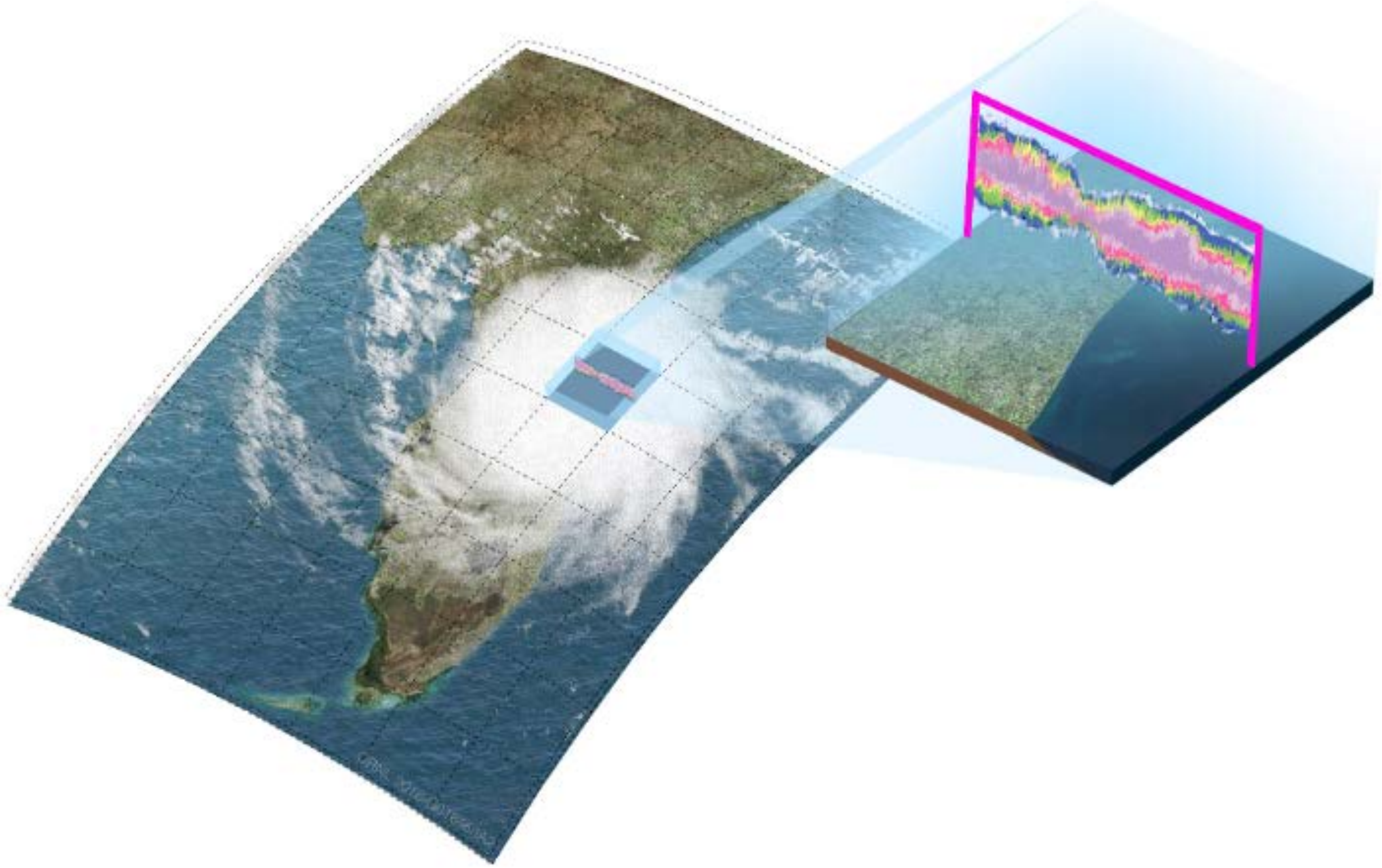
Johannes Mülmenstädt, Laura J. Wilcox,

First published: 29 August 2021 | https://doi.org/10.1029/2021MS002781

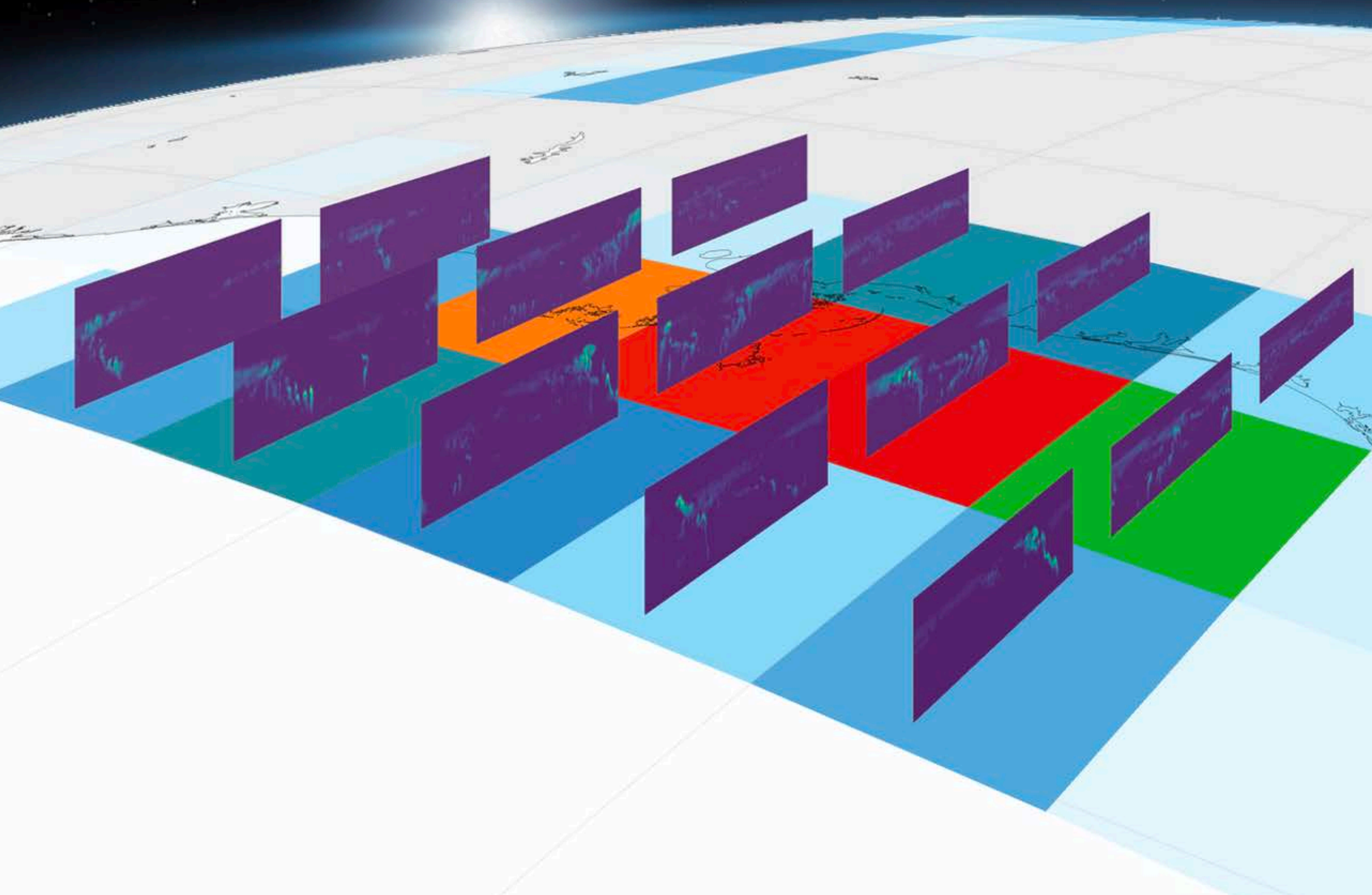
Get it at UC

SP-CESM: A Multiscale Modeling Framework

SP: "SuperParameterization" i.e. embedding small samples of explicit hi-res physics in a GCM.

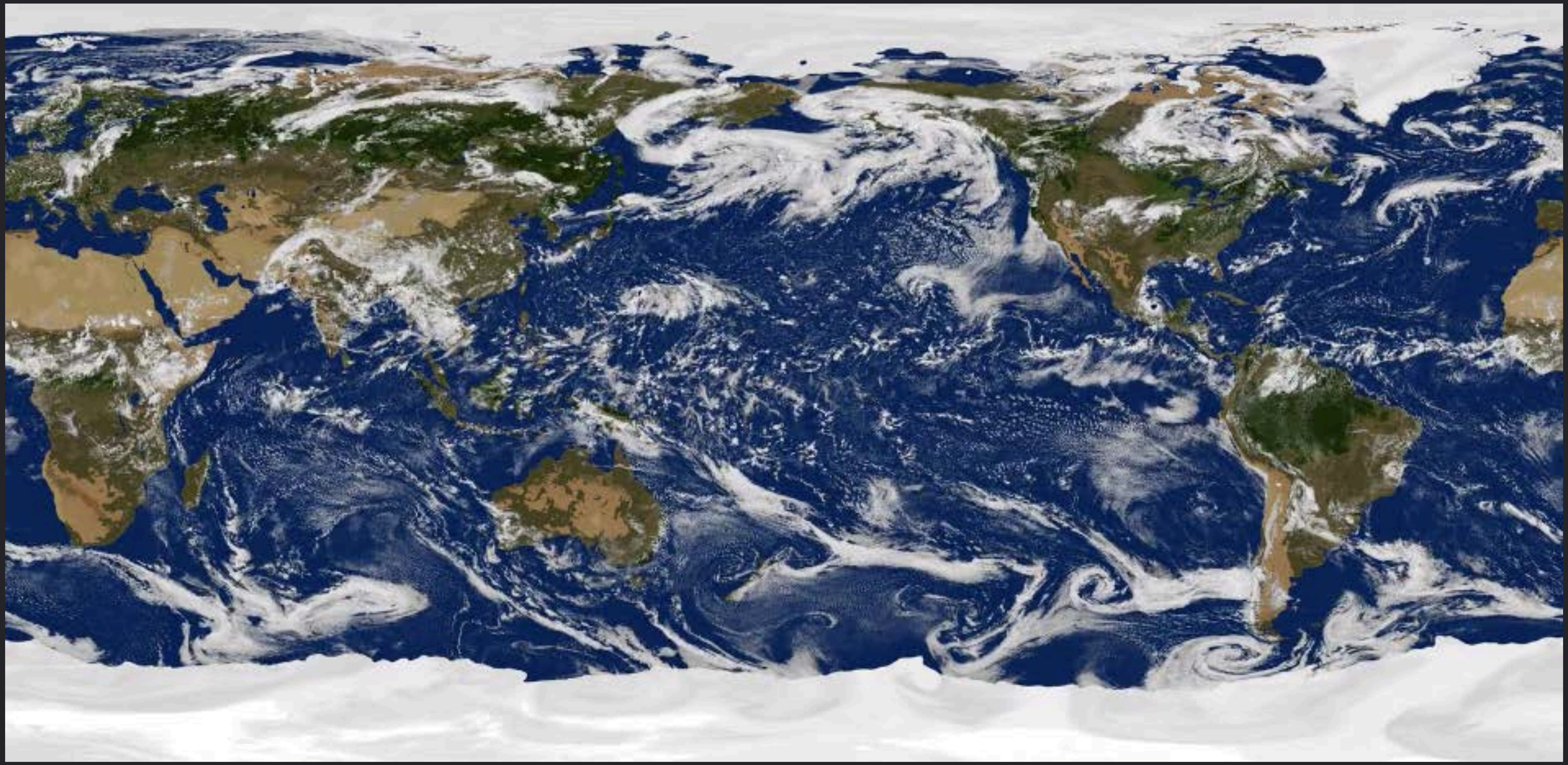


MMF: Thousands of "micro-models" of fast, fine-scale physics embedded in a host planetary atmospheric simulator.



