Probabilistic Machine Learning for Stochastic Parameterization of Deep Convection Triggering

Greta Miller [1], Hannah Christensen[1], Philip Stier [1], Judith Berner [2], Will Chapman [2]

[1] University of Oxford
[2] National Center for Atmospheric Research

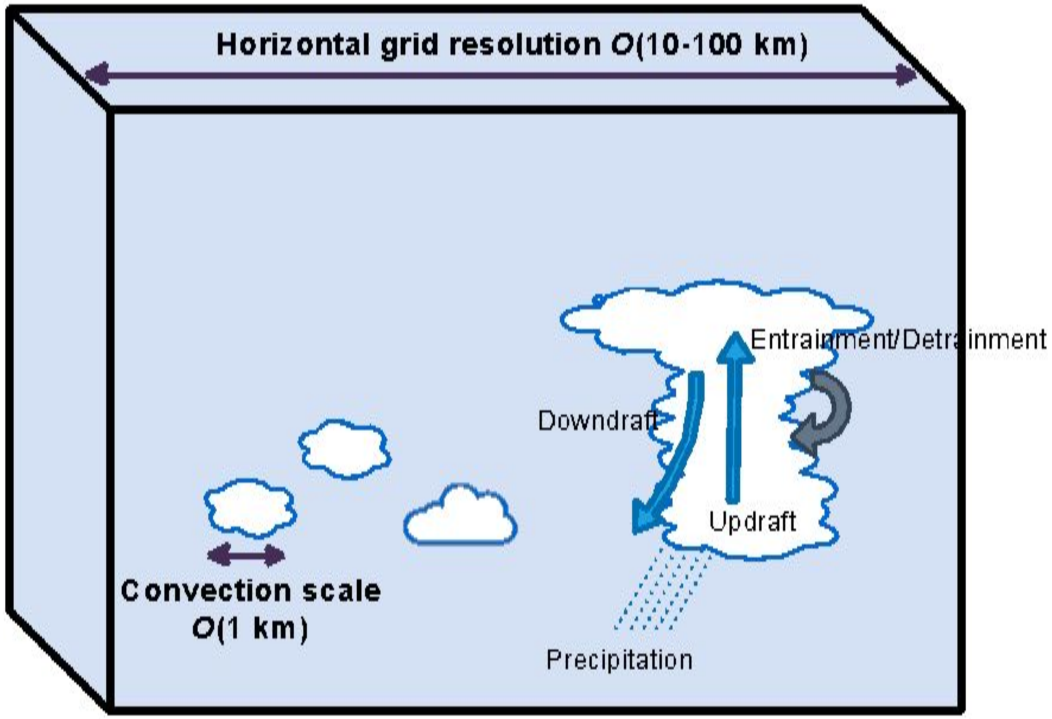
greta.miller@physics.ox.ac.uk

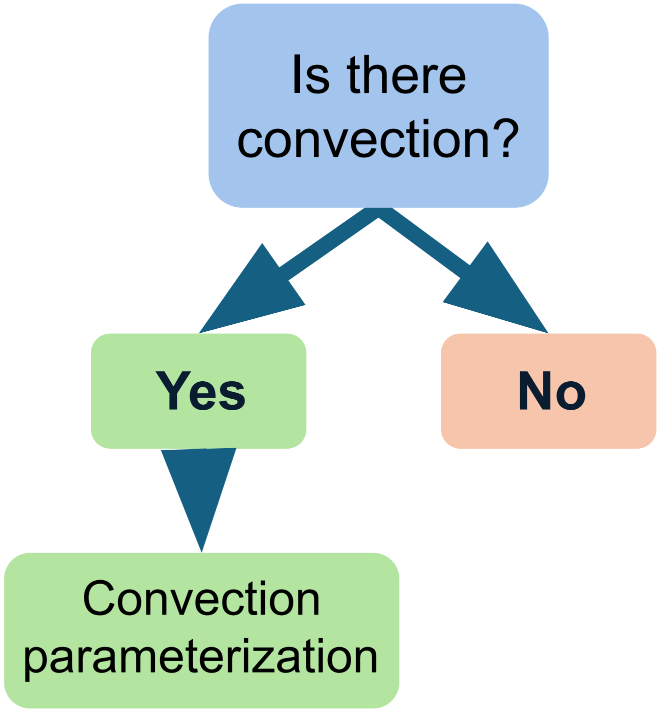
UNIVERSITY OF OXFORD

NCAR Graduate Visitor Program

M^2LInES

Convection trigger: where and when deep convection occurs in models





Parameterization of convection □ issues in model simulations

- Poor simulations of precipitation frequency and intensity (“drizzle problem”), diurnal cycle, ITCZ, etc.

- Most convection triggers in models are deterministic, without considering the uncertainty in the convective state as a stochastic process

Diurnal peak amplitude of

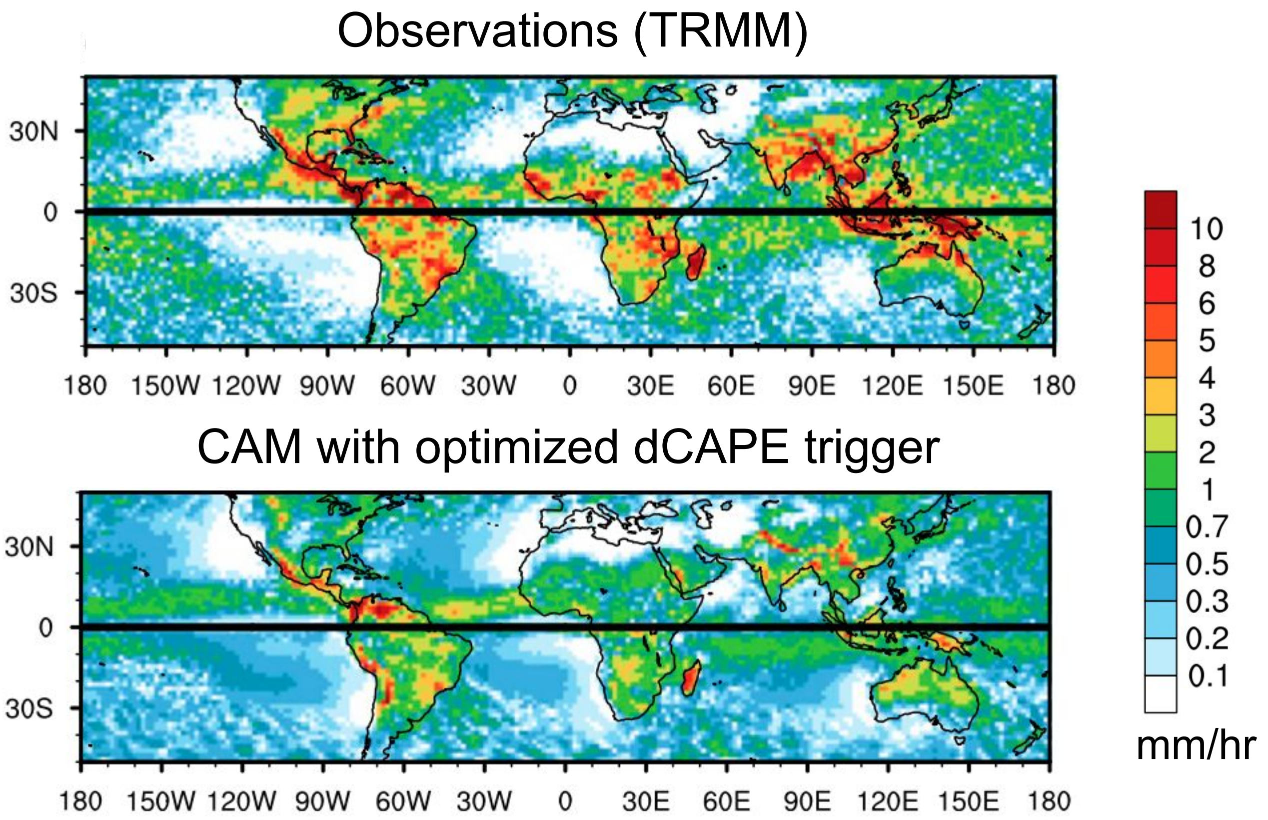


Figure from Cui et al. (2021)

