Basic Science and Actionable Science in Earth System Models: Current Challenges

Leo Donner
NOAA/GFDL, Princeton University

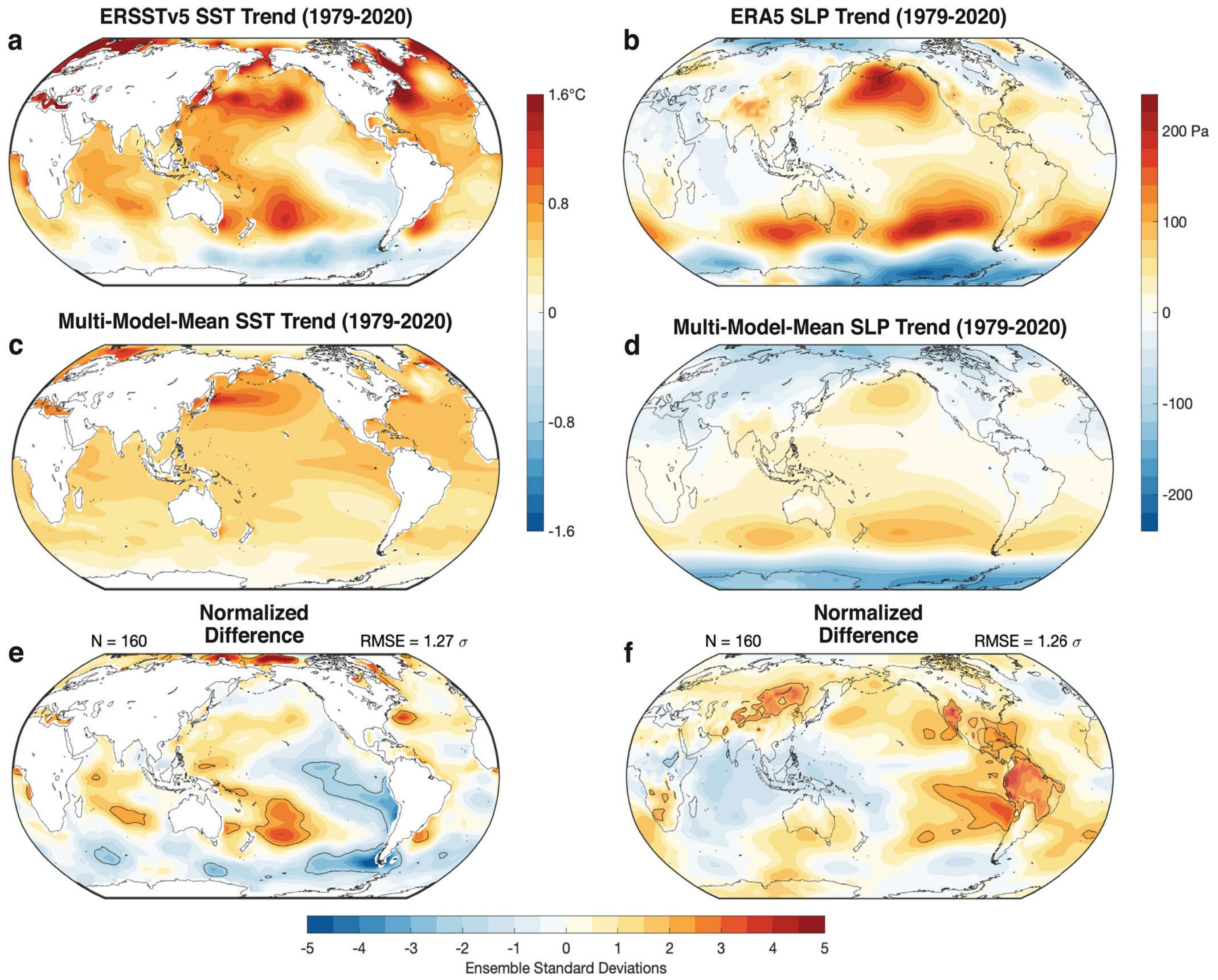
Spring DOE BERAC
April 11, 2024

Three Key Challenges

- Beyond Climate Sensitivity
- Going “Out of Sample” in Climate Projections
- “Fit for Purpose” Physics in Climate Projections

Beyond Climate Sensitivity

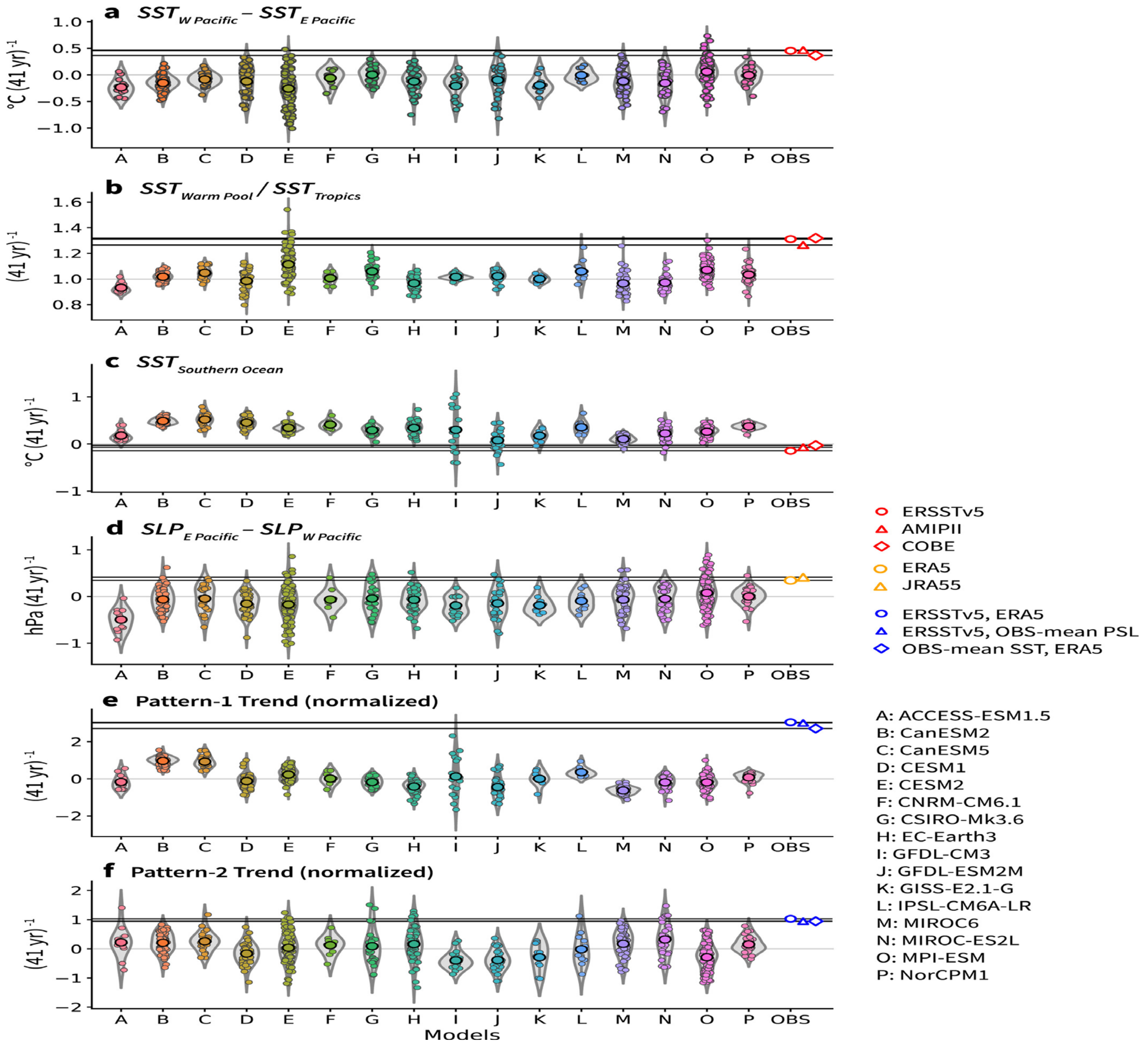
Systematic Climate Model Biases in the Large-Scale Patterns of Recent Sea-Surface Temperature and Sea-Level Pressure Change



Geophysical Research Letters, Volume: 49, Issue: 17, First published: 31 August 2022, DOI: (10.1029/2022GL100011)

Willis *et al.* (2022, *Geophys . Res. Lett.*)

Systematic Climate Model Biases in the Large-Scale Patterns of Recent Sea-Surface Temperature and Sea-Level Pressure Change



Geophysical Research Letters, Volume: 49, Issue: 17, First published: 31 August 2022, DOI: (10.1029/2022GL100011)