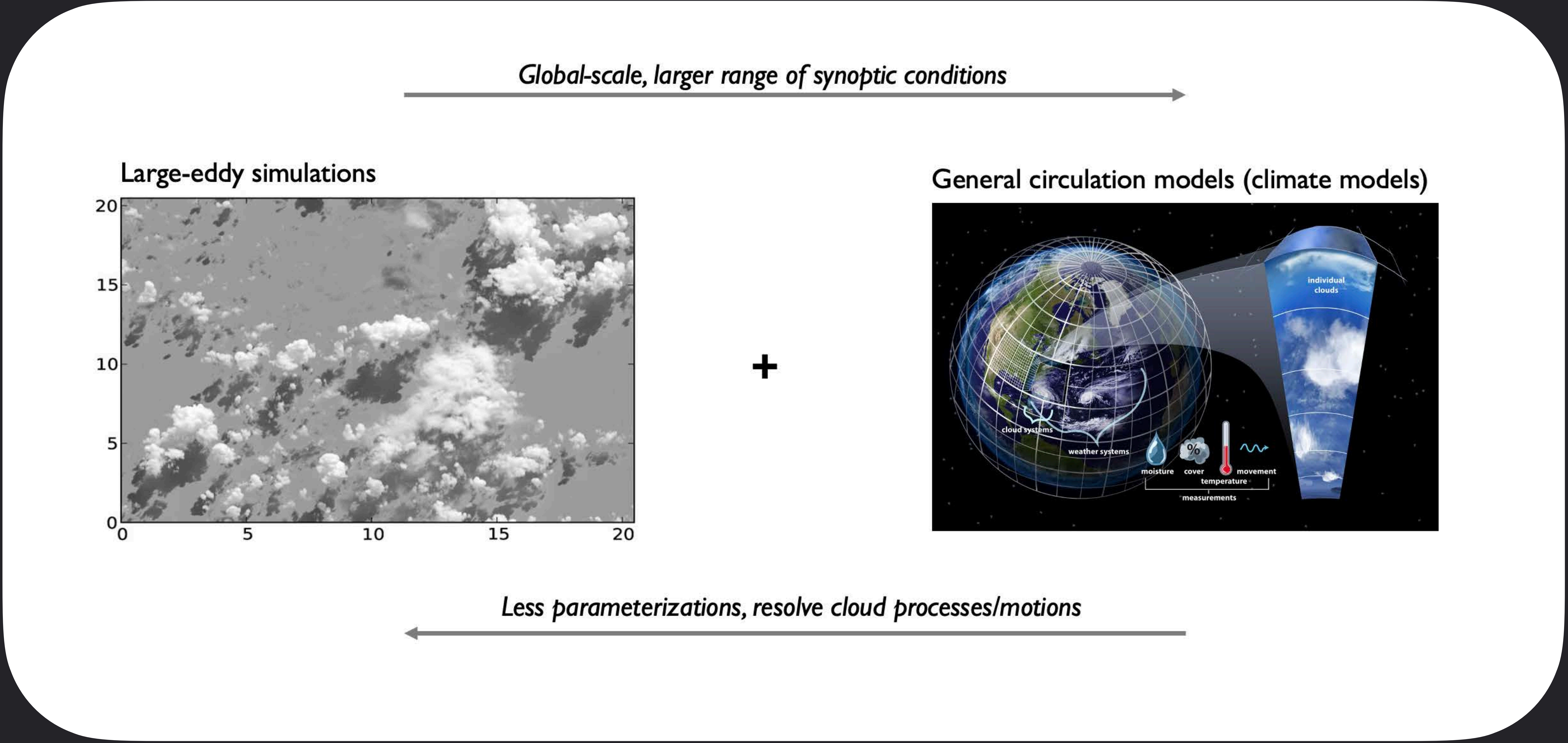
- Embeds spatio-temporal complexity.
- 100x more expensive than standard climate simulation.
- Explicit cloud physics emerging from appropriate governing equations.
- Low dimensional arteries that interface easily with climate model.
- Simple targets for machine learning.

In the past decade the 1-4 km resolution regime was a frontier that MMFs helped explore ahead of schedule.



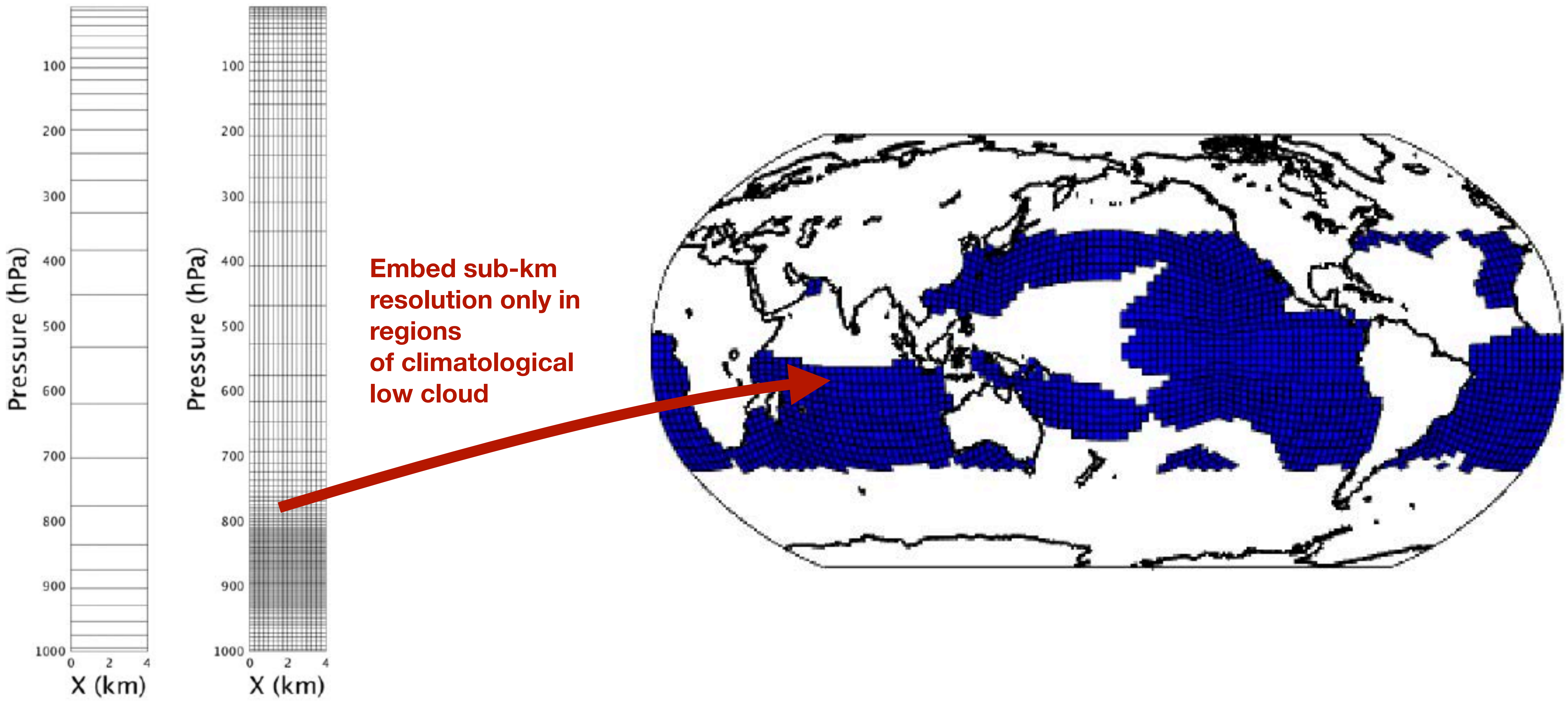
Now Global Storm Resolving Models handle this more elegantly, but only for the computationally privileged.

This decade, a new generation of high-res MMFs can also help penetrate the boundary layer turbulence frontier.

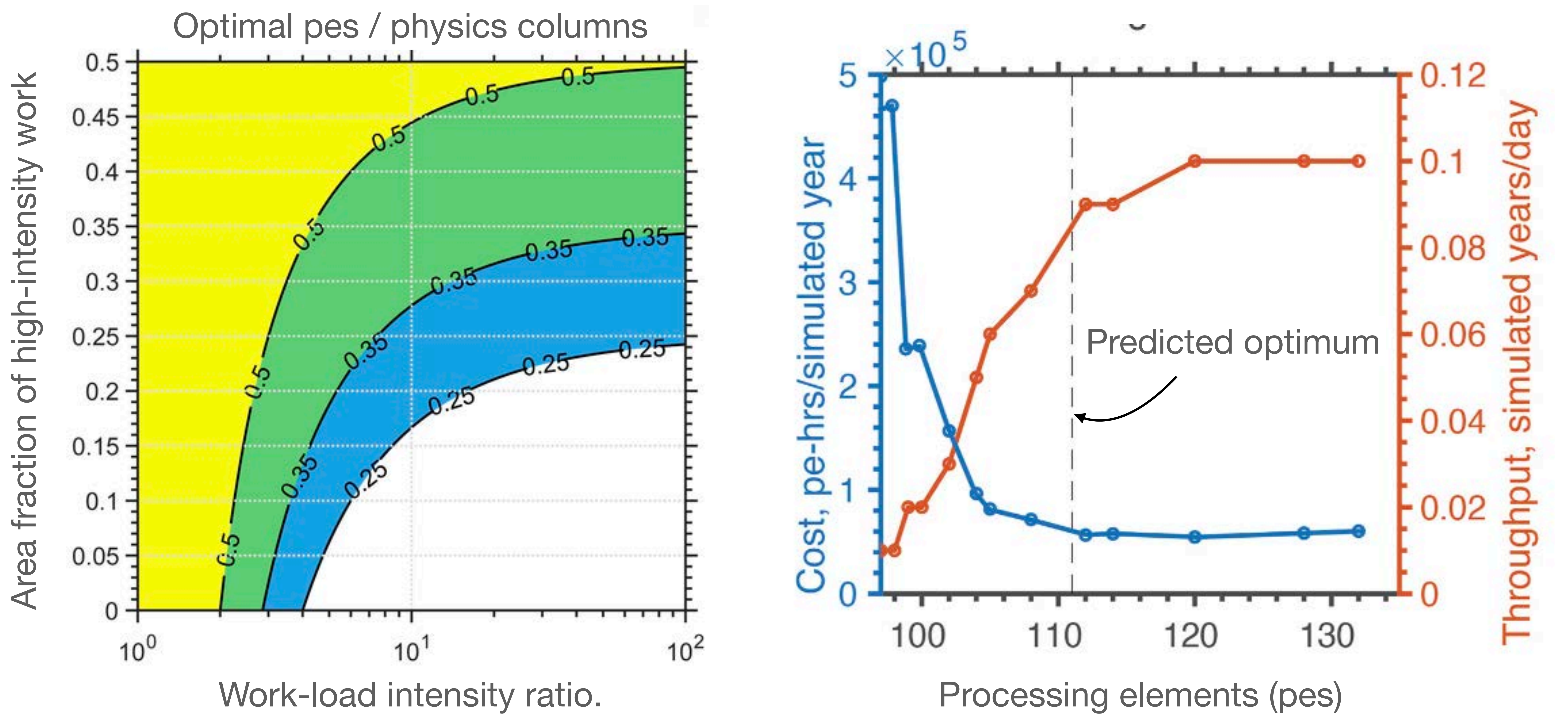


Meanwhile regionalizing 1-4 km can democratize & open flexible opportunities.