How can we improve the representation of convection in climate models?

Probabilistic machine learning

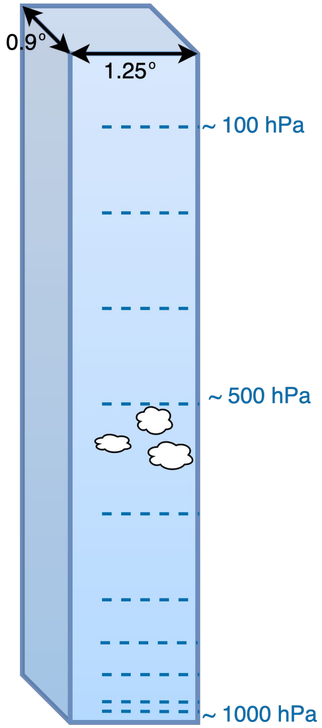
Stochastic parameterization

□ Develop a machine learning model that predicts the probability of deep convection occurrence

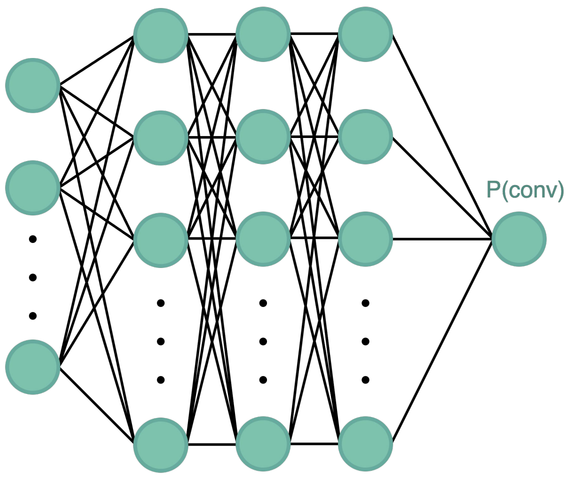
- Characterize sources of uncertainty for convection
- Implement as stochastic parameterization in operational climate model

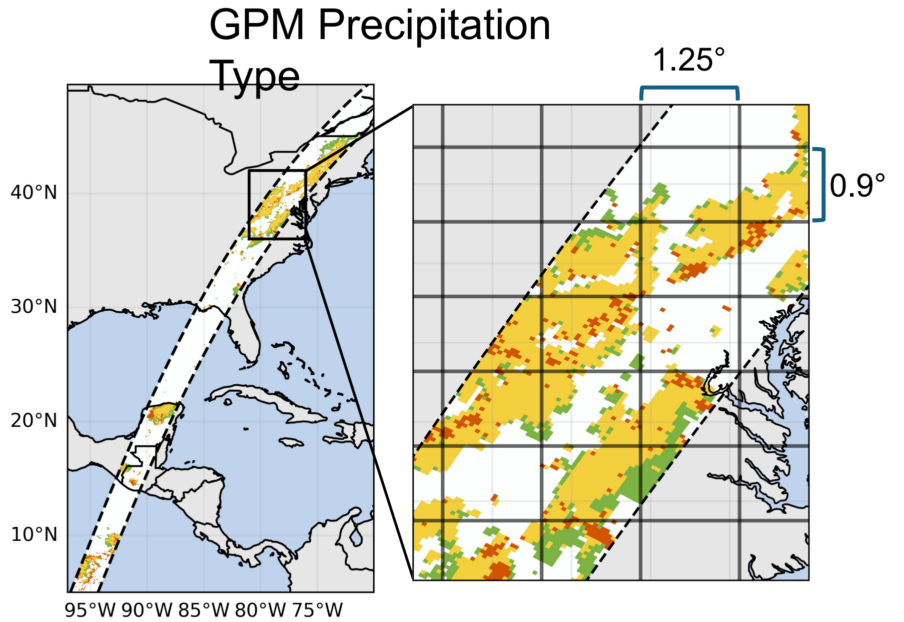
Methods

Inputs: Atmospheric state → **Neural Network** → **Output:** Probability of convection



Total column water vapor
Pressure (surface)
Temperature (levels)
Specific humidity (levels)
Vertical velocity (levels)





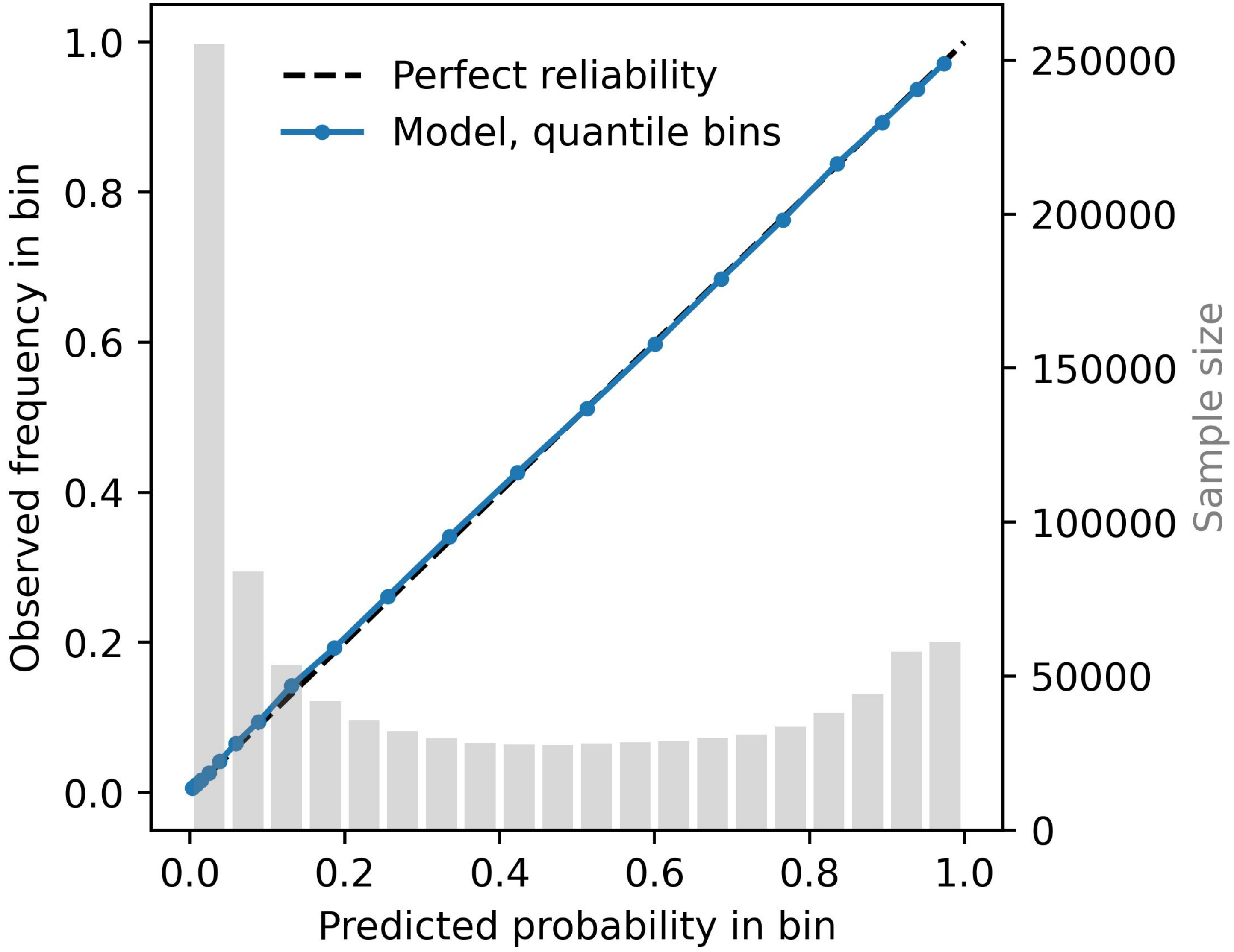
- ECMWF ERA5 reanalysis
- Hourly
- Re-gridded to CAM6 resolution:
  0.9° lat x 1.25° lon
  10/32 vertical model levels