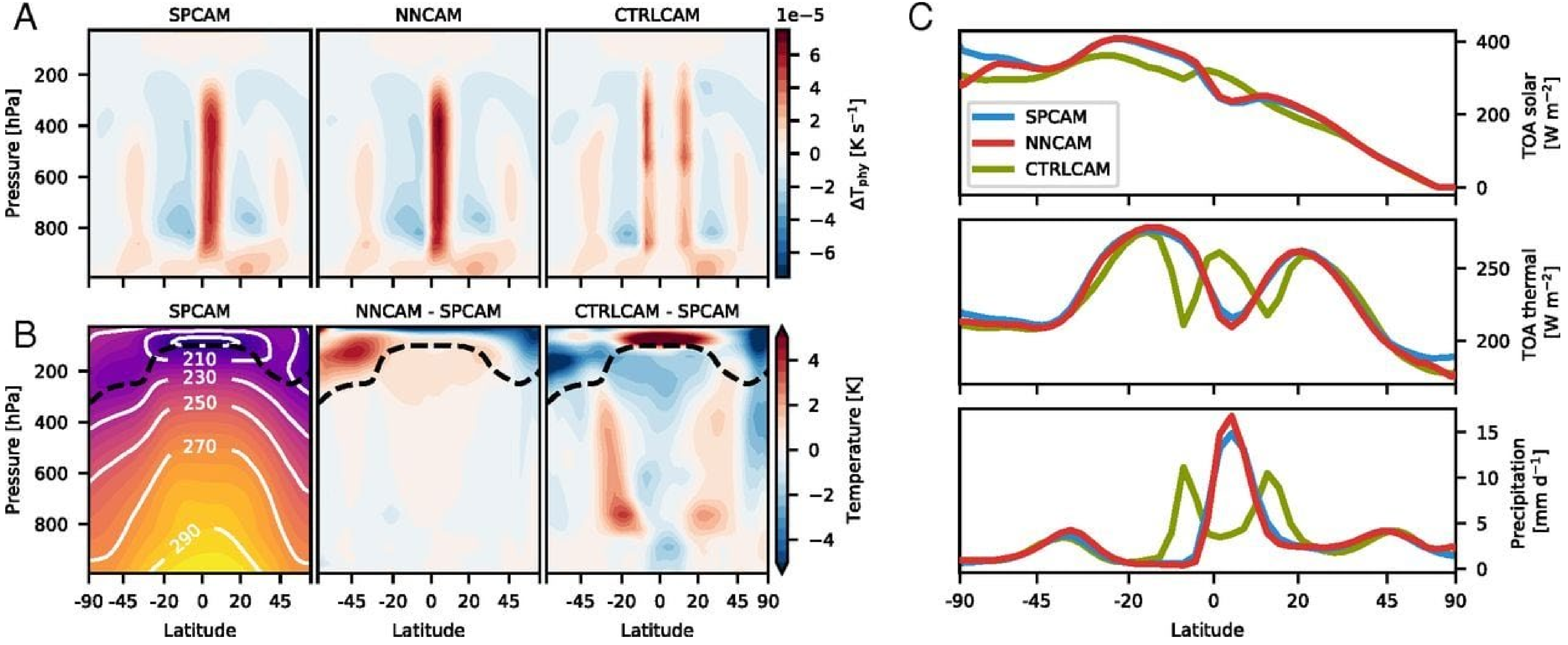
Beyond Climate Sensitivity

- Actionable science requires confident scenarios for high-impact events.
- For a given level of global-mean warming, high-impact events vary widely depending on SST pattern changes.
- Current Earth system models struggle to simulate SST pattern changes associated with historical warming.
- Climate sensitivity itself is closely linked to SST pattern changes!
- Earth system model development requires focus on SST pattern simulation. **Basic Science <-> Actionable Science!**

Going “Out of Sample” in Climate Projections

Neural nets can capture physical processes very well within training domain, both for mean climate:

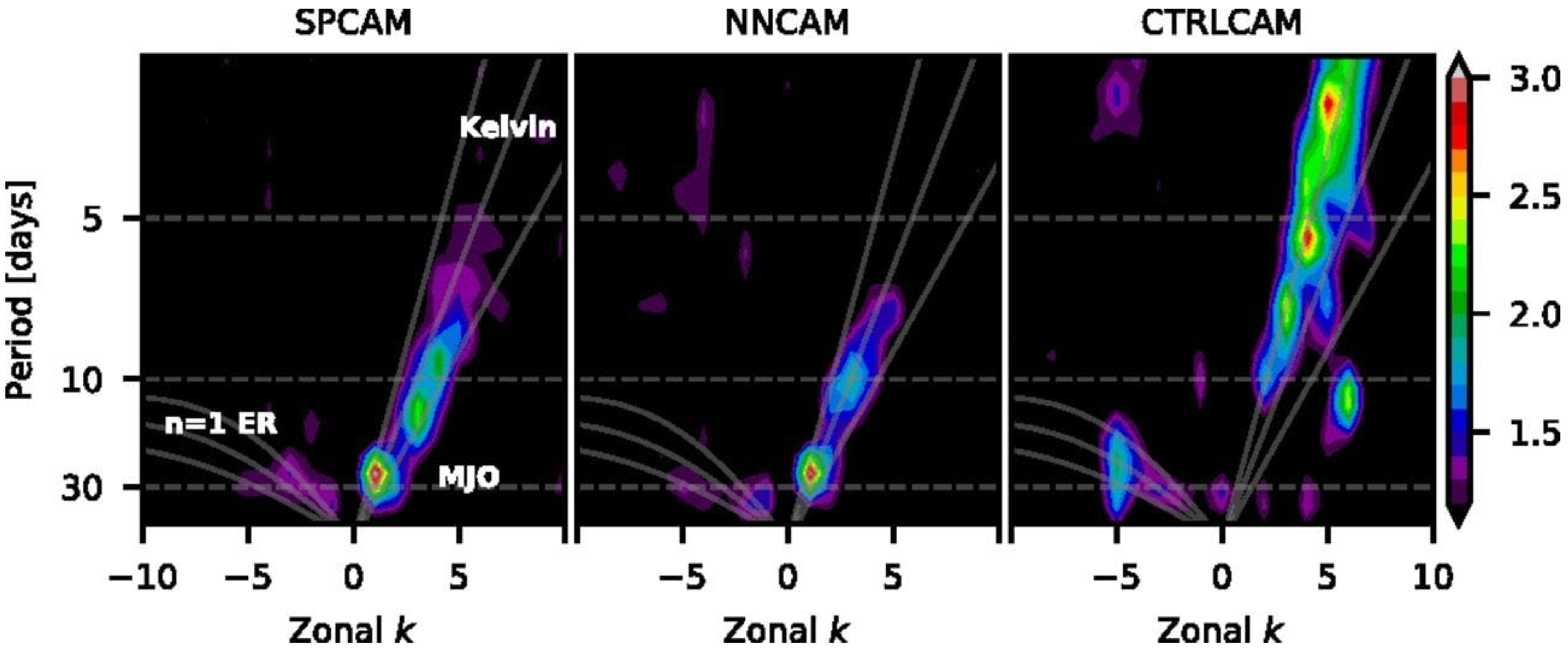
Convective + radiative heating



T differences

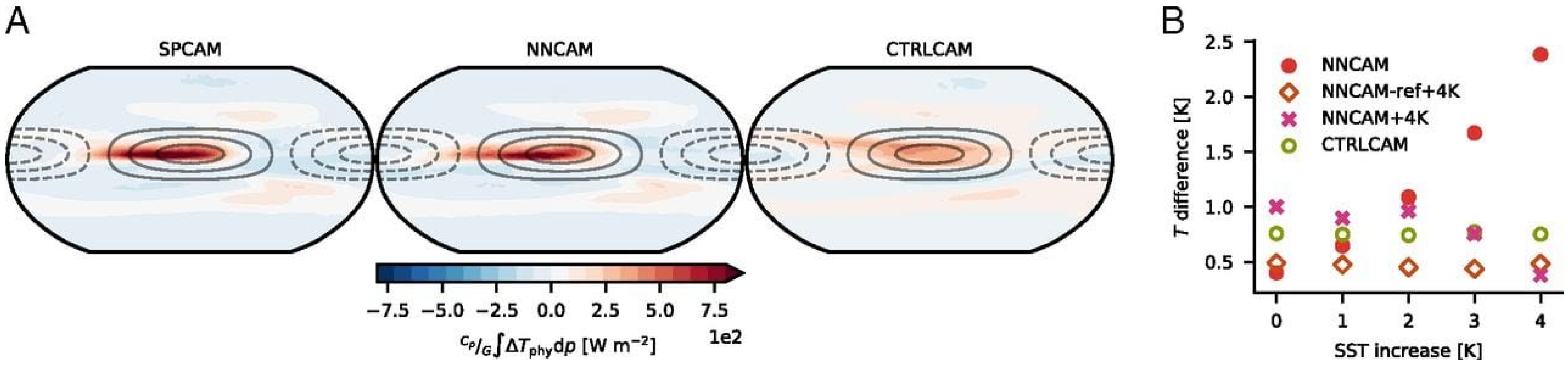
From Rasp et al. (2018, *PNAS*)

and variability important for sub-seasonal to seasonal prediction:



From Rasp et al. (2018, *PNAS*)

Neural nets can break down when applied out-of-sample.