Example: Regionalizing high resolution MMF, proof of concept.

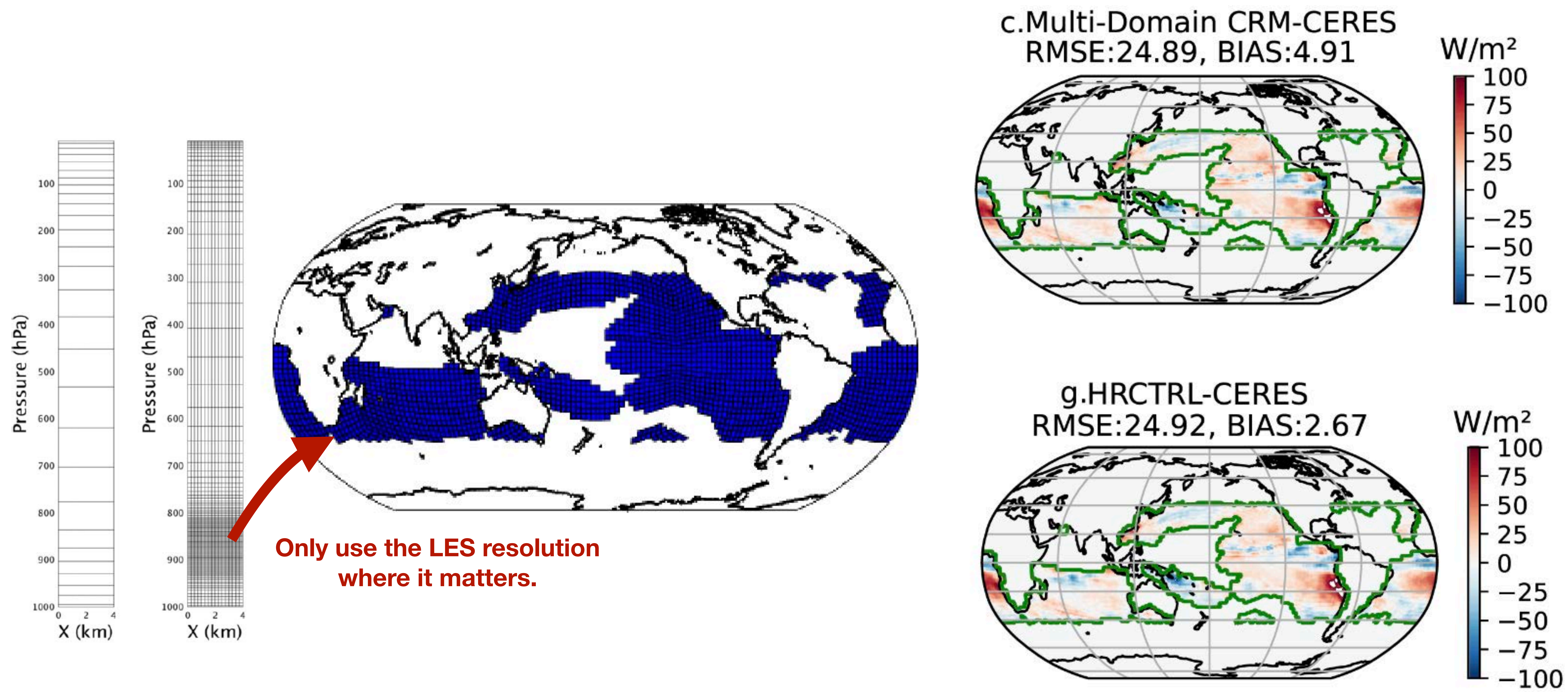


Adapting atm physics to load balance regionally intense calculations to maximize throughput and efficiency.



Peng, Pritchard et al., *JAMES*, 2022.

Same shortwave biases as uniform high-res but faster and cheaper. (> 80% of compute devoted to < 30% of planet)

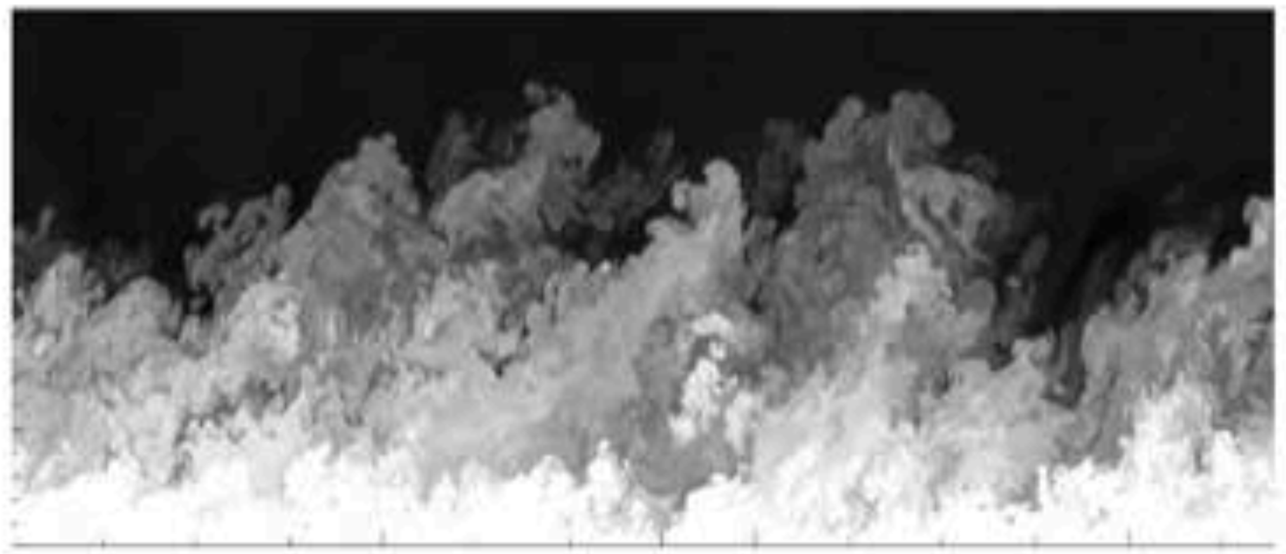


Peng, Pritchard et al., *JAMES*, 2022.

Upshot: A new way to regionally focus convective permitting resolution in subregions for niche impacts, dynamics studies.

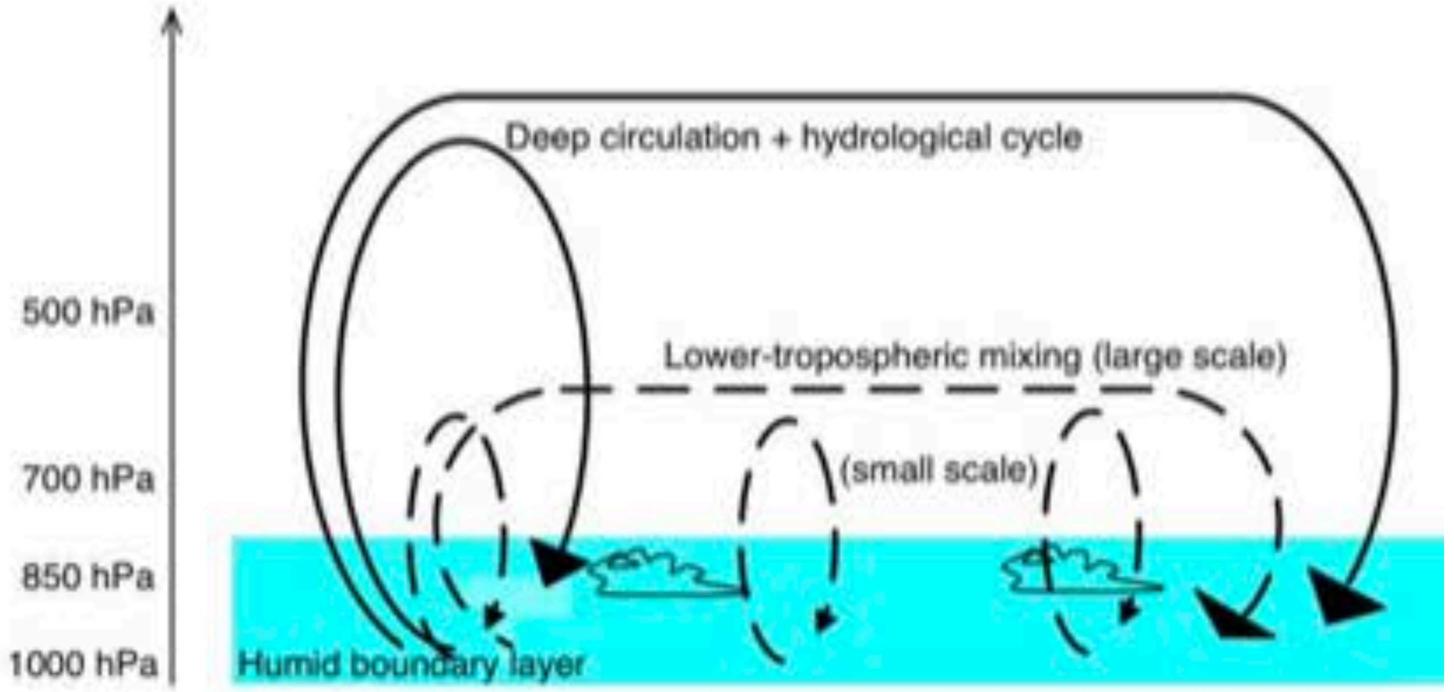
Problem: Lack of sub-km physics should keep us up at night.