- Train on 2018 and 2020, test on 2019 and 2021
- Loss function: binary cross-entropy

Convective
Stratiform
Other

- Convective precipitation from Global Precipitation Measurement (GPM) dual-frequency precipitation radar
- 65°S to 65°N

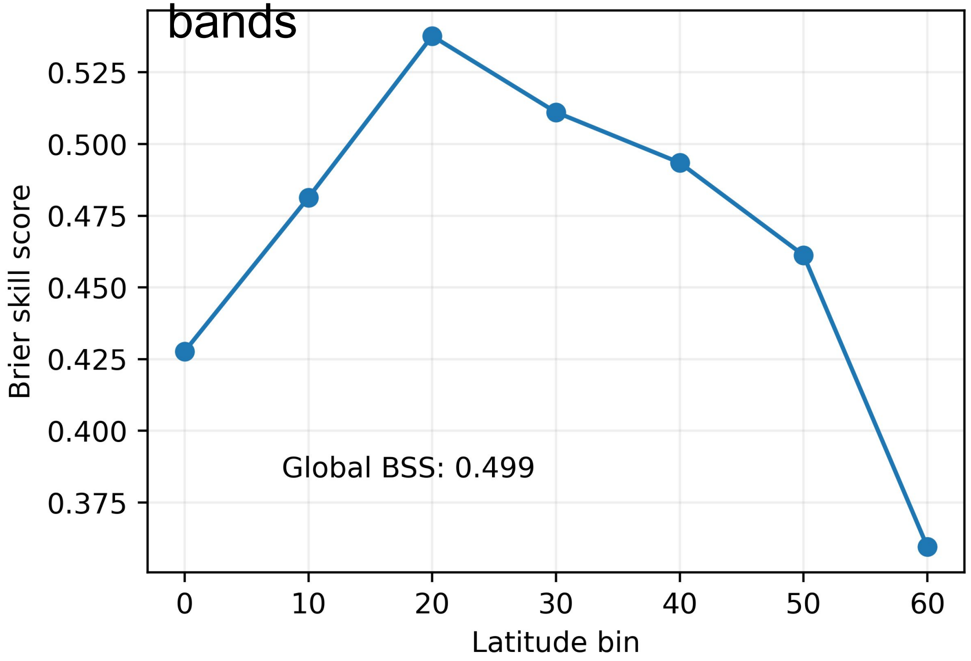
Model performance



Model performance

higher skill

BSS shows skill across latitude bands

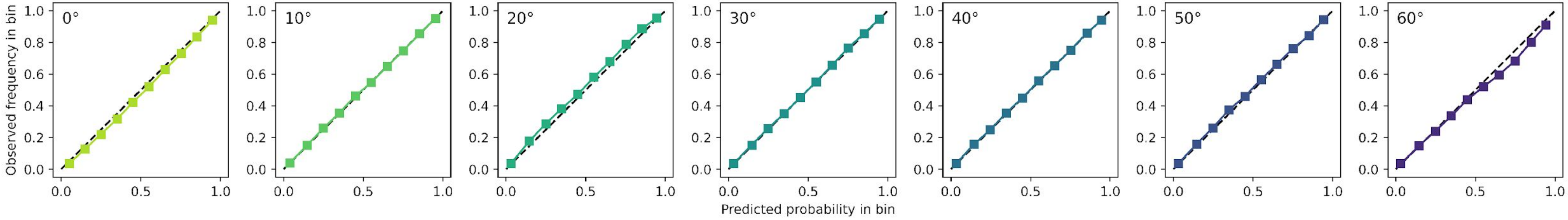


$$BSS = 1 - \frac{BS}{BS_{climatology}}$$

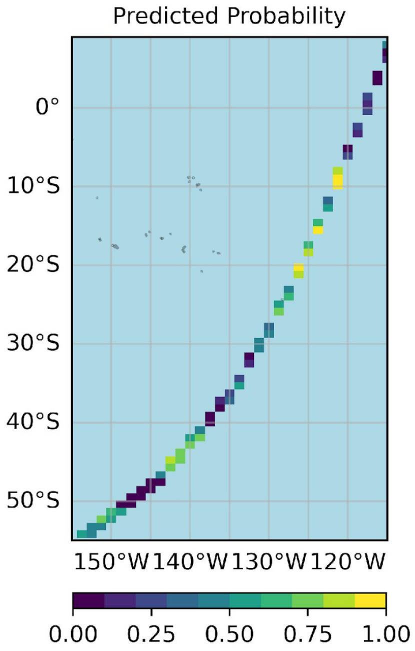
$$BS = \frac{1}{n}\sum_{i=1}^{n}(p_i - o_i)^2$$