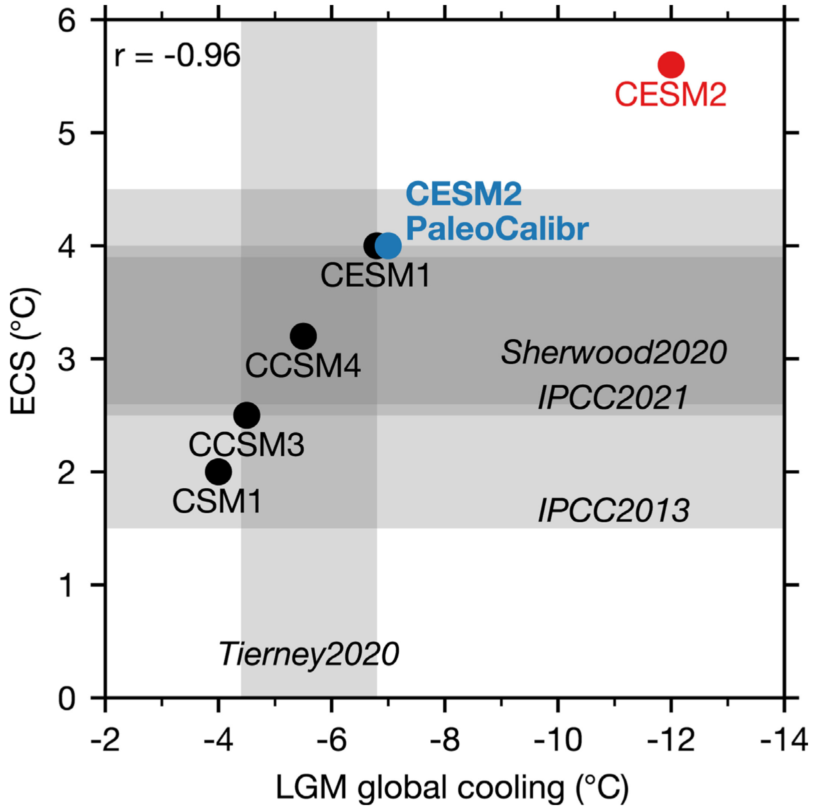
Neural net trained on zonally symmetric surface temperature able to simulate simulate heating on out-of-sample zonal wavenumber 1 asymmetry, 3-K amplitude.

from Rasp et al. (2018, *PNAS*)

Neural net (NNCAM) unable to simulate temperatures for global warming out of training domain.

Out-of-sample problem also applies to models developed conventionally, not using ML. CESM2, tuned and structured for pre-industrial and present-day climate, is not able to simulate the climate of the Last Glacial Maximum.

LGM Paleoclimate Constraints Inform Cloud Parameterizations and Equilibrium Climate Sensitivity in CESM2

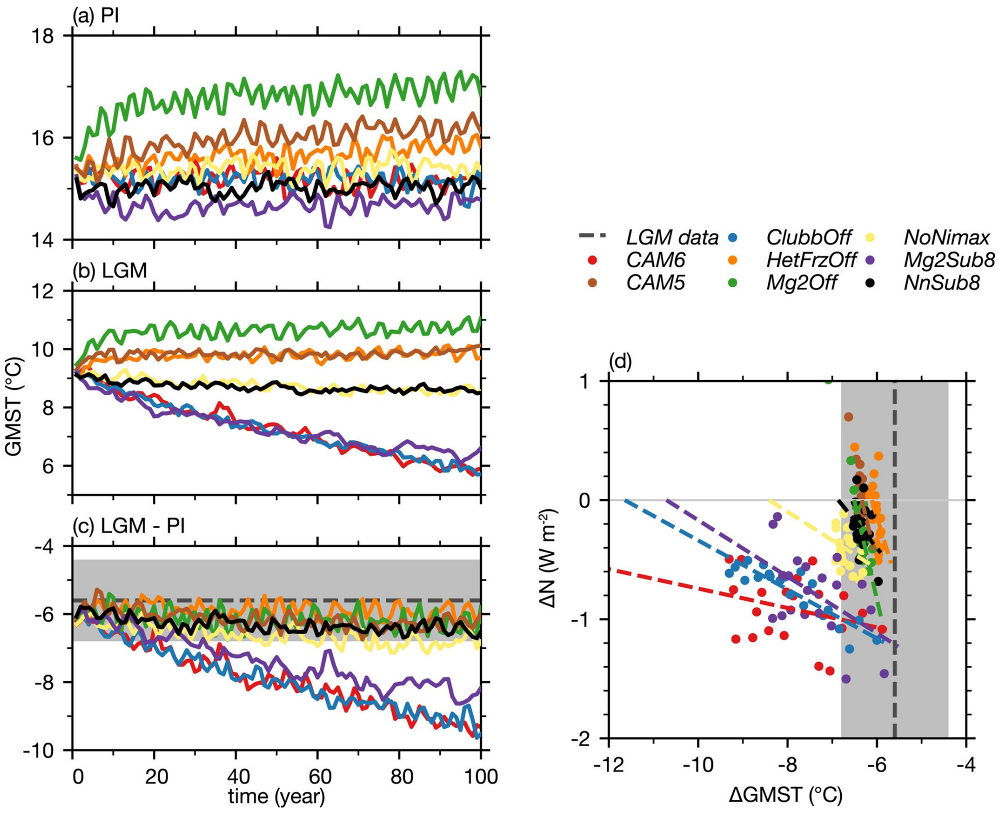


J Adv Model Earth Syst, Volume: 14, Issue: 4, First published: 25 February 2022, DOI: (10.1029/2021MS002776)

from Zhu et al. (2022, *JAMES*)

CESM2 LGM failure can be traced to microphysics (heterogeneous freezing, limiter on ice number concentration)

LGM Paleoclimate Constraints Inform Cloud Parameterizations and Equilibrium Climate Sensitivity in CESM2



J Adv Model Earth Syst, Volume: 14, Issue: 4, First published: 25 February 2022, DOI: (10.1029/2021MS002776)

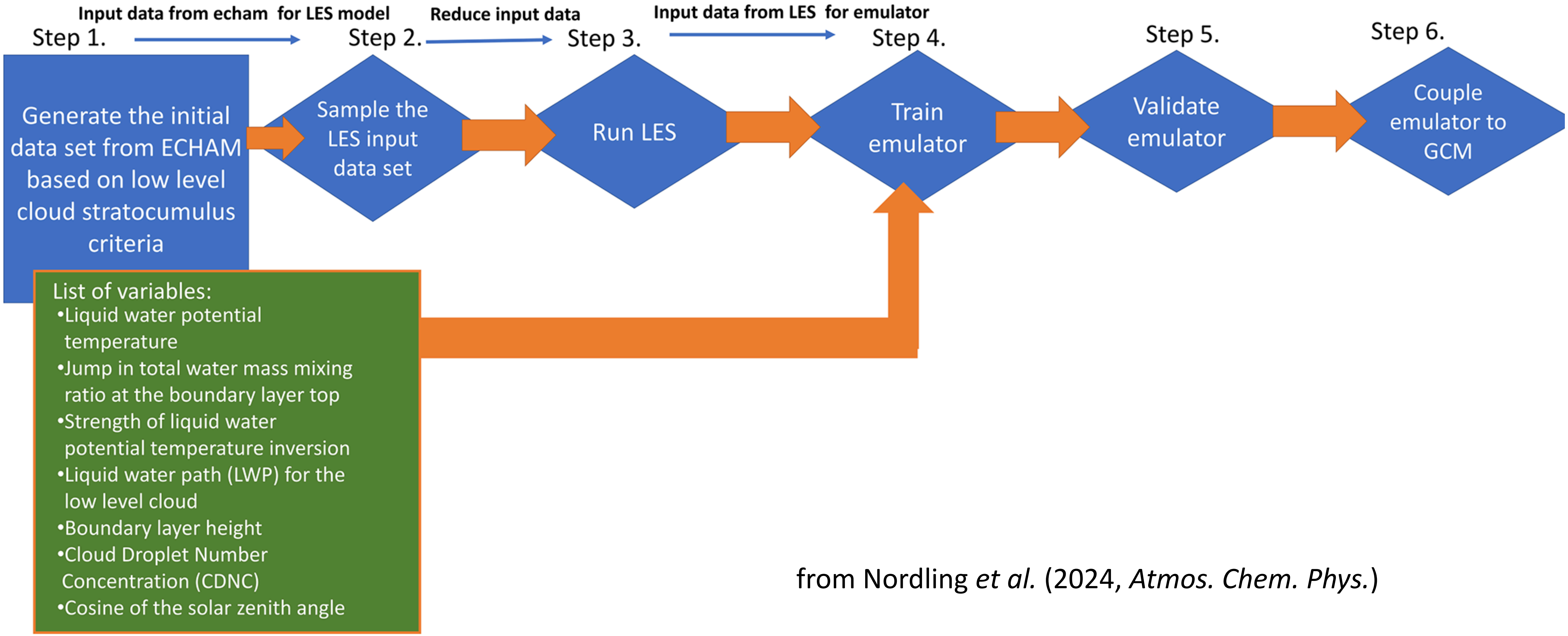
from Zhu et al. (2022, *JAMES*)

Going "out of sample" in climate projections

Robust physics, machine-learned from process models validated by field and lab observations, may provide a path forward. Physics must be constrained within training data, which can be scoped (perhaps iteratively) using current-generation Earth system models.