Boundary layer turbulence

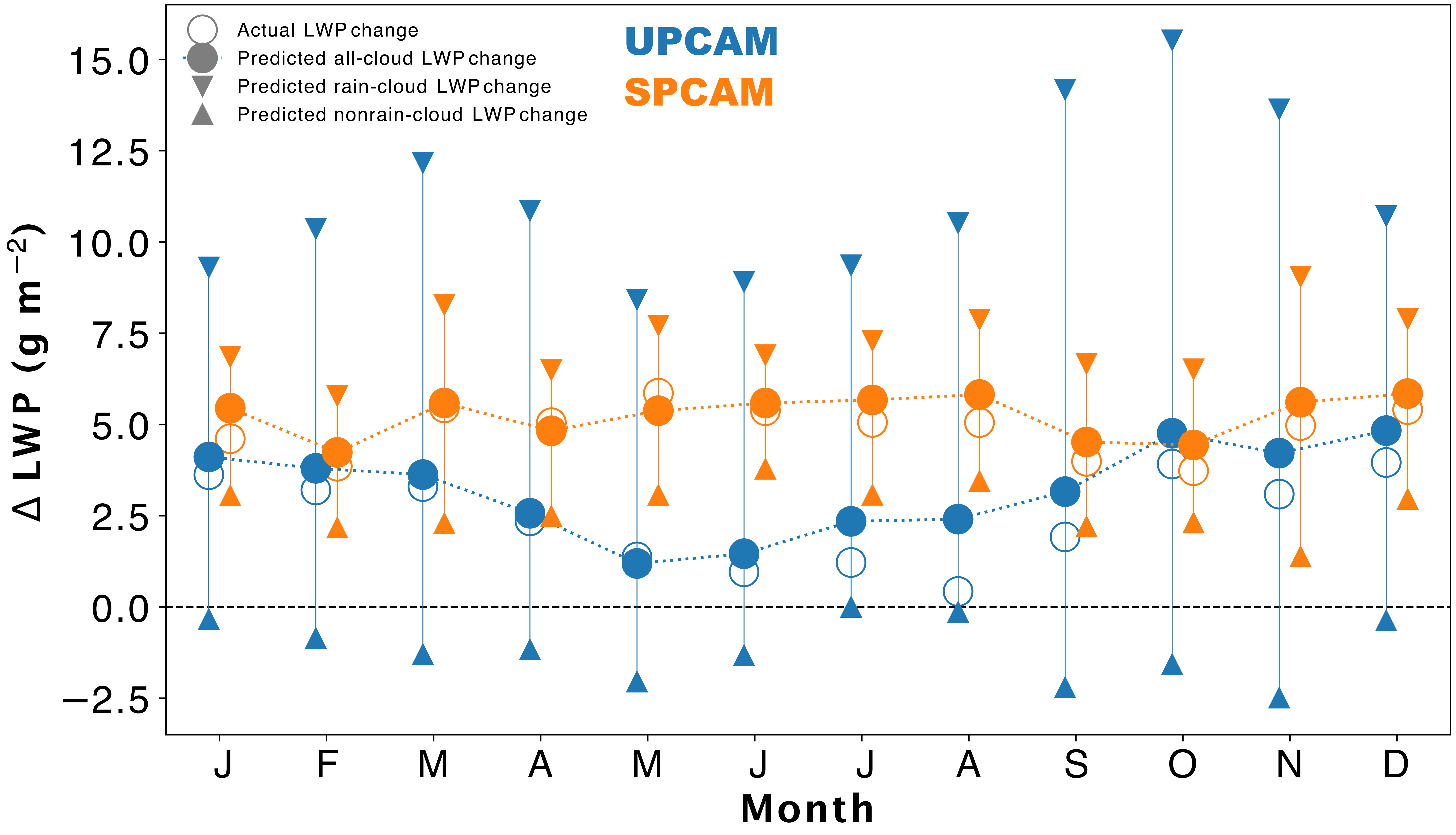


Low cloud formation & aerosol-cloud interaction

Mixing between the PBL & free troposphere

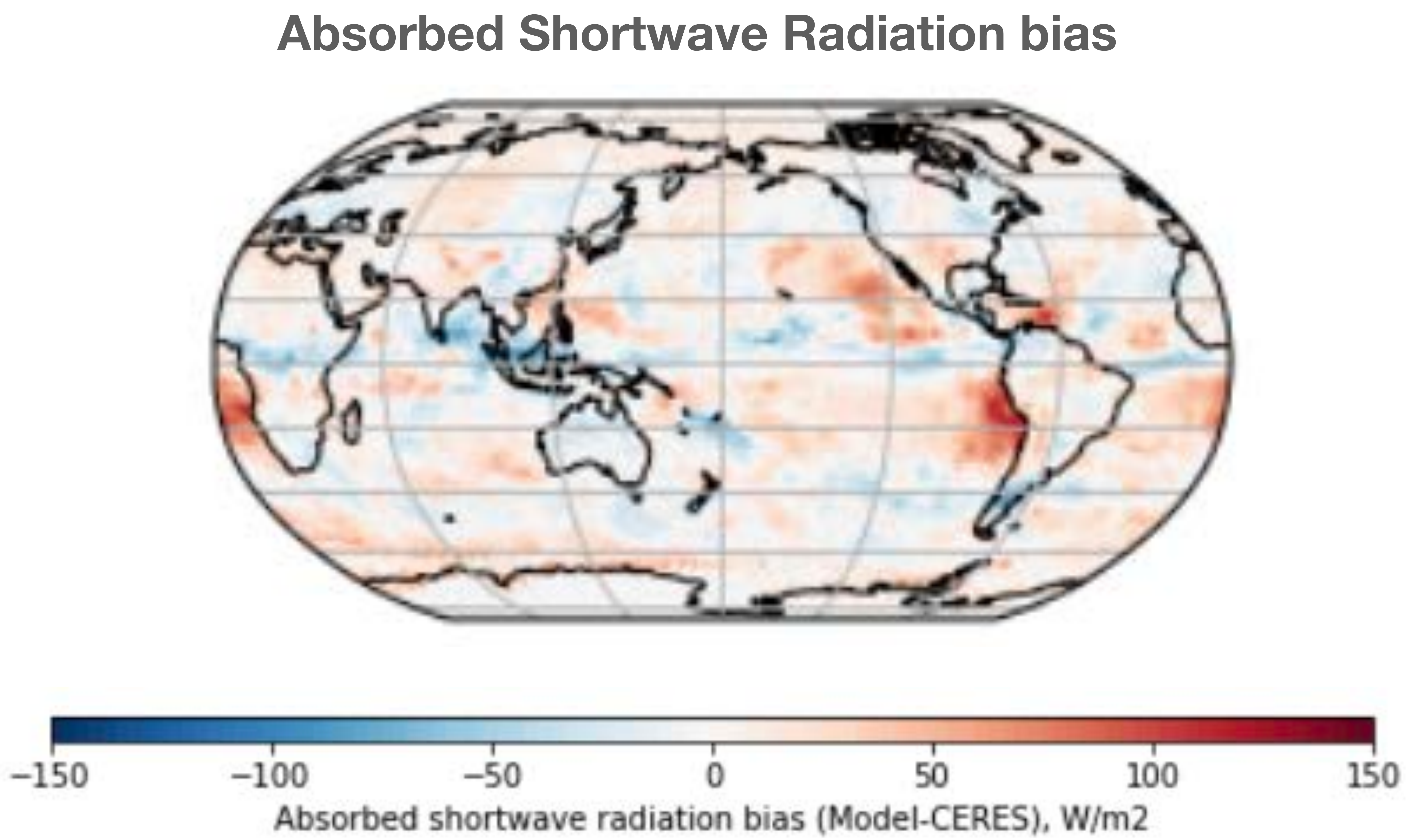


High-res MMF modifies ACI: Regime-dependence of aerosol-cloud feedback emerges.



See Terai, Pritchard et al., *JAMES*, 2020.

Problem: Even high-res MMF suffers a stratocumulus dim bias.



We are learning to tune the inner models to control chronic overentrainment.

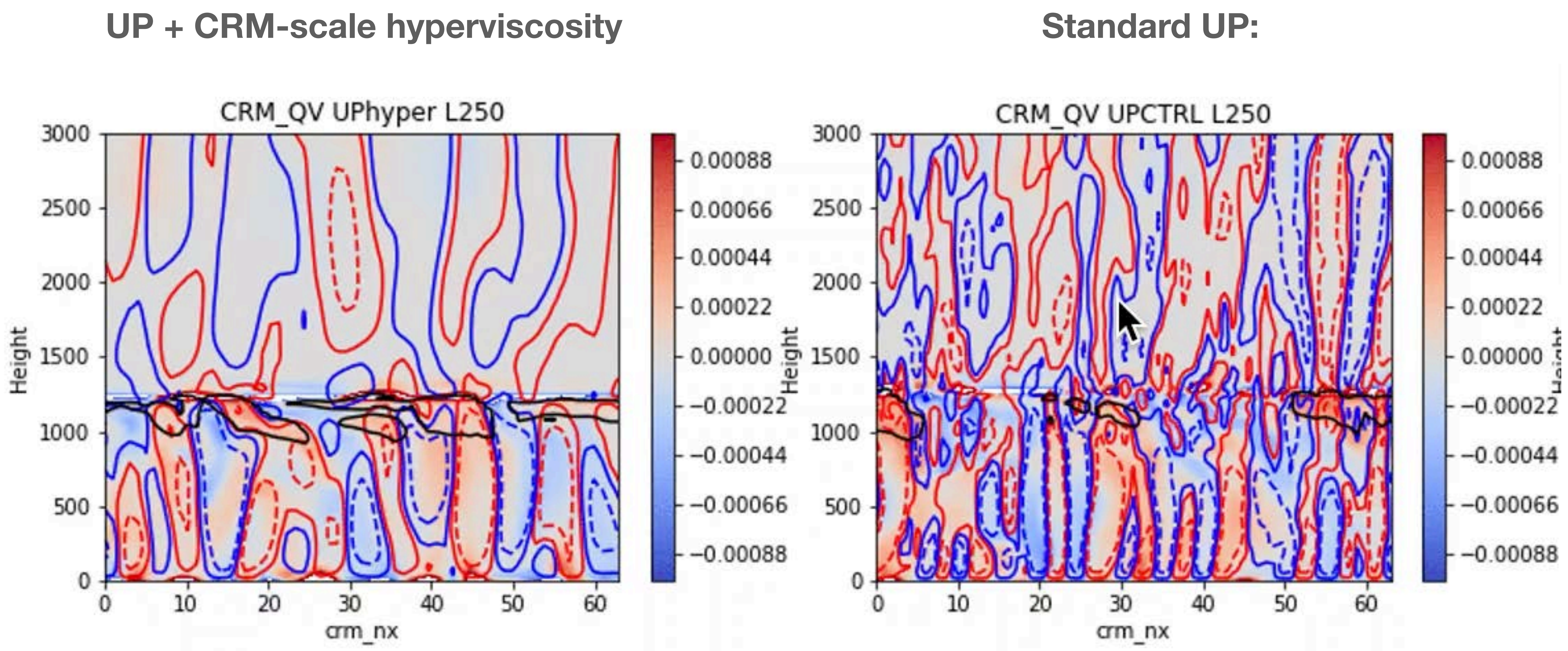


Figure courtesy of Liran Peng, UCI Postdoc collaborating w/ UW.

Promising effects: Brightening MMF's marine stratocumulus.

Change in ASR due to tuning.