Consistent reliability across latitude bands

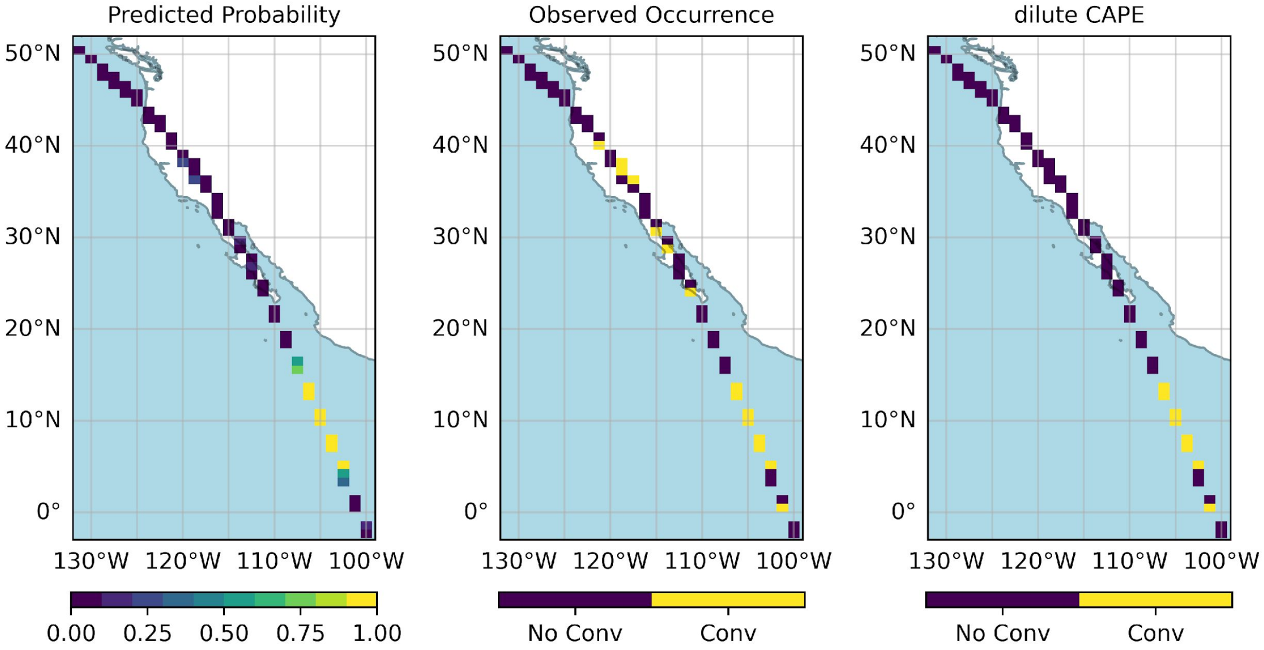
latitude increasing



Neural network compared to spatial observations

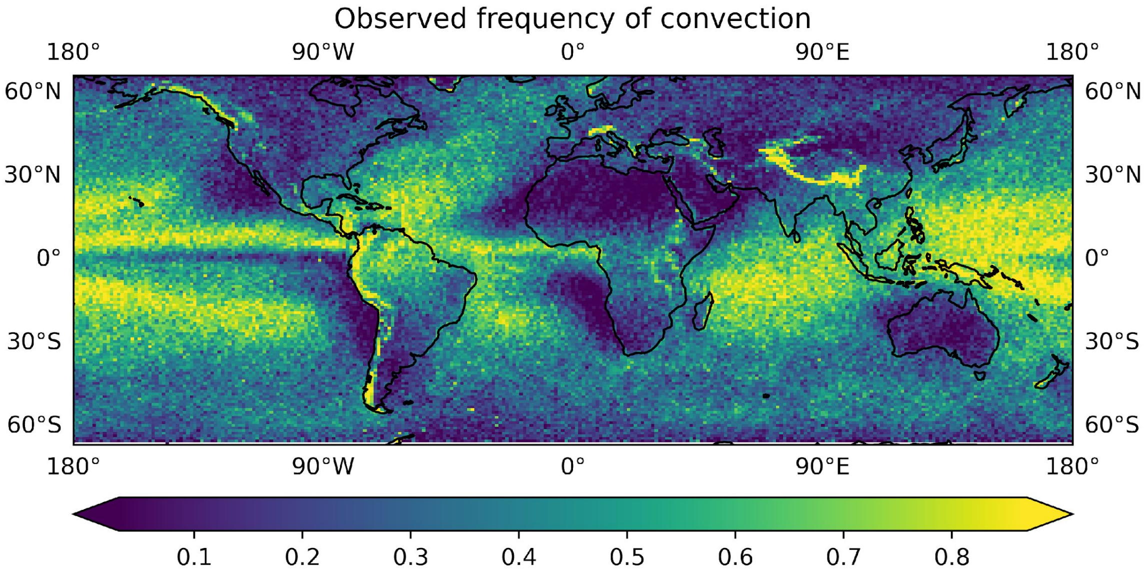


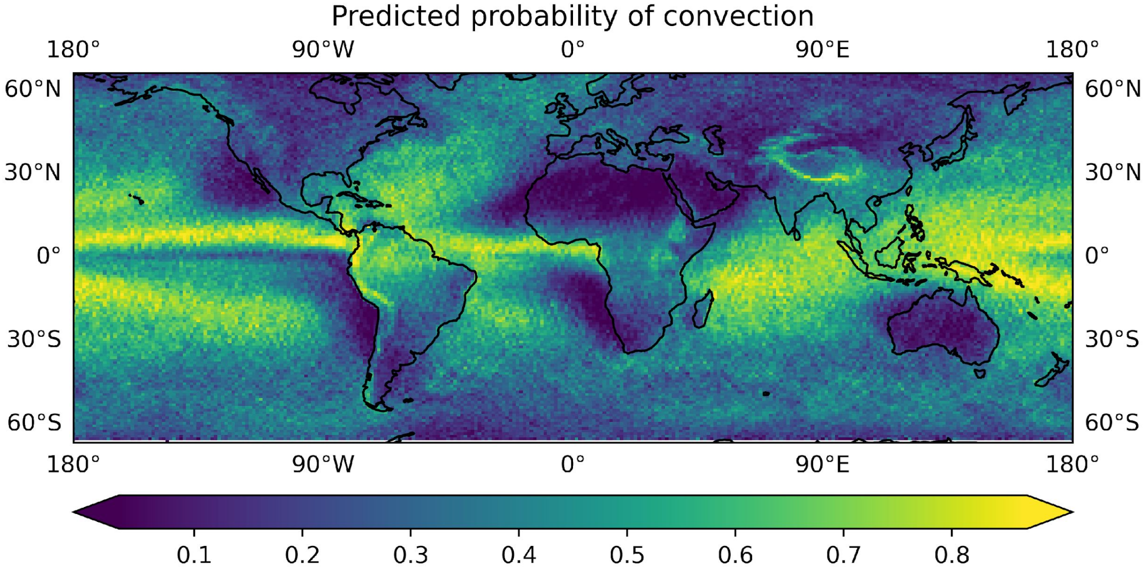
2019/07/07 5:00 UTC