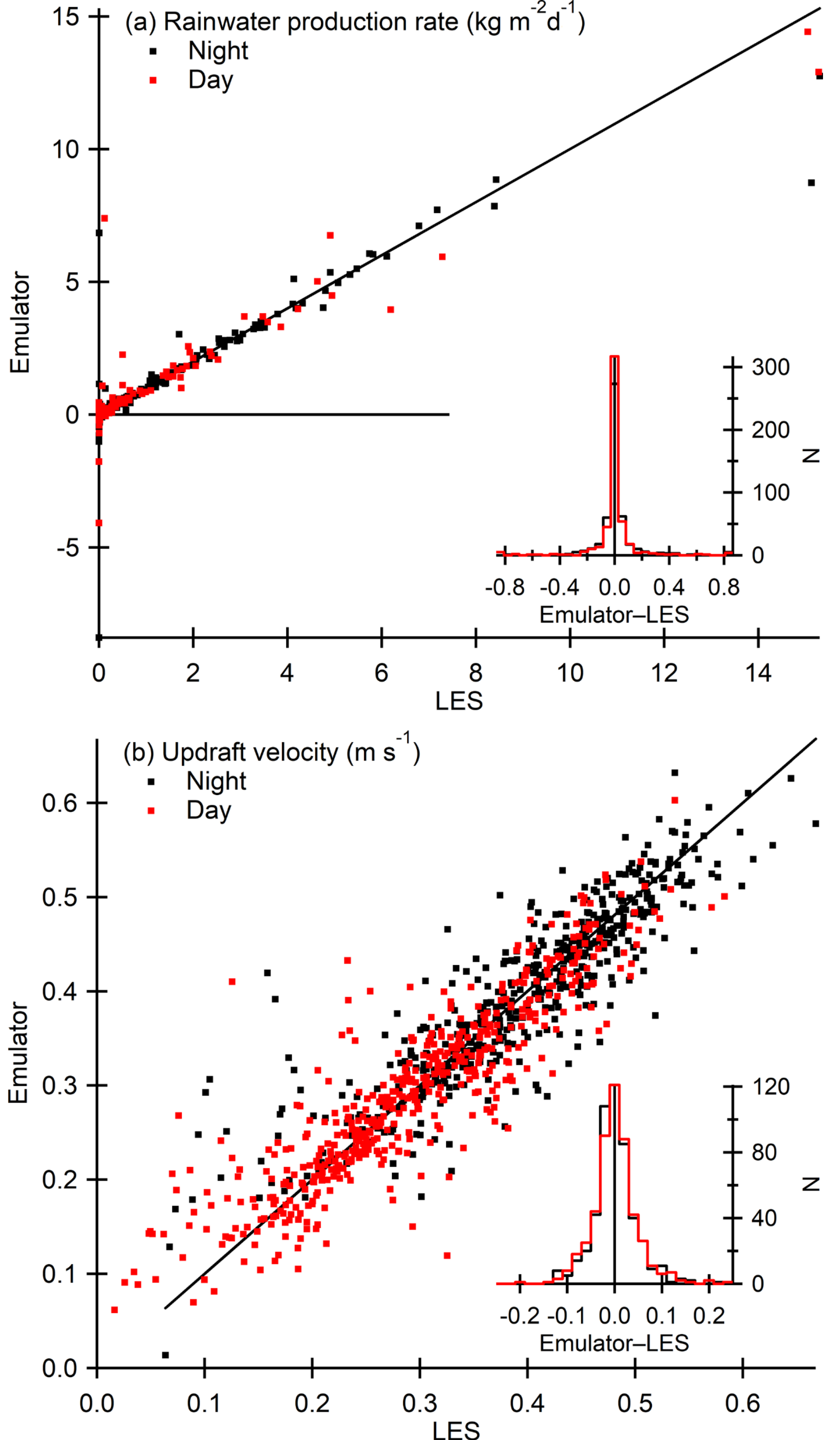
Proof of concept: Nordling *et al.* (2024, *Atmos. Chem. Phys.*)

ML over training space spanning all possible physics invocations in Earth system model



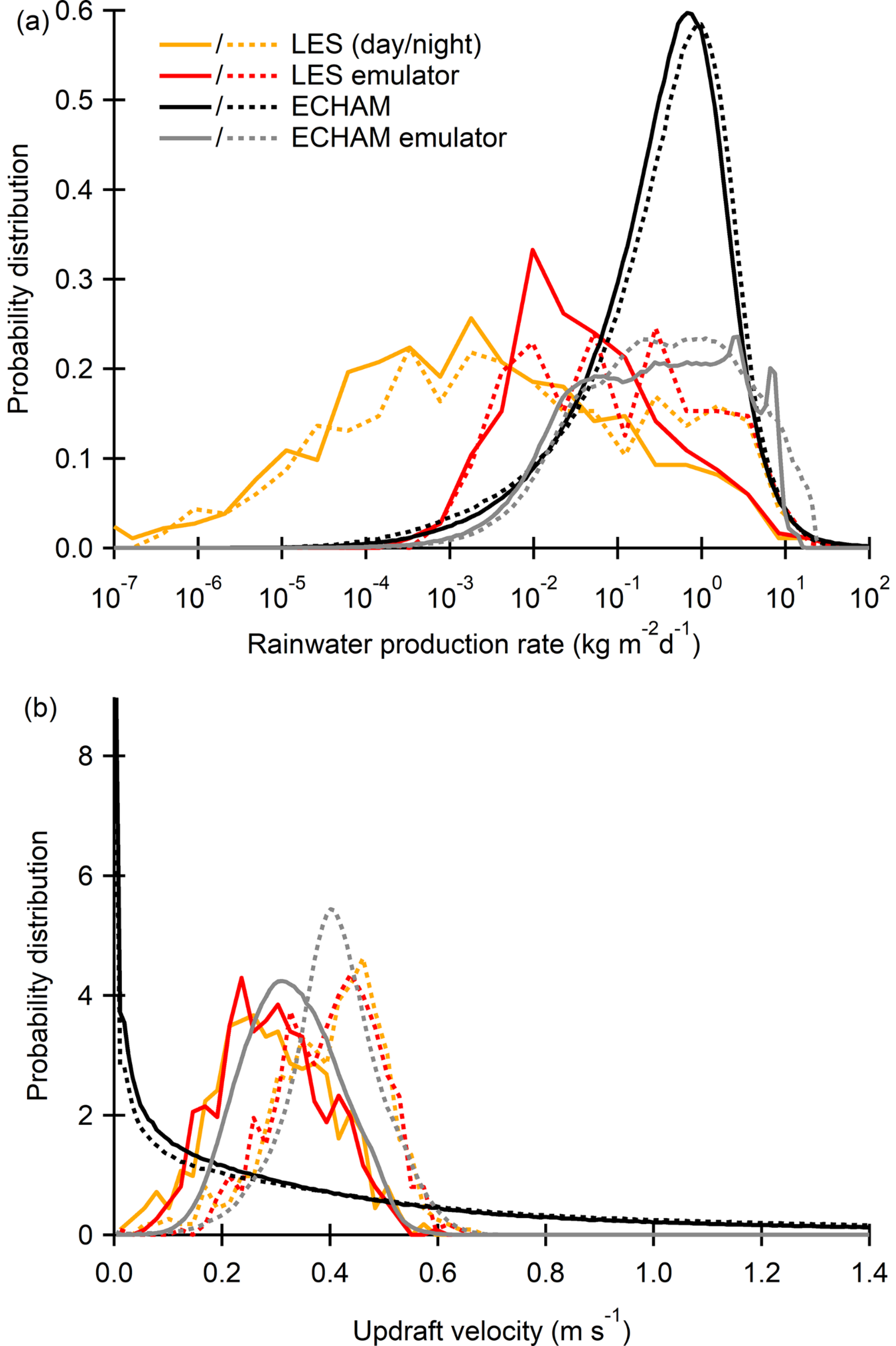
from Nordling *et al.* (2024, *Atmos. Chem. Phys.*)

Emulator successfully simulates properties of marine stratocumulus essential for realistic modeling of cloud-aerosol interactions-updraft velocities and rain production.



from Nordling *et al.* (2024, *Atmos. Chem. Phys.)*

When fully implemented in the ECHAM-M7 climate model, the ML emulators simulate LES updraft velocities very well. Rain production rates in the full model are less successfully simulated, even by LES.