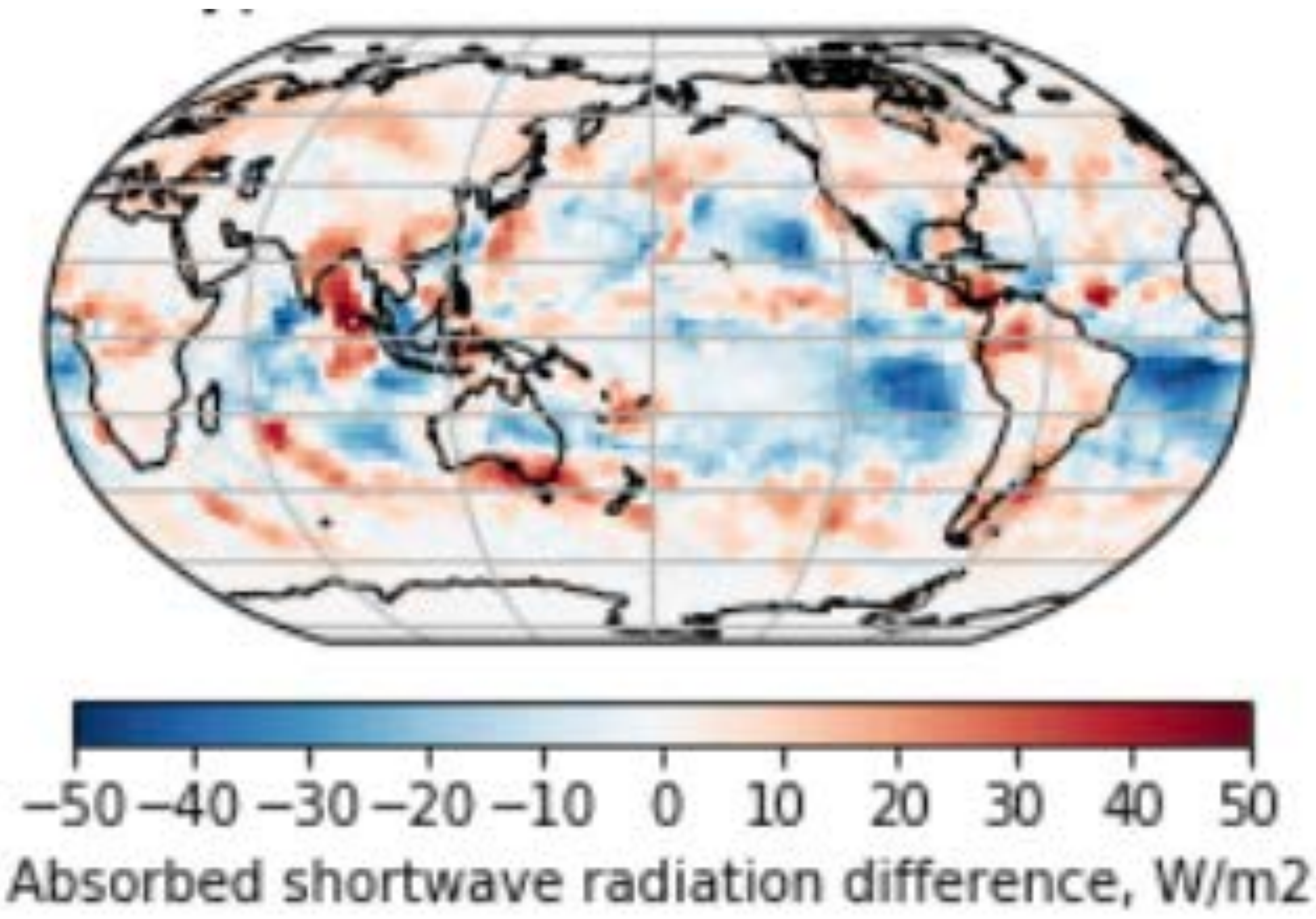
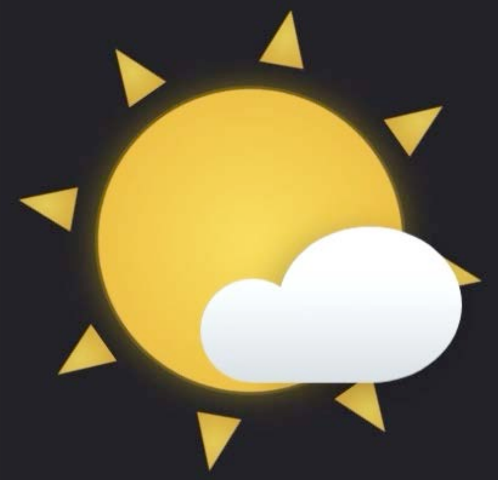


Figure courtesy of Liran Peng, UCI Postdoc collaborating w/ UW.

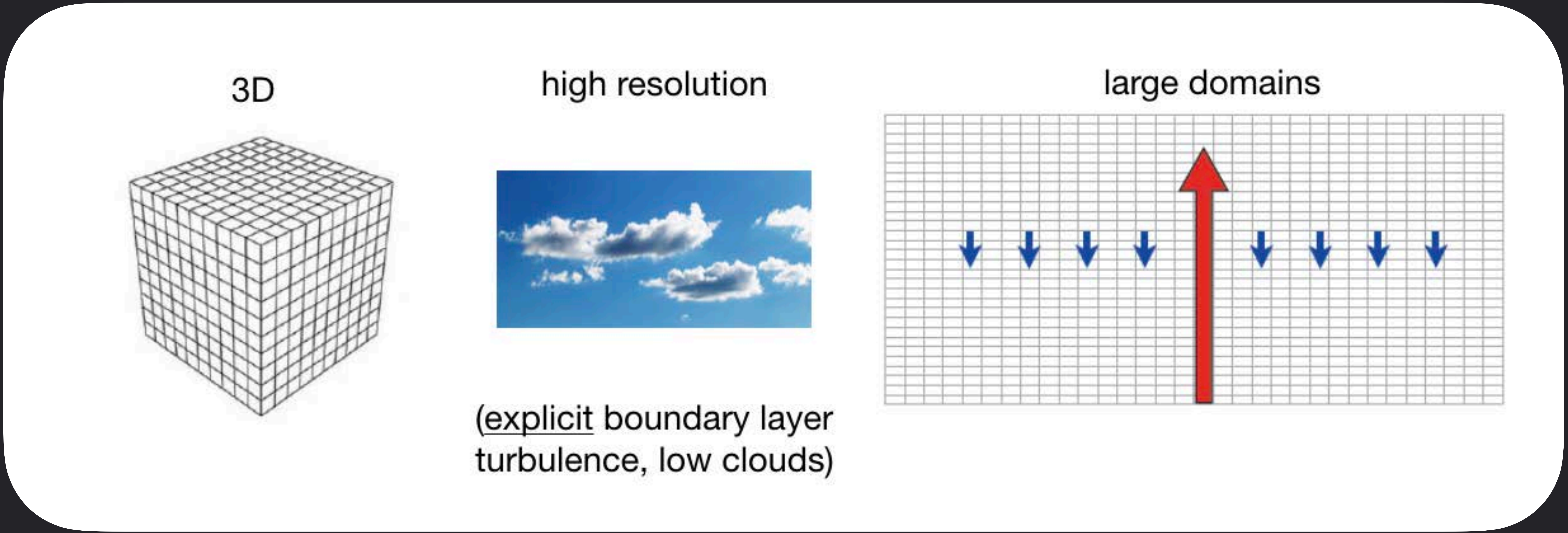


LOOKING AHEAD

In the era of
petascale computing

There is room to relax many unsatisfying compromises made in high-res MMF's prototype turbulence.

And explore the actual potential of high-res MMF



Raw computing power is not the issue — still exploding nationally including in NSF clusters beyond NCAR that CESM community could further exploit.

Example: “Frontera” at the Texas Advanced Computing Center
~ 8,300 Intel Cascade Lake nodes (56 cores/nodes)

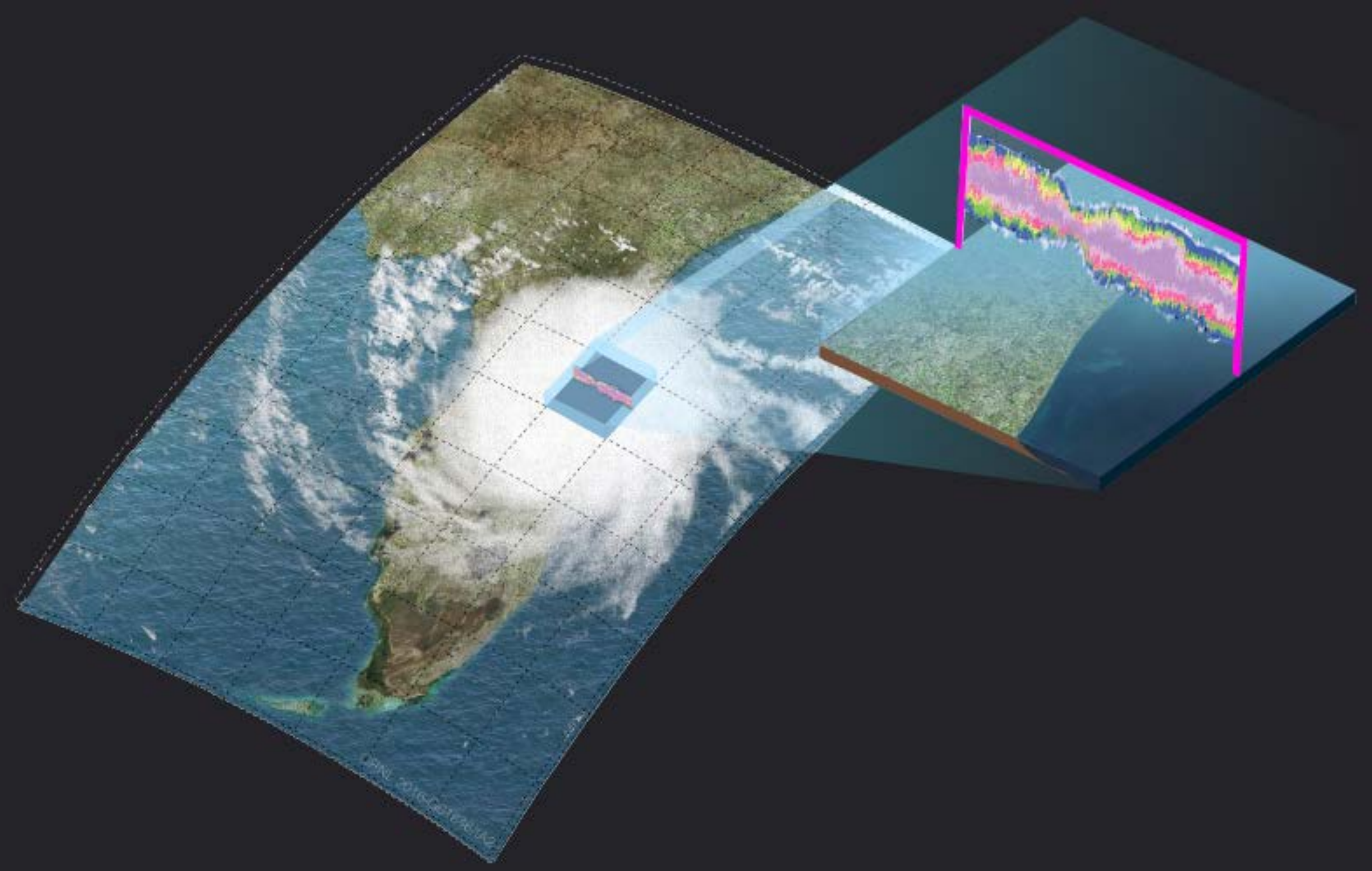
The actual potential of turbulence-permitting MMF has not been explored at parity with the model classes it is meant to complement.