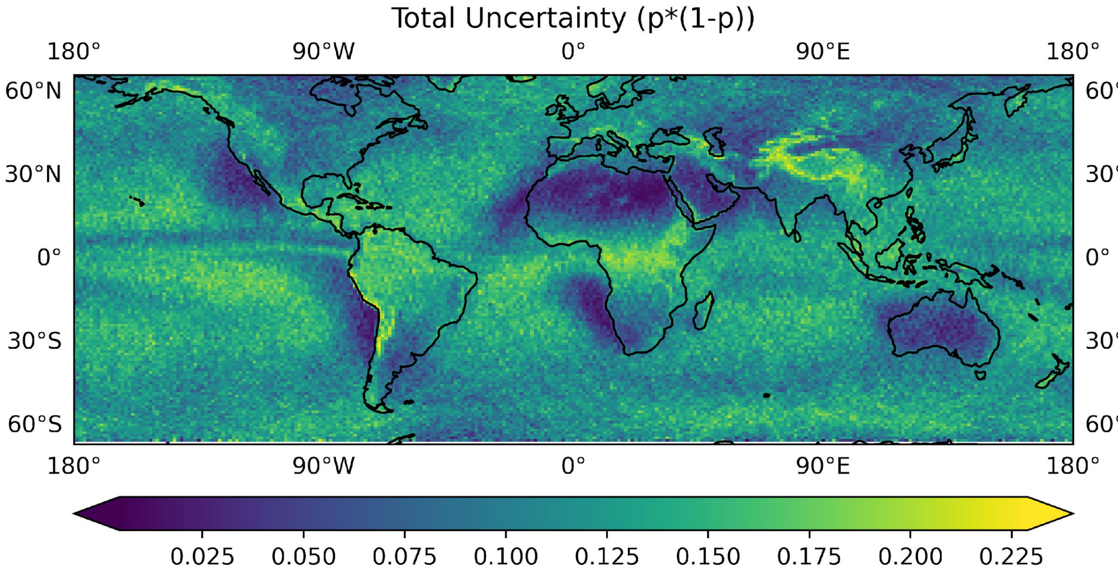


2019/07/03 17:00 UTC

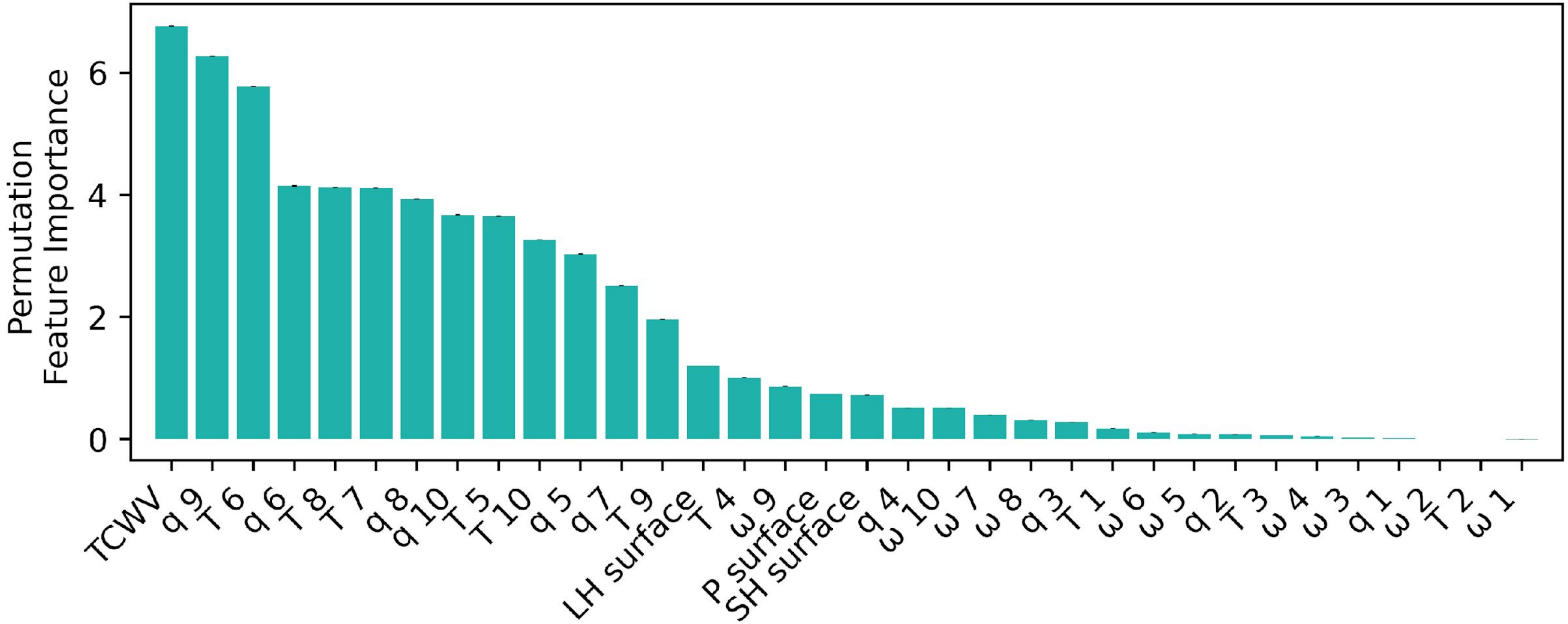
Global annually averaged maps of neural net vs observations

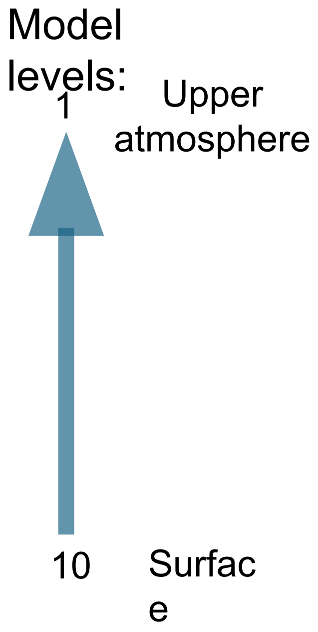






Neural network feature importance

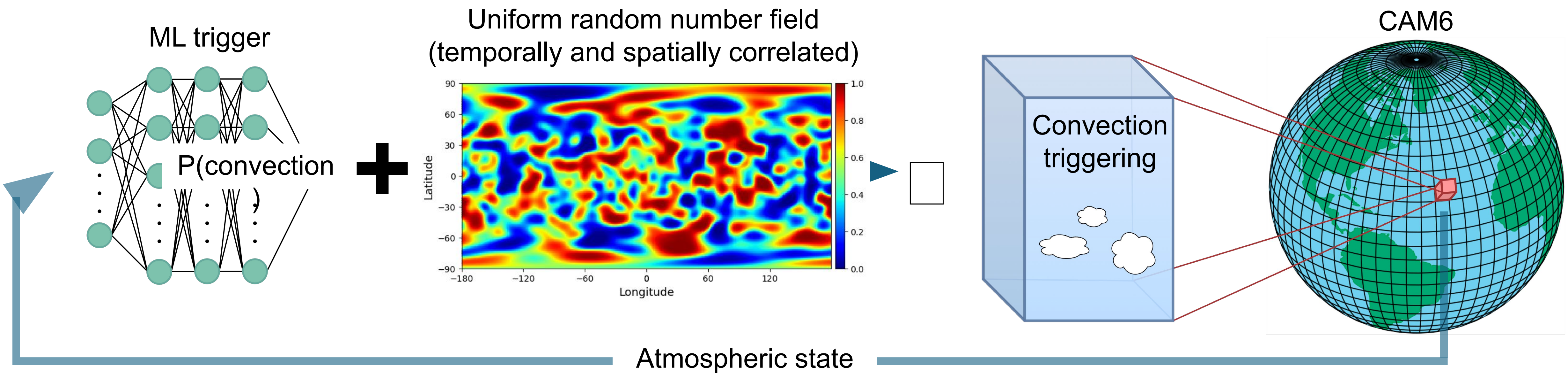




Part 2:

Implement the probabilistic ML model as a stochastic parameterization for the convection trigger in the Community Atmosphere Model (CAM6)

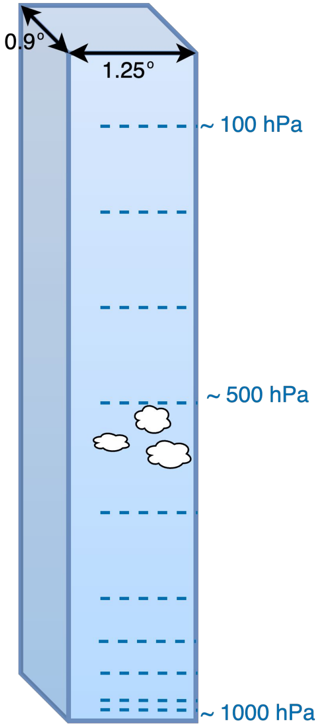
Implementing ML stochastic convection trigger in CAM6



- Quantifies uncertainty directly from observations
- Mass and energy conservation
- Generalization to warmer climates?

First implementation in CAM: simplified NN

Inputs



Temperature (levels)

Specific humidity (levels)

Total column water vapor

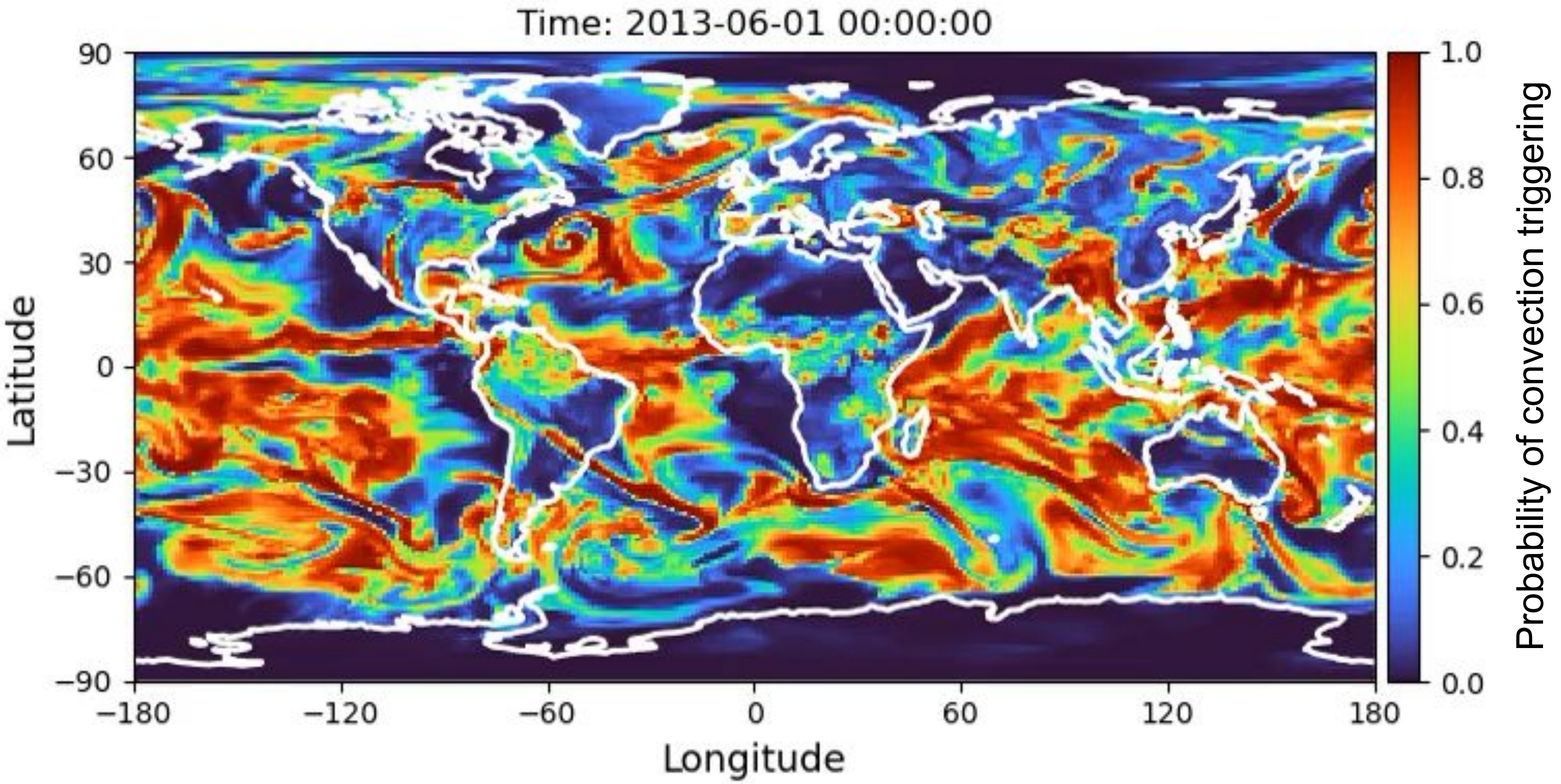
Pressure (surface)

- Replace Zhang McFarlane (ZM) deep convection trigger: dilute CAPE > 70 J/kg
- ML implementation in CAM: FTorch (Cambridge ICCS)
- Compare CAM 3-year (2011-2013) runs:
  1. Original CAPE trigger
  2. Deterministic ML trigger (P=0.5 threshold)
  3. Stochastic ML trigger (random number field)

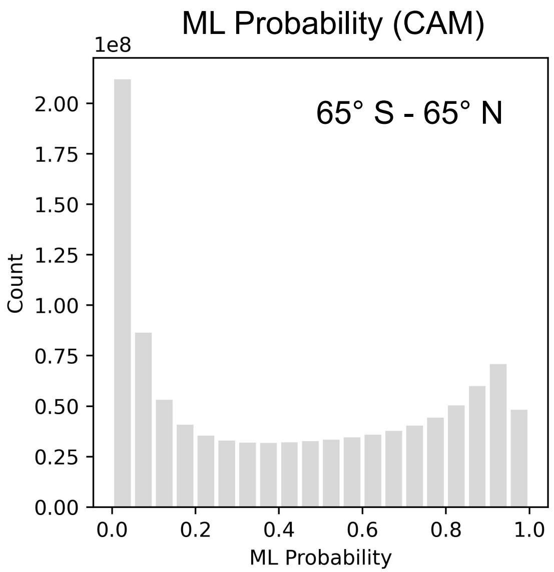
Part 3:

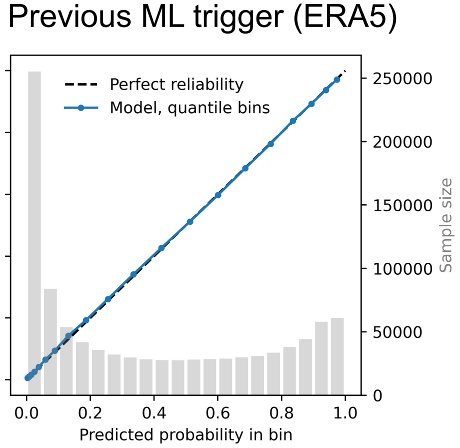
Preliminary results: ML probability of convection as a diagnostic output (not implemented as trigger)