from Nordling *et al.* (2024, *Atmos. Chem. Phys.)*

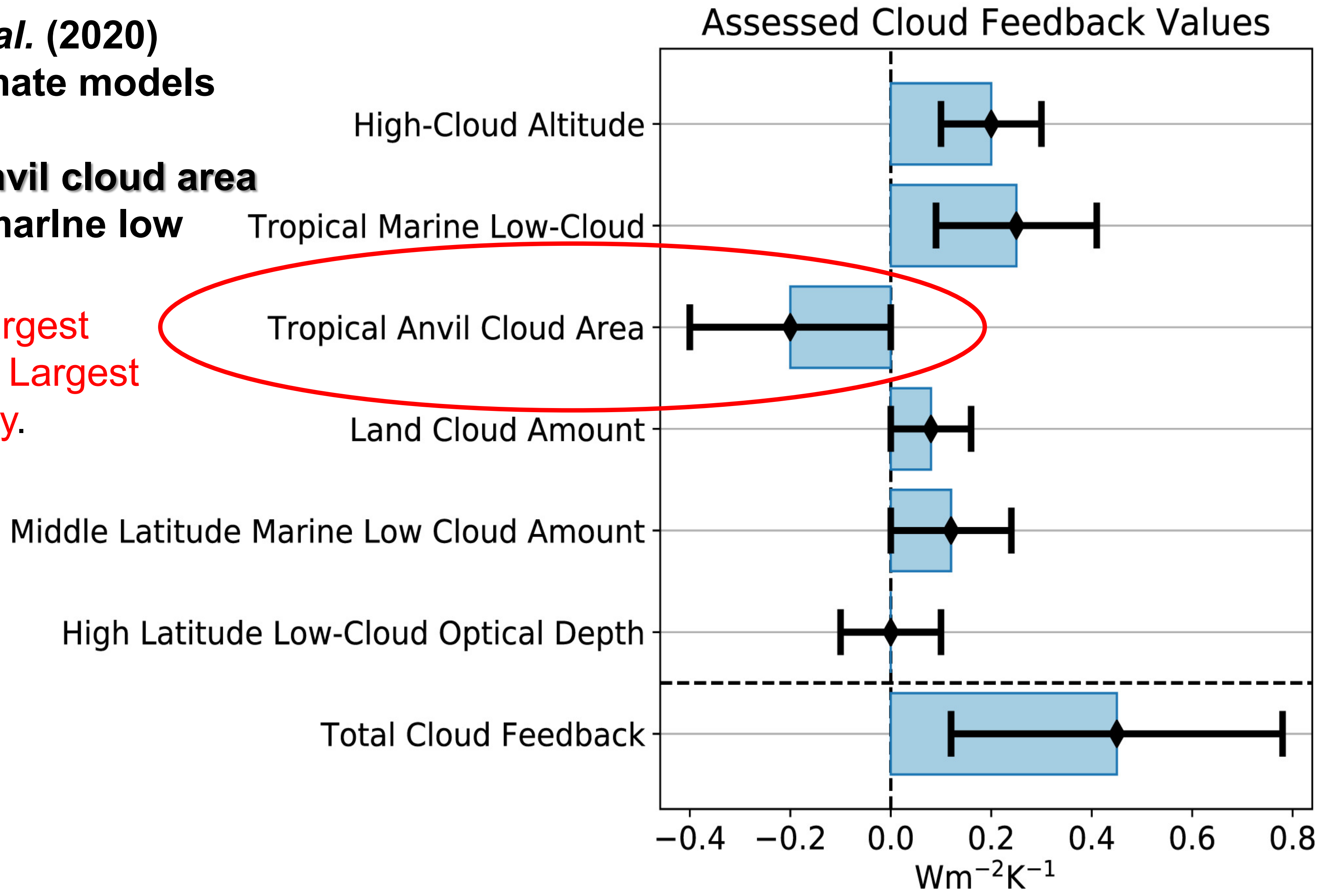
Going “Out of Sample” in Climate Projections

- Both convectional approaches to optimizing (“tuning”) Earth system models and ML methods can lead to serious “out of sample” problems when the models are applied outside of their “training data.”
- Weather versus Climate (cf., *Escape from Model Land,* Erica Thompson, Basic Books, 2022)
- Challenge models with observations for which they were not developed, e.g., paleoclimates for Earth system models.
- Models using emulators should be monitored carefully to detect straying out of training domains used in model development.
- Where robust physical basis is available, exploit in models, likely in conjunction with ML. **Basic Science <-> Actionable Science!**

“Fit for Purpose” Physics in Climate Simulations

Sherwood *et al.* (2020) assessed climate models *inadequate* for tropical anvil cloud area and tropical marlne low cloud.

Second largest feedback. Largest uncertainty.



Reviews of Geophysics, Volume: 58, Issue: 4, First published: 22 July 2020, DOI: (10.1029/2019RG000678)

from Sherwood et al. (2020, *Rev. Geophys*).

Tropical Anvil Cloud Area Feedback: N(-0.20;0.20) W m^{-2} K^{-1}

Lines of Evidence: *Observations*; Theory

- Microphysics changes with warming in deep convection (iris hypothesis) and increased convective aggregation with warming could reduce LW CRE by anvil clouds
- GCM convective microphysics and mesoscale parameterizations too simple/absent to capture feedback
- CRMs produced mixed results depending on domain size and microphysics
- CERES shows local negative feedbacks resulting from reduced LW trapping in slight excess of reduced SW reflection with warming. Scale local to global to obtain likelihood function. Large uncertainty because other observational studies with less accurate radiation measurements or older analysis periods yield neutral or slightly positive feedbacks.

from Sherwood et al. (2020, *Rev. Geophys.*)

Towards better understanding the tropical anvil cloud area feedback-

- Sherwood et al. (2020, *Rev. Geophys.*) lines of evidence for understanding climate sensitivity and cloud feedbacks are GCMs, observations, process-resolving models, and theory.
- Unlike the tropical marine low cloud feedback, which relies on process-resolving models as a line of evidence, the tropical anvil cloud area feedback relies only on observations and theory.
- Integrating observations with LES of unprecedented domain size and resolution to simulate both deep convection and anvils proposed as path forward.

Observed (Solid Black) & CRM Vertical Velocities (Varble *et al.*, 2014, *JGR*)