Global cloud resolving model

Typical scale: ~ 100,000 cores
Max limit: ~ 500,000,000 cores

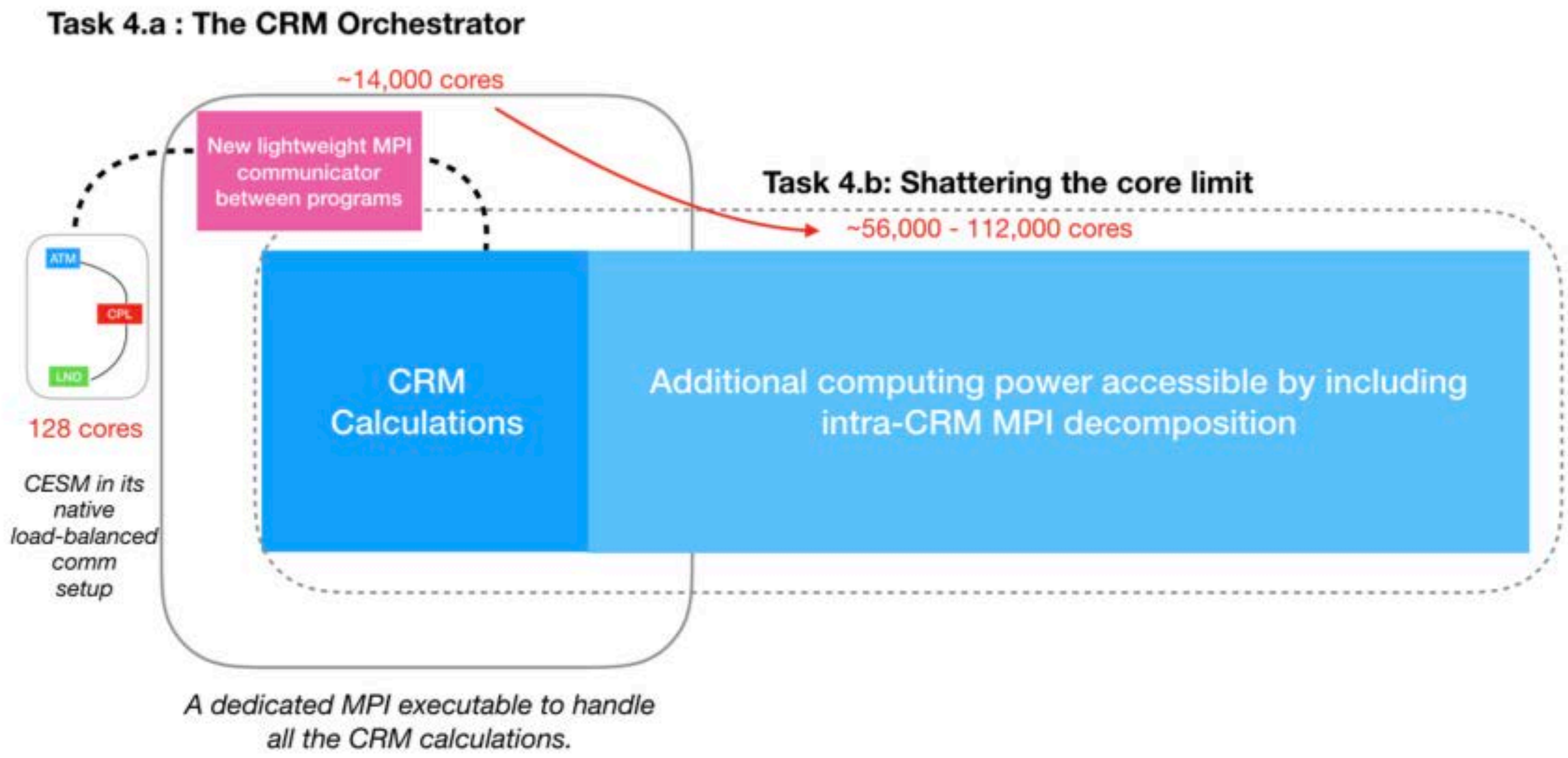
>>

2-degree high-res MMF
(i.e. ~ 14,000 embedded LES)

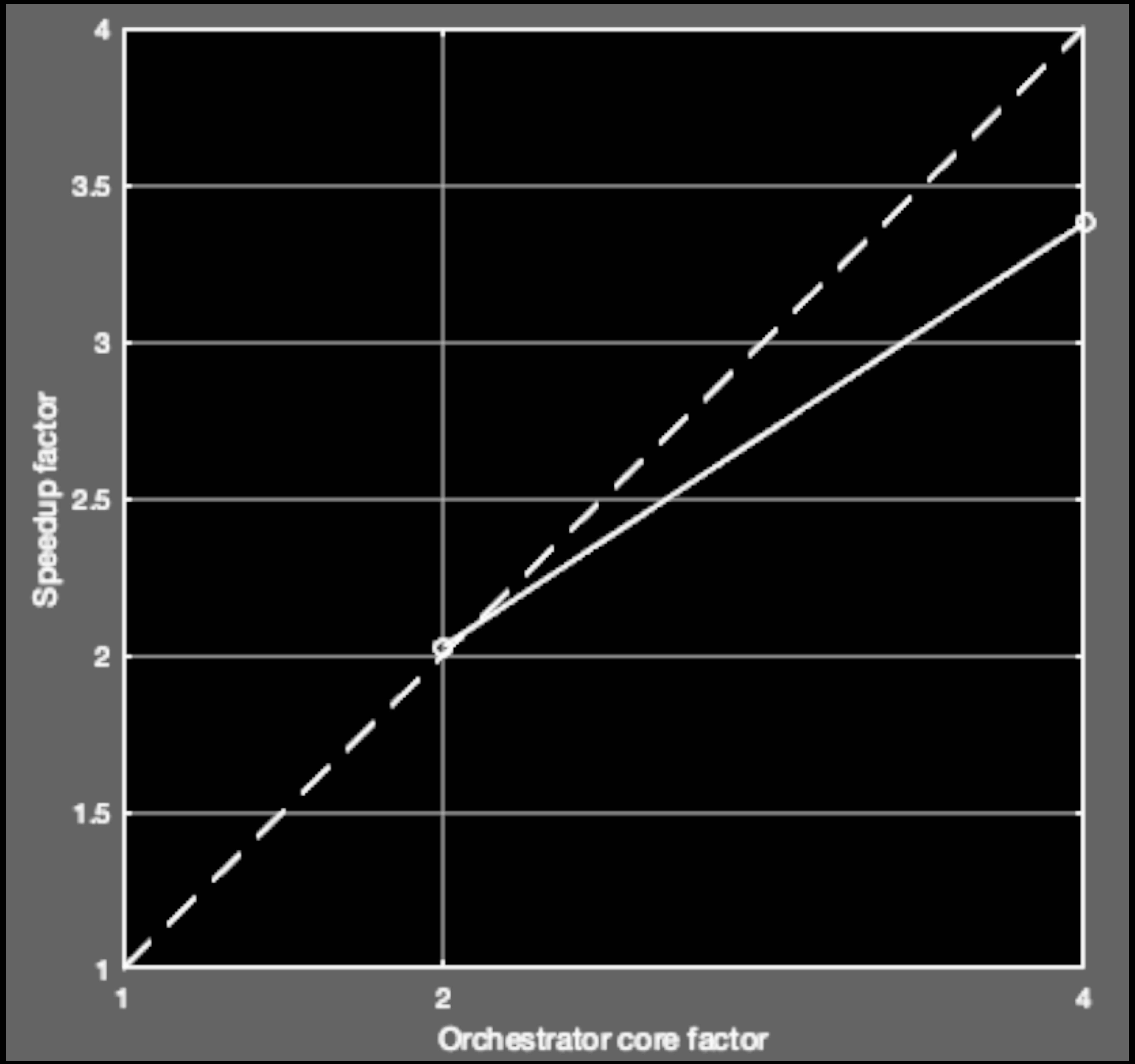


Typical scale: ~ 1,000-6,000 cores
Max limit: ~ 14,000 cores

New software: “Orchestrator” to outsource CESM MMF’s calculations to efficiently parallelize & access more cores.



This spring, Liran Peng & Peter Blossey have proved the concept.
Speeding up high-res CESM-MMF with quadruple the cores on Cheyenne.



Parallel decomposing cloud-resolving models

Exploding national HPC

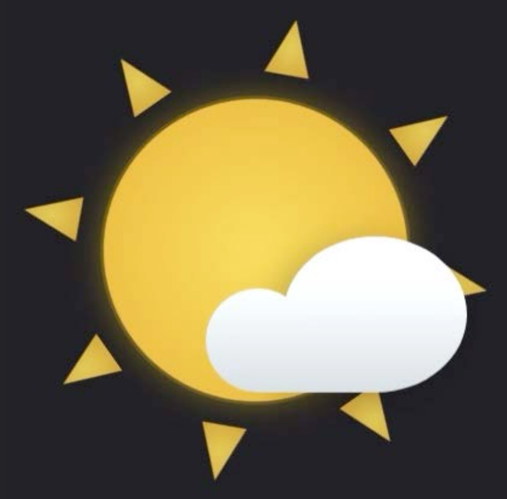
Regionalization (Multi-Domain CRM)

8-16x compute

3x compute savings

This decade, MMFs can help penetrate the boundary layer turbulence frontier.

How will our view of low-cloud and turbulence-mediated feedbacks change?



LOOKING AHEAD

In the era of
machine learning

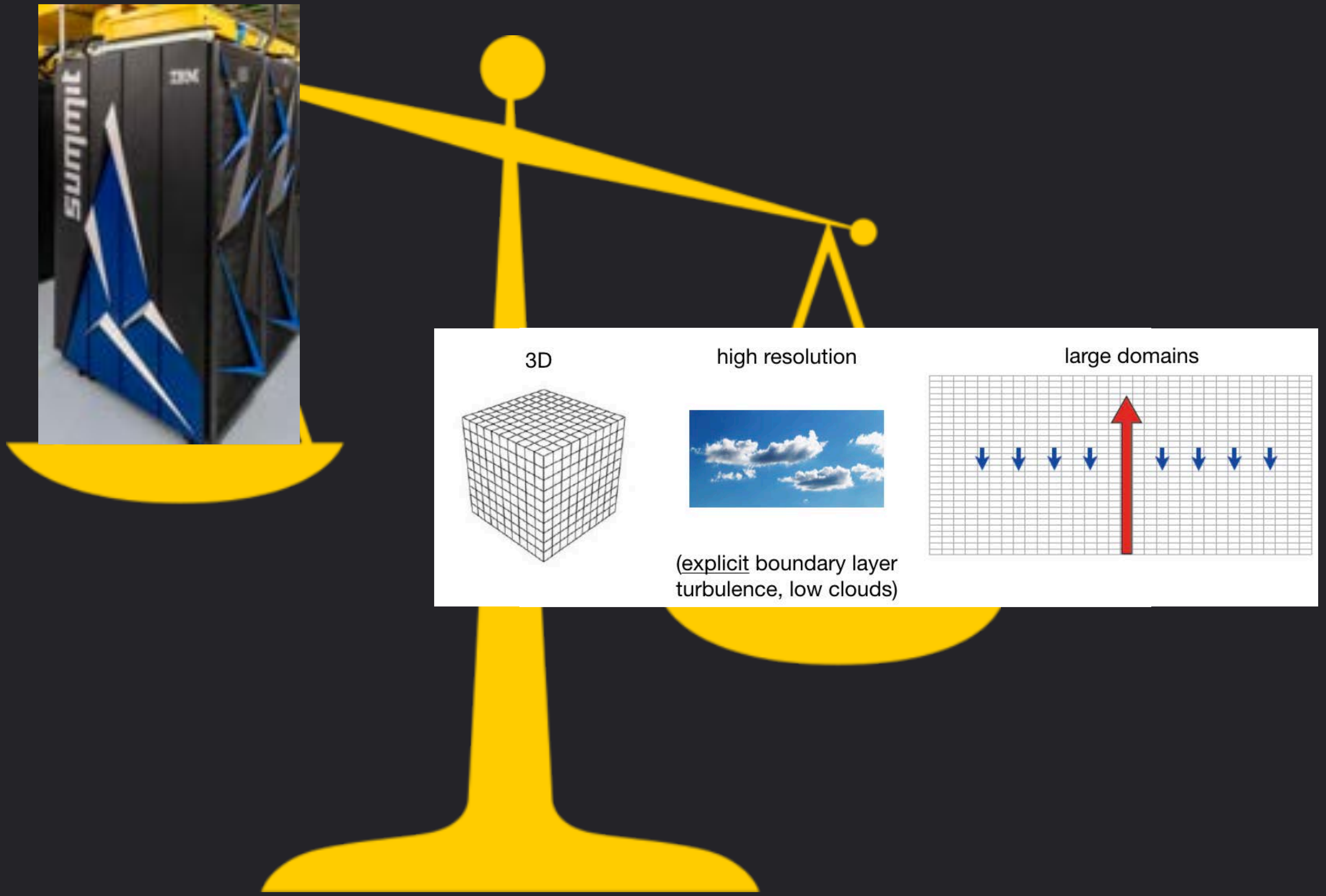
ML parameterization is exciting (might have breakthrough potential)

If the job is hard, e.g. simulating the whole atmosphere for decades...