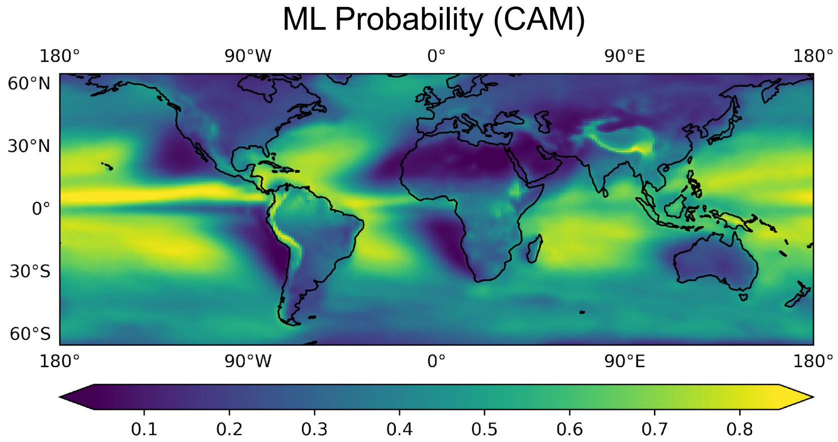
Preliminary Results: ML probability (diagnostic)

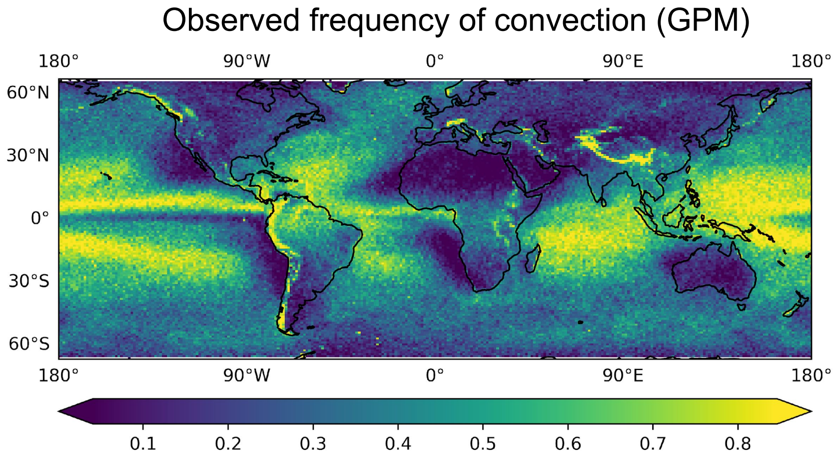


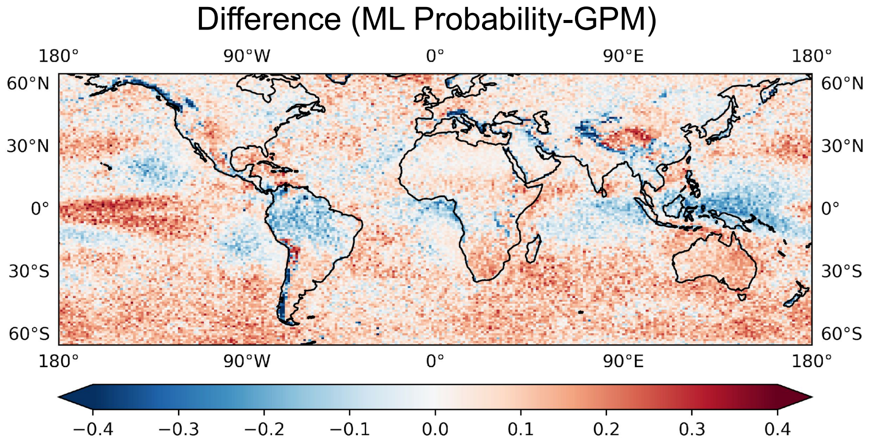
Preliminary results: ML probability (diagnostic) vs Observations





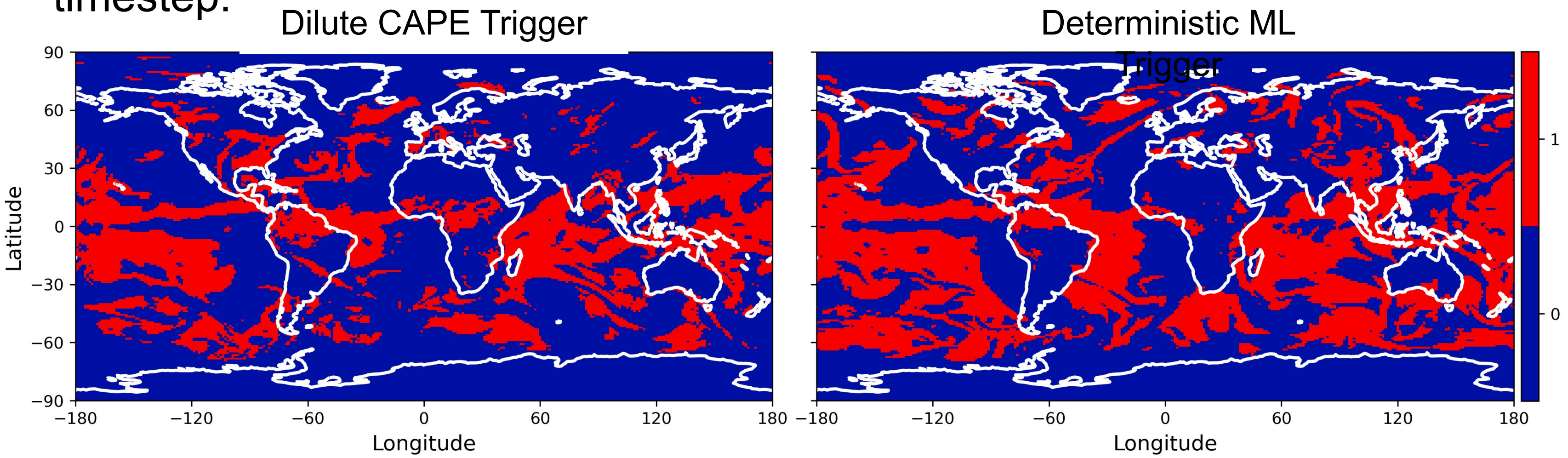




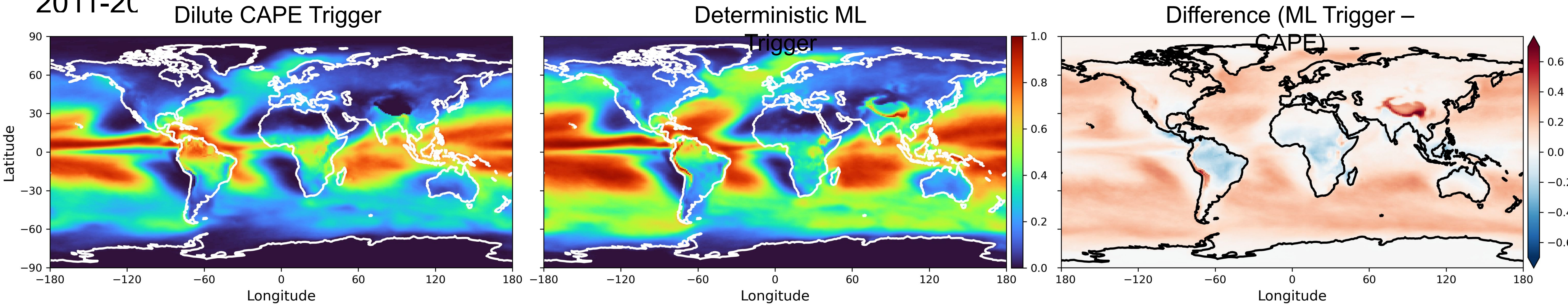


How does the ML trigger (diagnostic) compare to the CAPE trigger?