TWP-ICE
Experiment

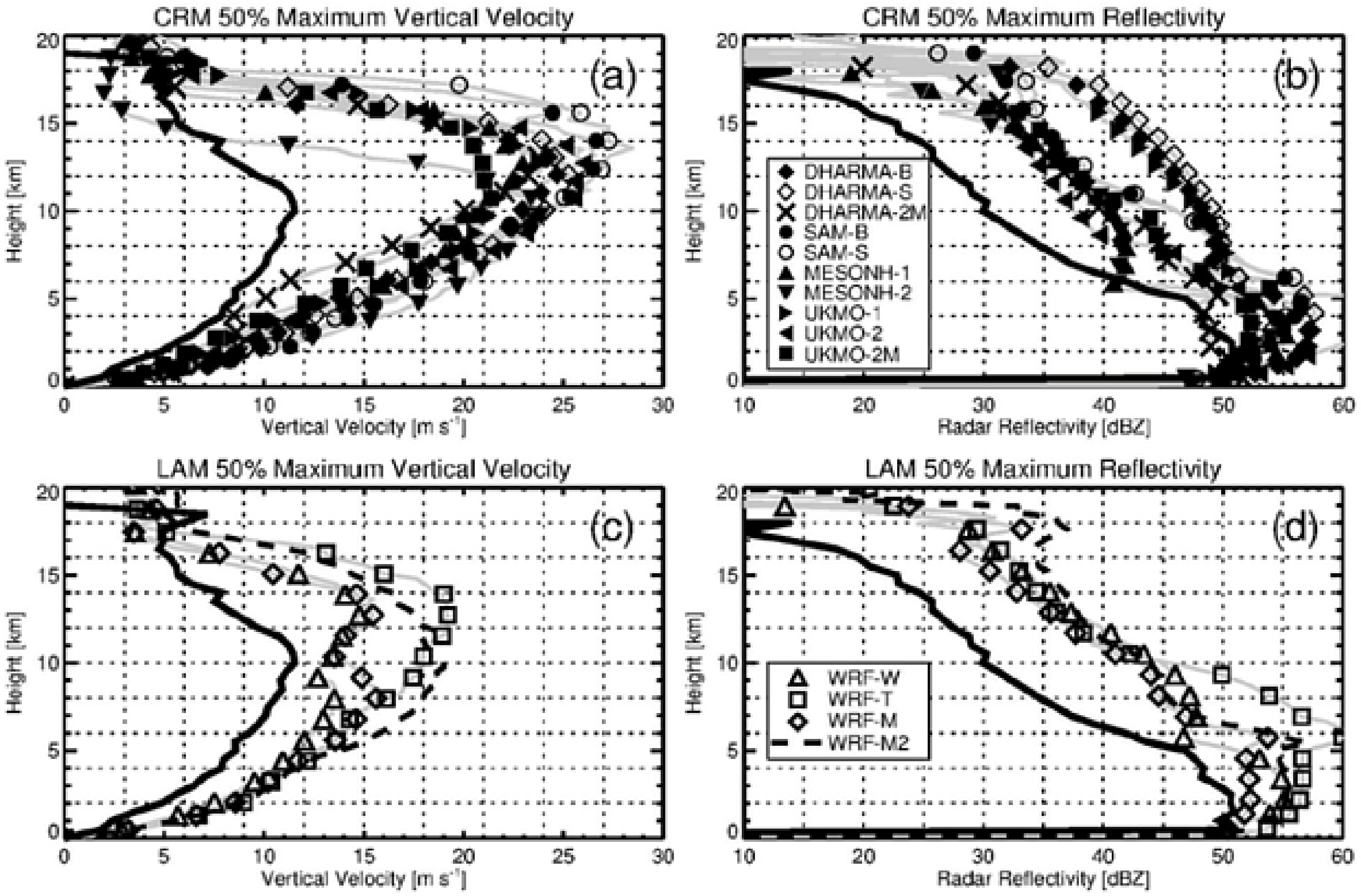
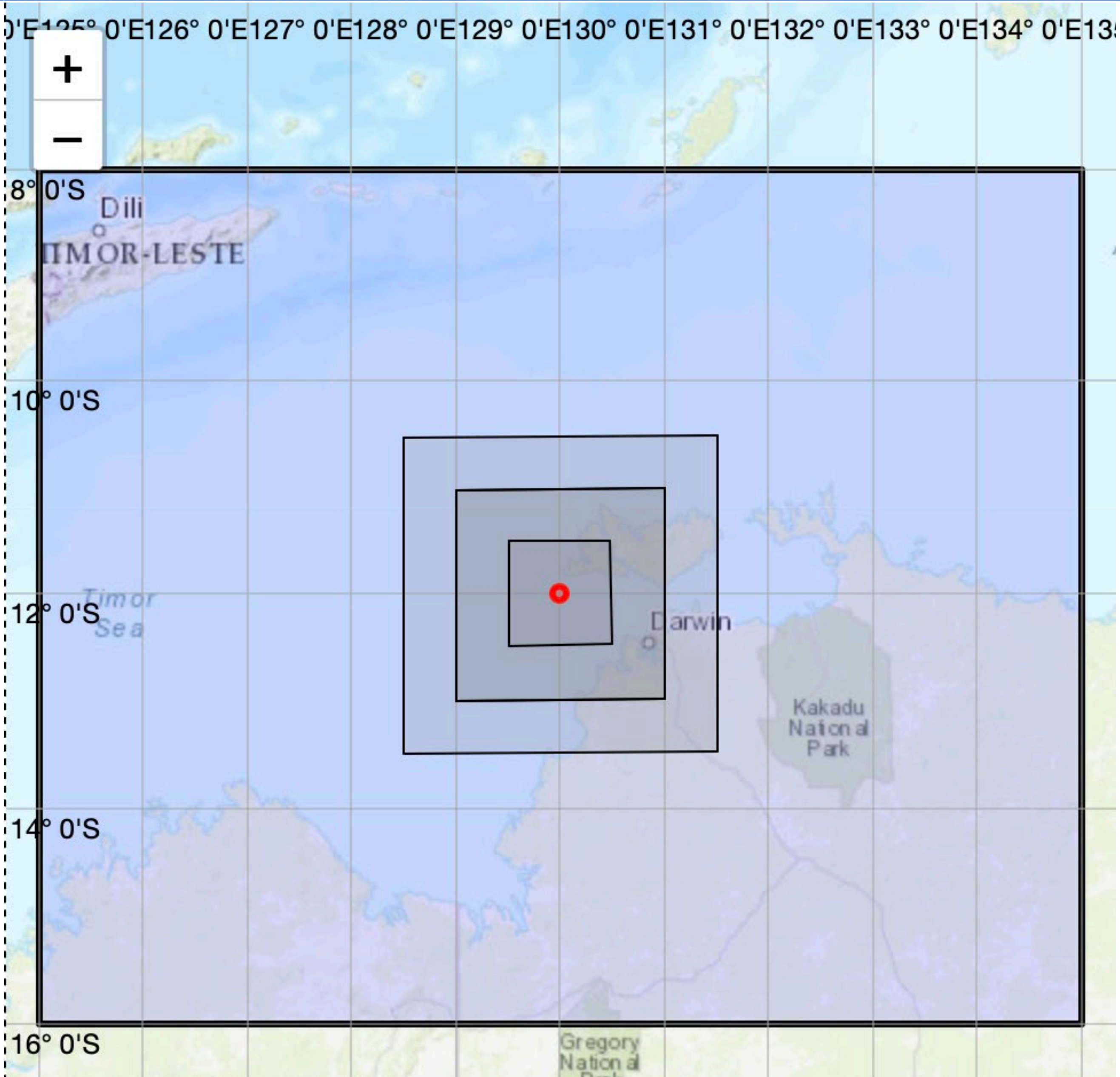


Figure 9: Median profiles of maximum vertical velocity (a,c) and radar reflectivity (b,d) for three-dimensionally defined convective updrafts beginning below 1 km and ending above 15 km for the period of 1310Z to 1750Z on 23 January 2006. CRM statistics are shown in (a-b) and LAM statistics are shown in (c-d). Gray lines with symbols and the dashed black lines represent simulations. Observations are represented by solid black lines.

Tropical Warm Pool – International Cloud Experiment (TWP-ICE) ICON-LAM simulation 625m-77m



dx: 625m/312m/156m

dz: 150m for 8-14km

dt: 8s

double-moment microphysics:
  Seifert and Beheng (2006)

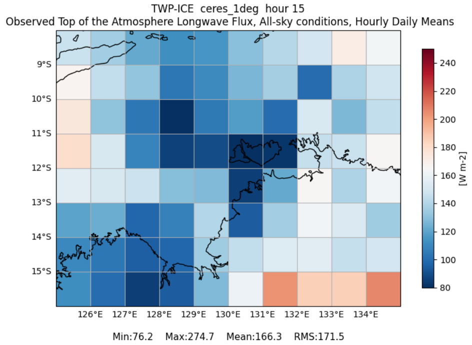
radiation: ecRad, dt=360s

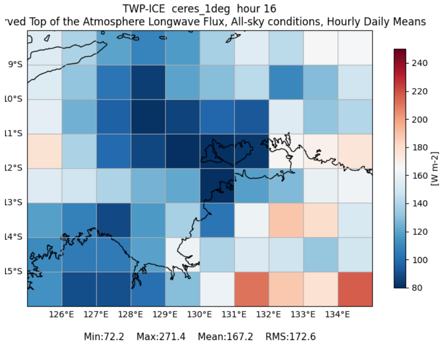
same setup with FV3 (later)

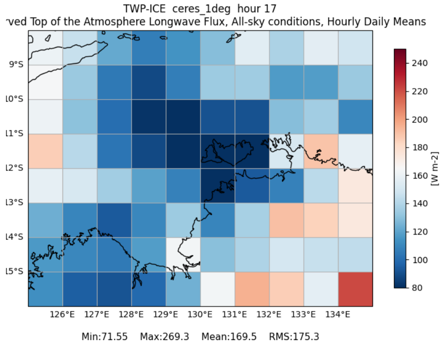
from Martin Köhler, DWD

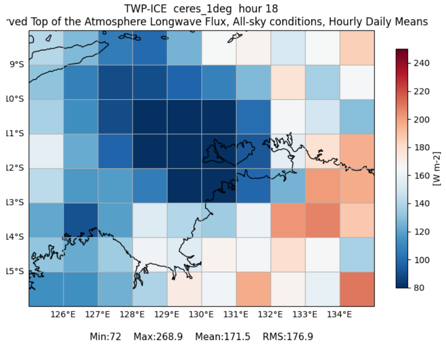
0-20S
120-140E
year 2006

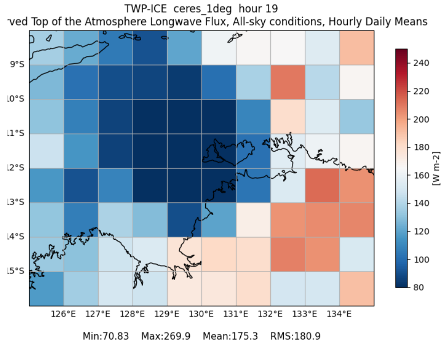
Tropical Warm Pool – International Cloud Experiment (TWP-ICE)
CERES SYN1deg TOA – LW (analysis by Martin Köhler, DWD)