...satisfying 3D turbulence calculations can seem too much even for powerful computers.

ML parameterization is exciting (might have breakthrough potential)

If the job changes to making short simulations just for training machine learning emulators...

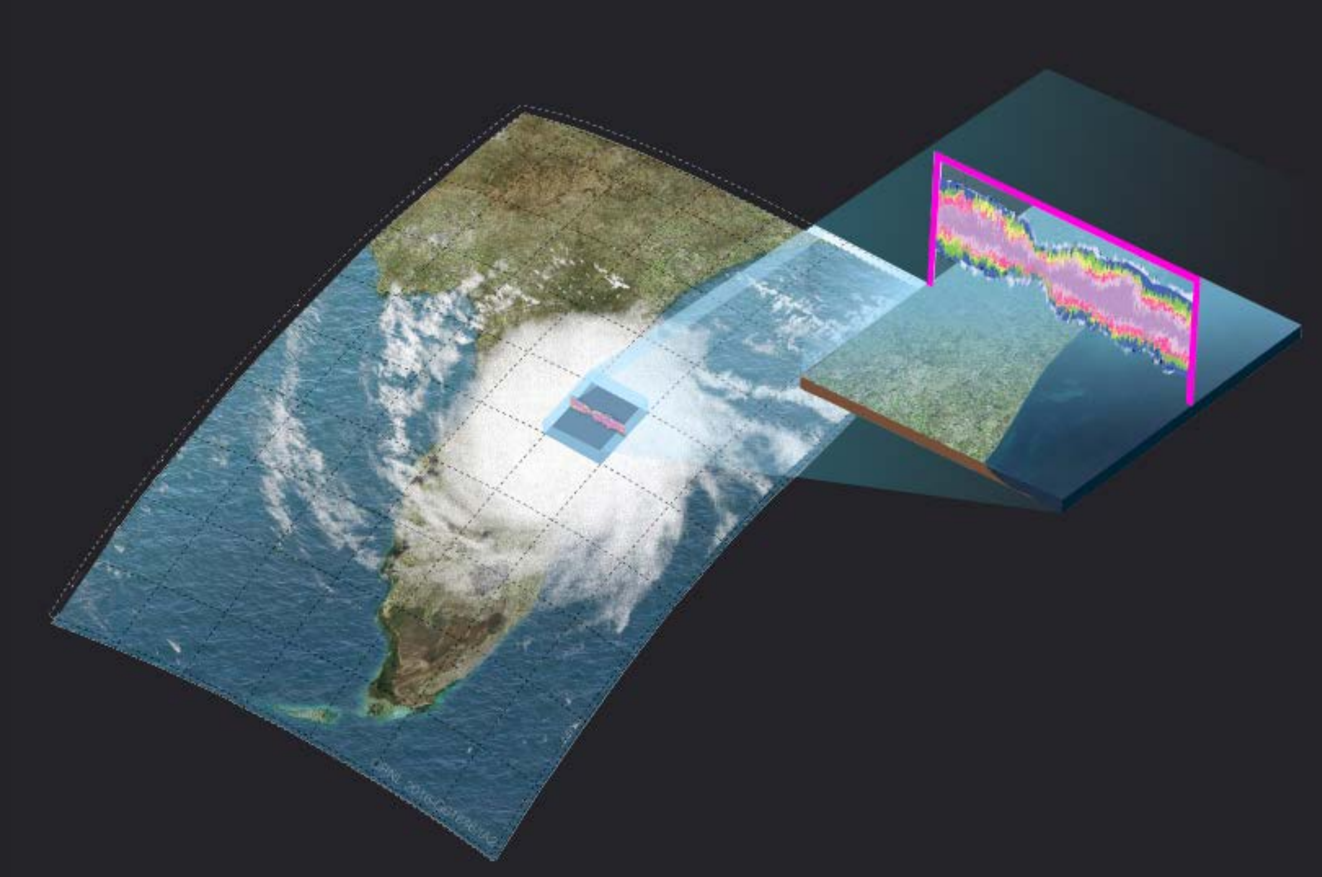


...we can do much more justice to turbulence physics.

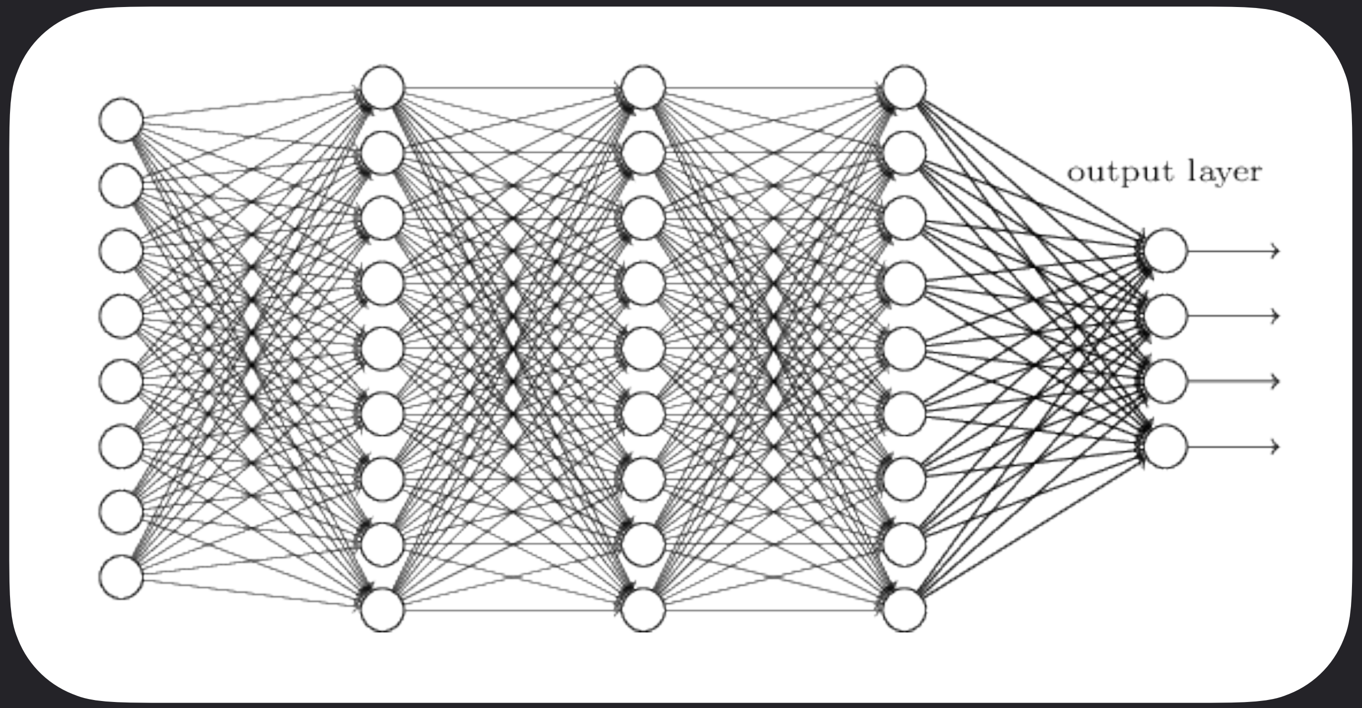
Is deep learning viable for emulating superparameterization?

Global aquaplanet testbed

Can 140,000,000 outputs from 1 year
of ~ 10,000 cloud-resolving models...



Be fit by a deep neural network?

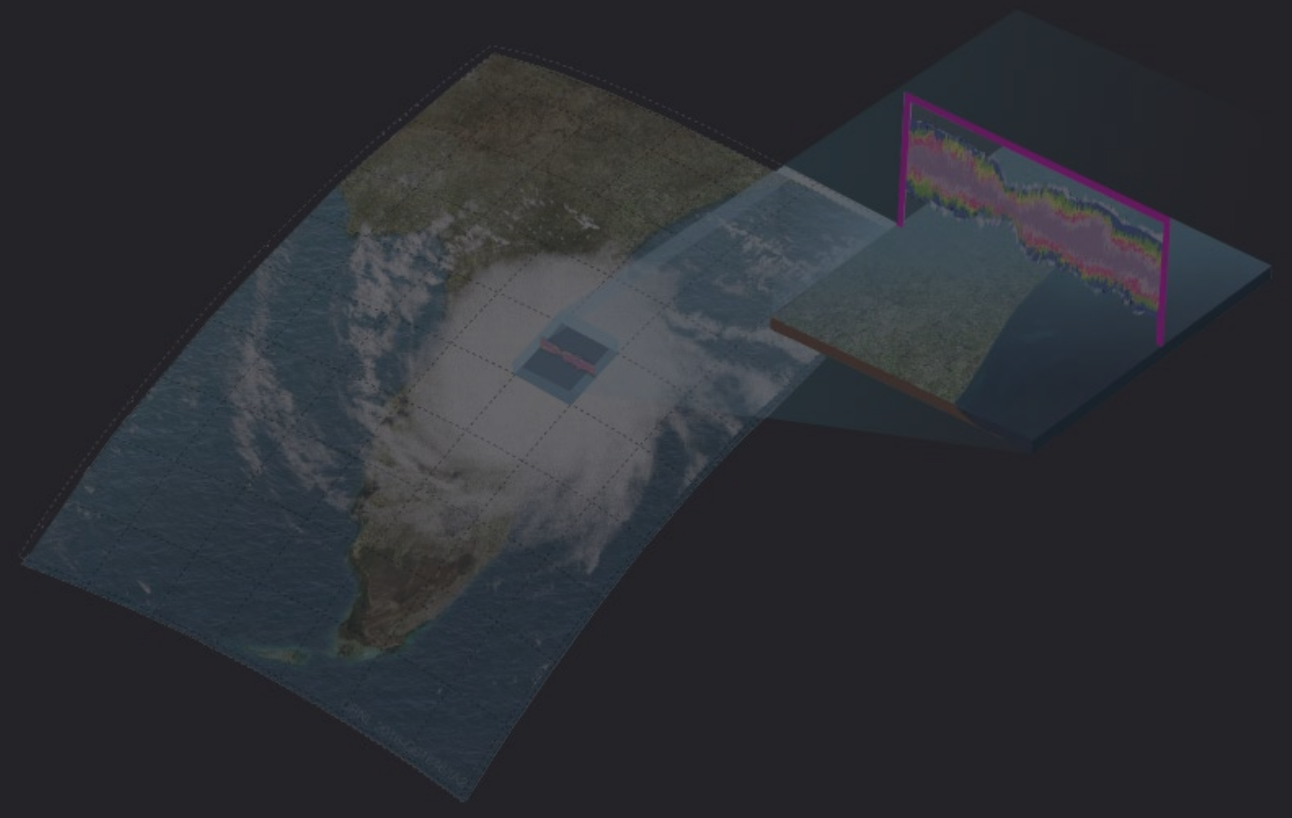


Gentine, Pritchard, Rasp et al., *GRL*, 2019.