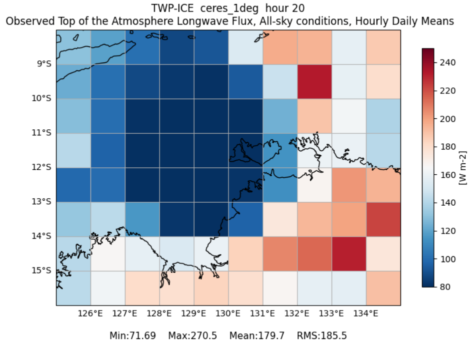
1/23 15Z


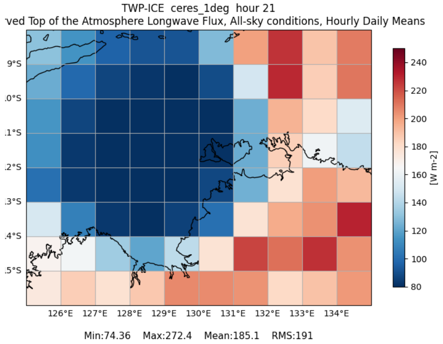
1/23 16Z


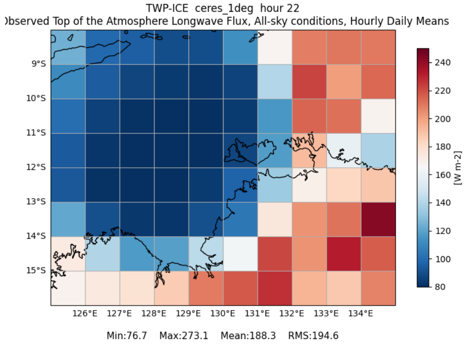
1/23 17Z


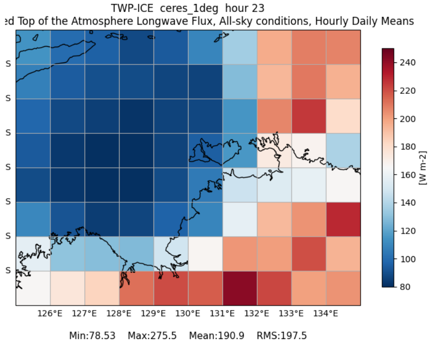
1/23 18Z


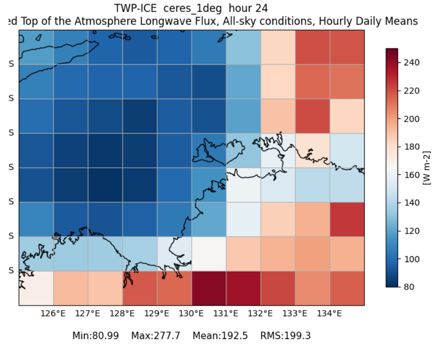
1/23 19Z


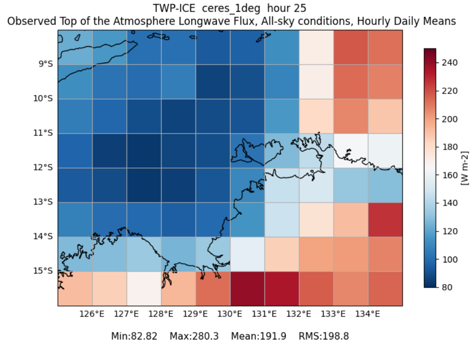
1/23 20Z


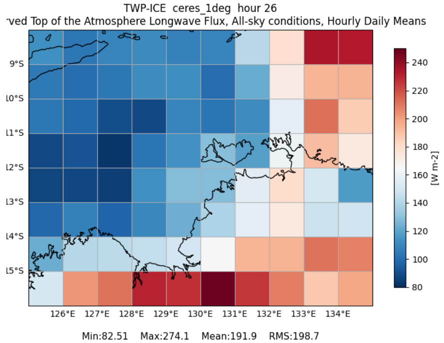
1/23 21Z


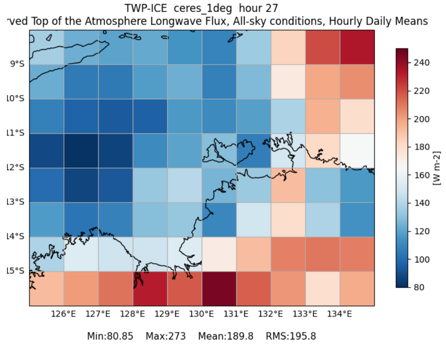
1/23 22Z


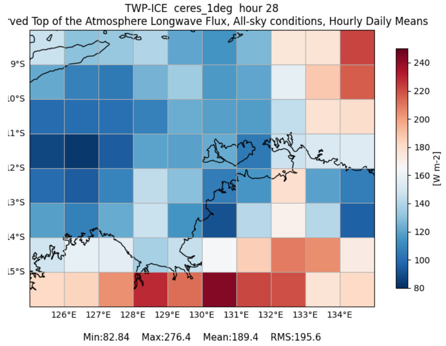
1/23 23Z


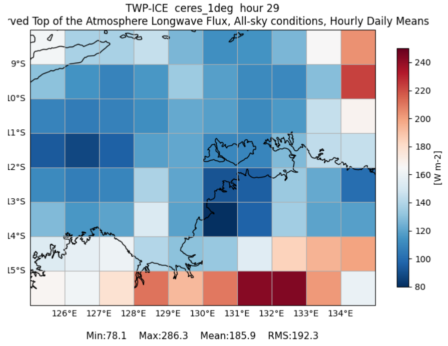
1/24 00Z


1/24 01Z


1/24 02Z


1/24 03Z


1/24 04Z


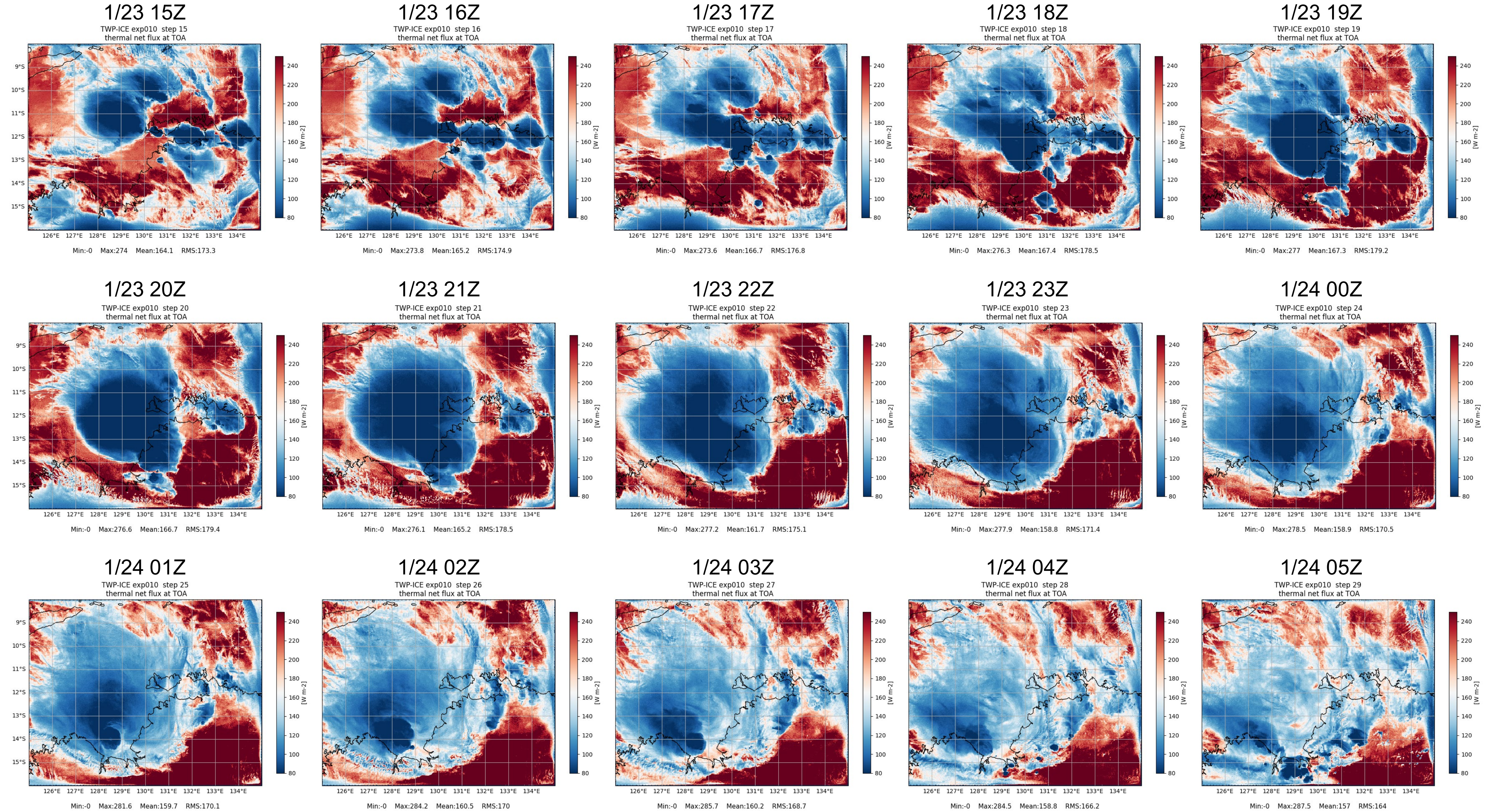
1/24 05Z


Tropical Warm Pool – International Cloud Experiment (TWP-ICE)

0-20S
120-140E
year 2006

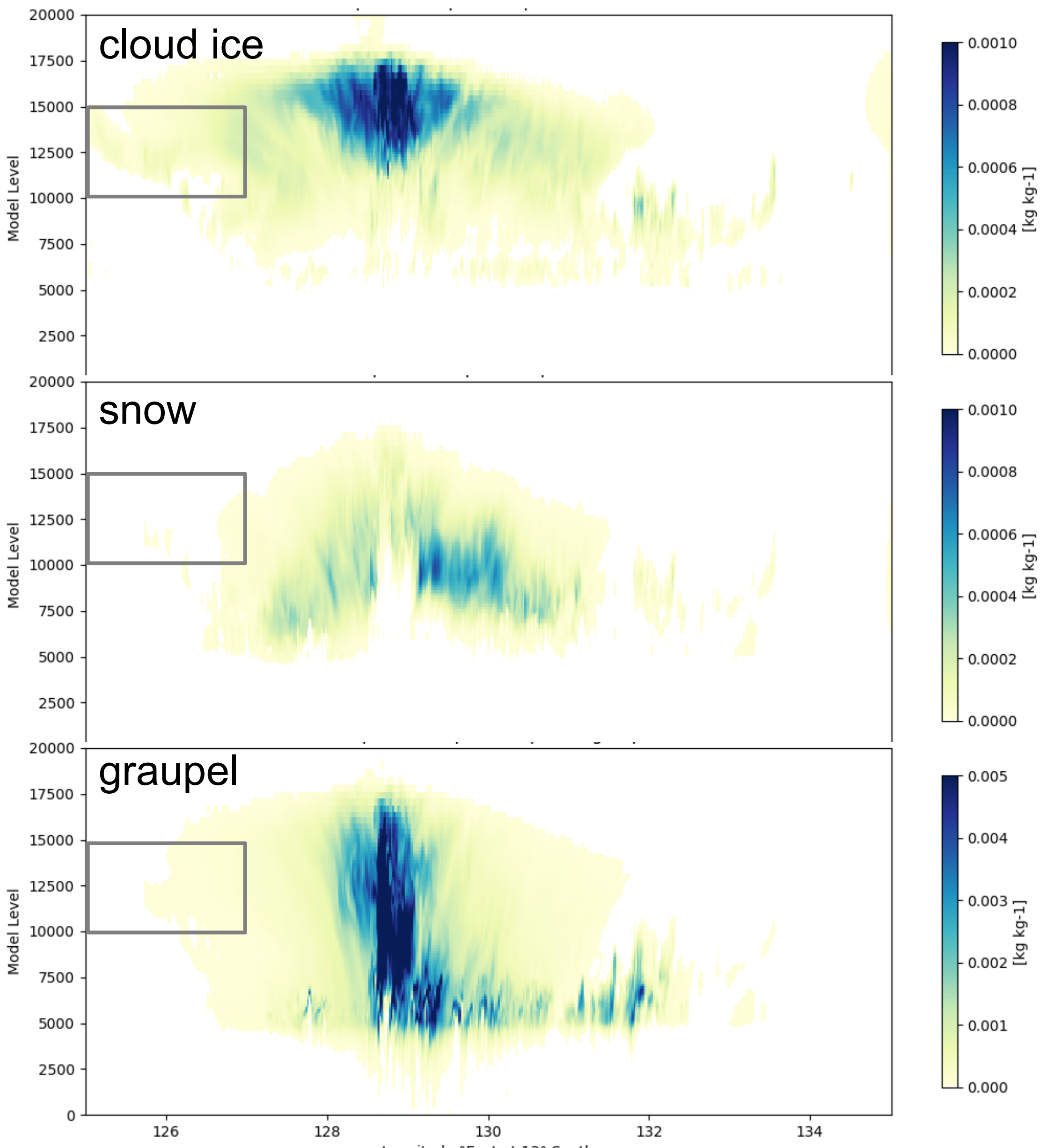
ICON 625m TOA – LW

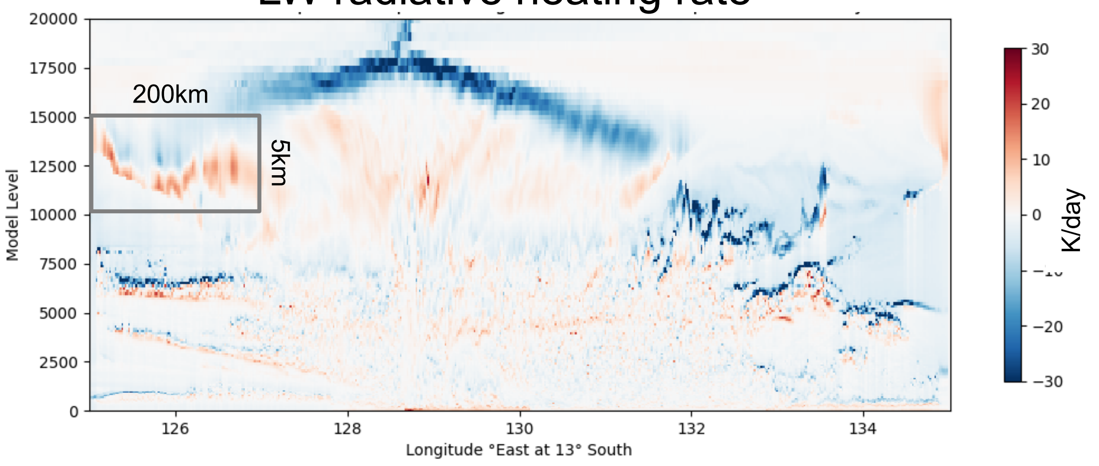
Integration by Martin Köhler, DWD

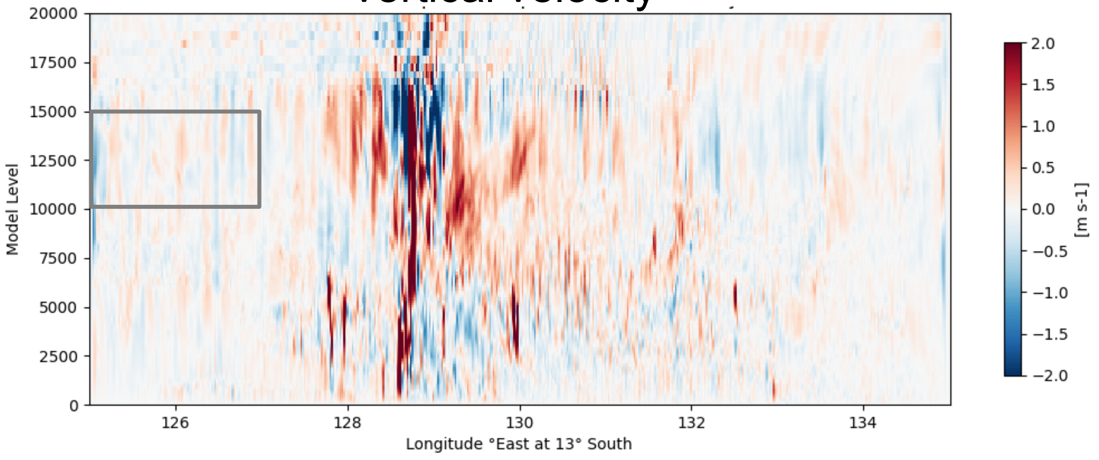


ICON TWP-ICE: 2006-01-23+24h (09LT) 13S

exp008







Integration by Martin Köhler, DWD

dx=615m

“Fit for Purpose” Physics in Climate Simulations

Imperative: Process-resolving models able to successfully simulate potentially large and important feedback phenomena, i.e., tropical convection with anvils. Extensive observational validation necessary.

Once such models are in place, they offer a feedback line of evidence and potentially a basis for parameterization in climate models. ML could be promising in doing so. **Basic Science <-> Actionable Science!**

What about kilometer-scale global models?

What about kilometer-scale models?

- Global, resolved models for mesoscale
- Largest convective updrafts approach kilometer scale, but many are smaller
- Many cloud systems (cirrus, stratiform) are turbulent at the scale of tens of meters
- Unprecedented spectrum of atmospheric motions explicitly resolved but subgrid turbulence and numerical treatments remain issues
- Cloud feedbacks and aerosol-cloud interactions require realistic updraft velocities
- Kilometer-scale models challenged by observational constraints on updrafts

SST Patterns, Increasing "Out of Sample" Accuracy and Confidence, "Fit for Purpose" Physics

- New model diagnostic strategies to discern reasons for poor performance of Earth system models in simulating SST pattern changes during global warming are required.
- "Out of sample" issues have been exposed by paleoclimate modeling and arise due to structural and optimization choices, not only in ML.
- Increasing realism of physical processes and unresolved processes is ultimate solution to "out of sample" and "fit for purpose" challenges.
- Robust observational base and process-level modeling are necessary, but challenge remains incorporating this knowledge in Earth system models. ML offers great promise, though balanced effort with domain knowledge and respecting limits of training data essential.
- **Basic Science <-> Actionable Science**