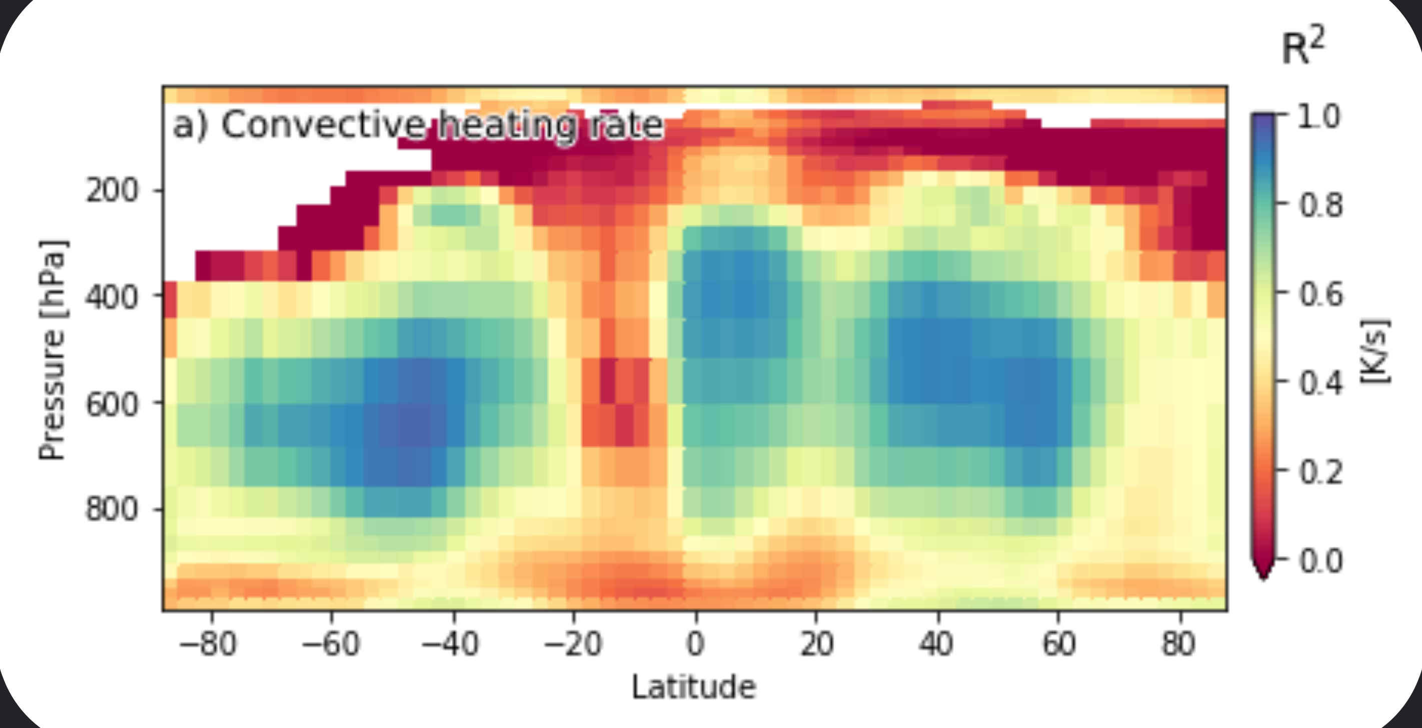
Can it work? Offline training skill achieved in 2017.

Global aquaplanet testbed

Can 140,000,000 outputs from 1 year of ~ 10,000 cloud-resolving models…



Be fit by a deep neural network?



Yes, e.g. $R^2 > 0.7$ for mid-tropospheric heating by convection and radiation.

Quite possibly!

The "Cloud Brain"

Encouraging ML proofs of concept from the MMF testbed

2017: Reasonable offline fits achieved with crude NNs in regions where it matters to the atmosphere.

2018: Prognostic tests that couple the NN to a host planetary model proved in concept. Realistic coupled wave spectra encouraging.

Rasp, Pritchard and Gentine, *PNAS*, 2018.

2019: Mass and energy conservation solved by architecture re-design; no trade-off with fit accuracy.

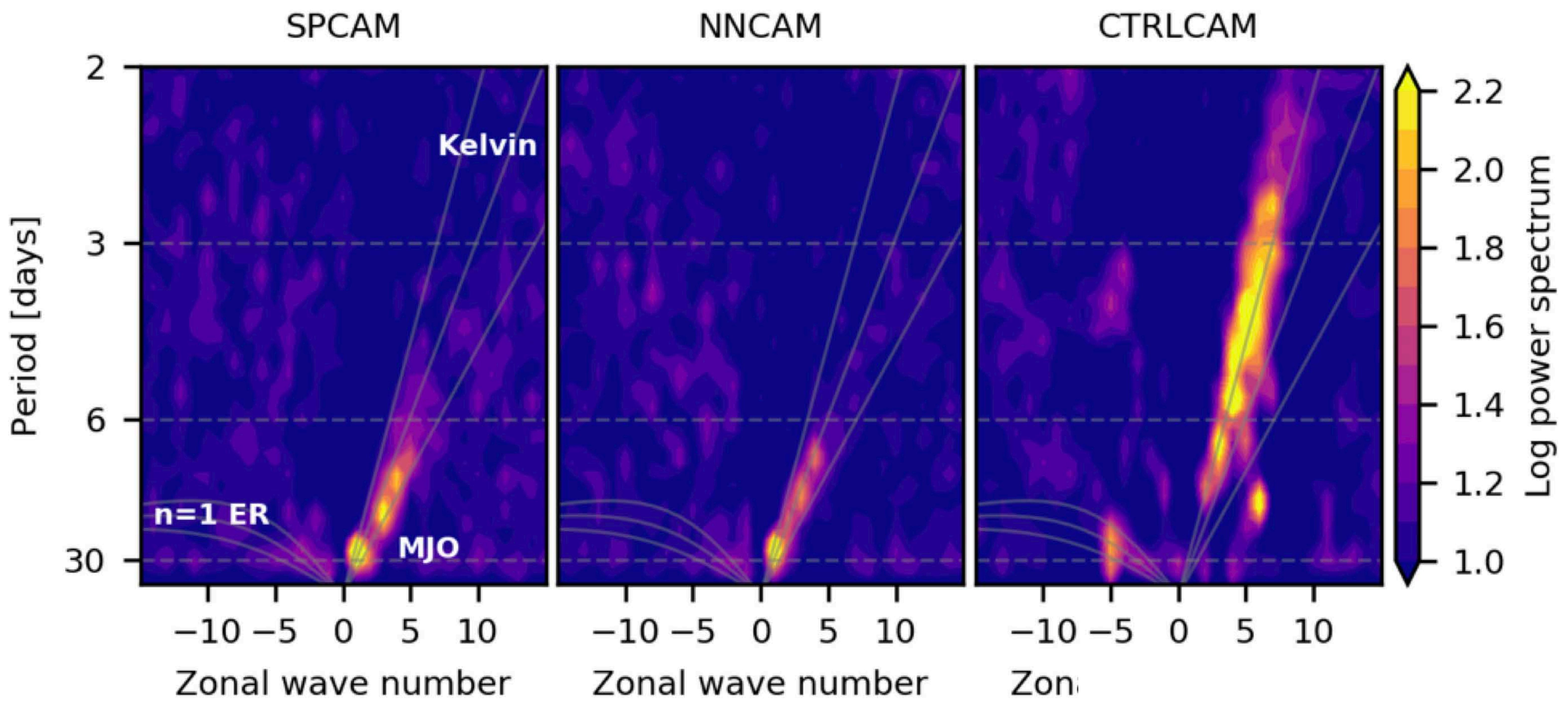
Beucler, Pritchard et al., *PRL,* 2020.

2021: Physical renormalizations that shield from extrapolation error developed.

Beucler, Pritchard et al: arXiv:2112.08440

Shiny prognostic results do exist for some skillful fits.

Equatorial wave spectrum produced by a NN-emulator of superparameterization on an aquaplanet GCM



Rasp, Pritchard and Gentine, *PNAS*, 2018.

But ML parameterization is also finicky and empirical (limits of operational potential uncertain)

Example of one of our neural network powered climate model blowing up in prognostic mode.

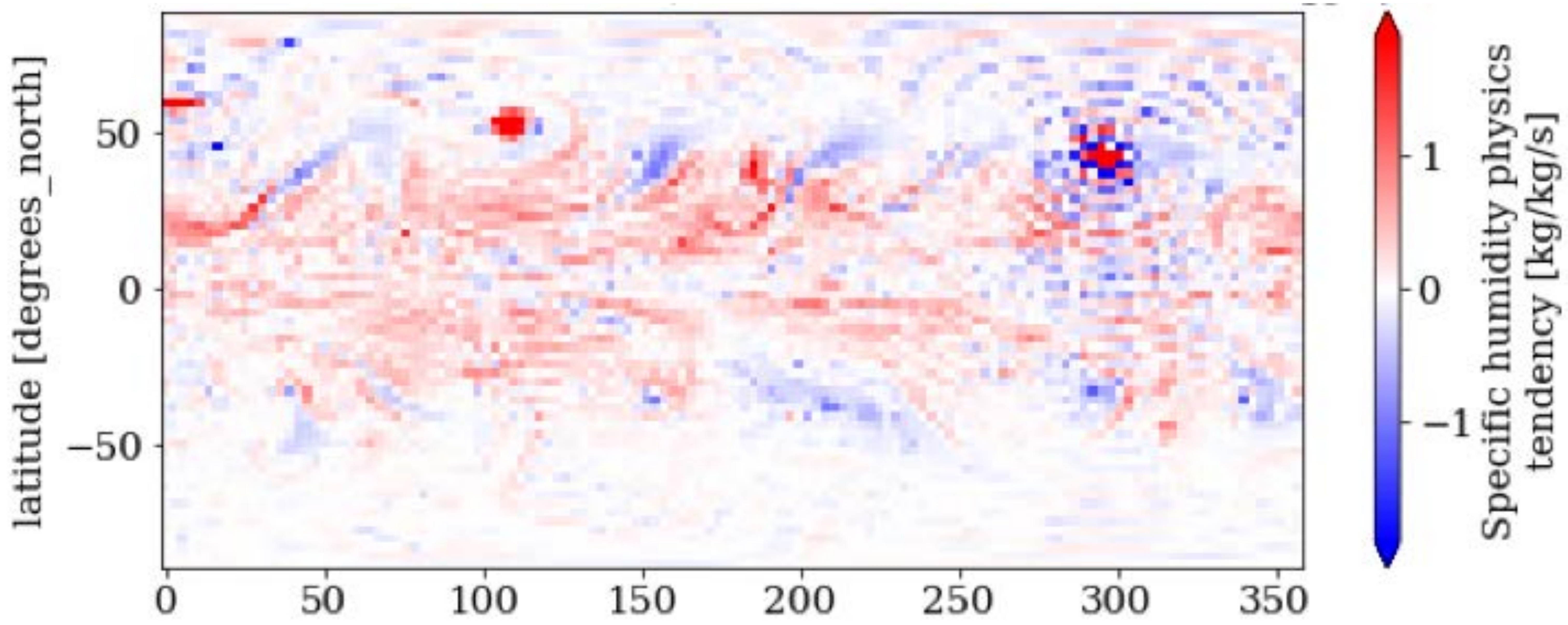


Figure courtesy of Tom Beucler.

OUTLOOK

MMF provides a perfect ML testbed. Many lessons emerging, rapid progress.

The LEAP Science & Technology Center is creating community MMF benchmarks to engage data scientists in competitions to solve reliability.

Maintaining SP within CESM will be vital to these efforts

If successful, a way to truly revolutionize & democratize high-resolution physics in CESM!

TAKE-HOME POINTS

- The MMF capability of CESM remains interesting for a variety of use cases.
- Regionalized MMF: High-res physics at reduced costs, faster throughout.
- High-res MMF: Explicit low cloud physics at increasingly ambitious computational scales.
- Art of tuning MMFs is maturing – fixing rainfall biases & low cloud amount.
- Machine learning parameterizations trained on MMF could upend cost trade-offs of high-resolution physics.