For one timestep:



Average 2011-20

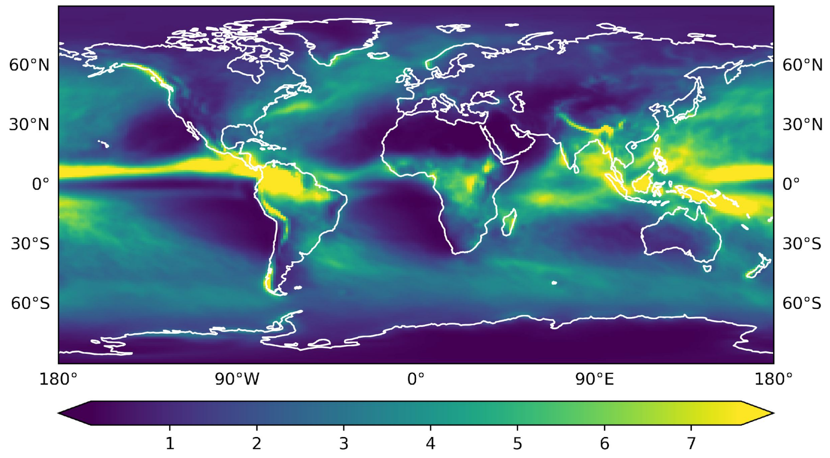


Part 4:

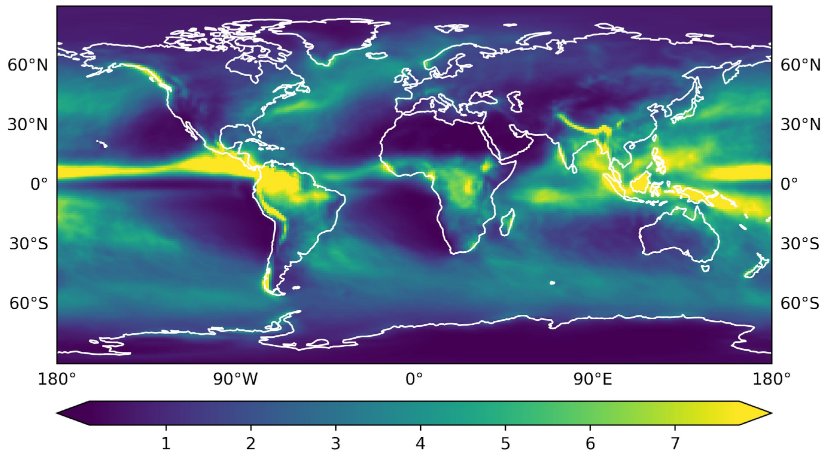
Preliminary results: ML trigger feedbacks in CAM

Preliminary Results: Global Average Precipitation

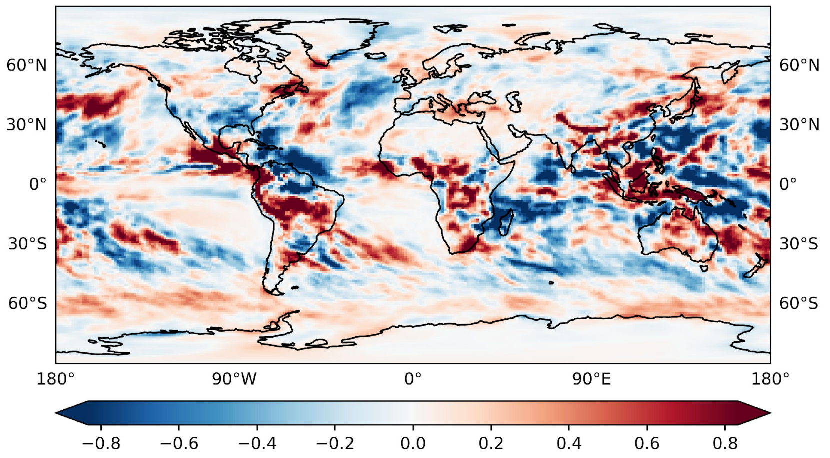
Control Precip (mm/day)



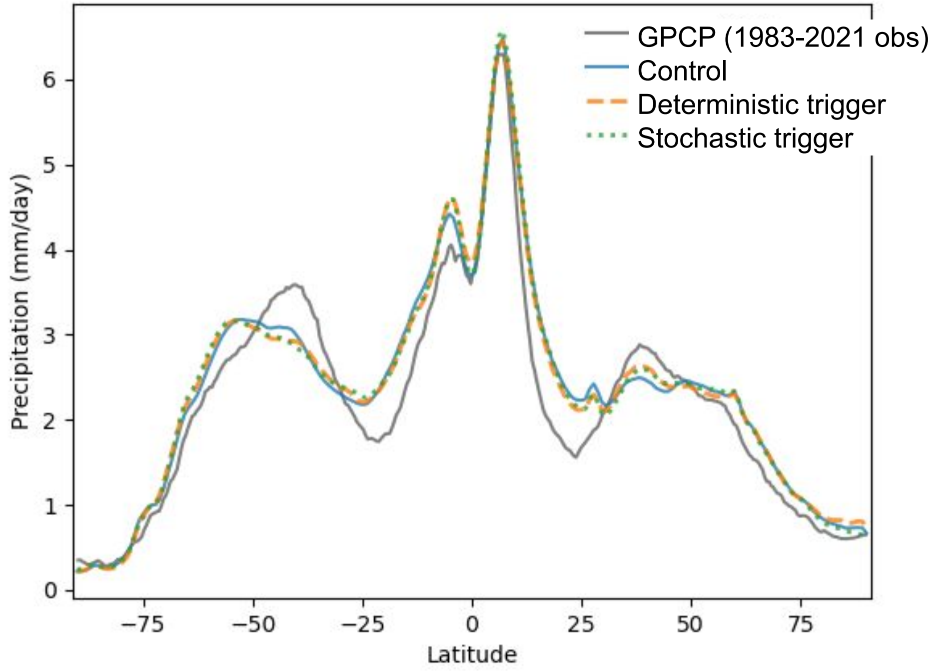
Stochastic ML Trigger Precip (mm/day)



Stochastic – Control (mm/day)



Zonal Average Precipitation



2011-2013

Preliminary Results: Diurnal Cycle (JJA)

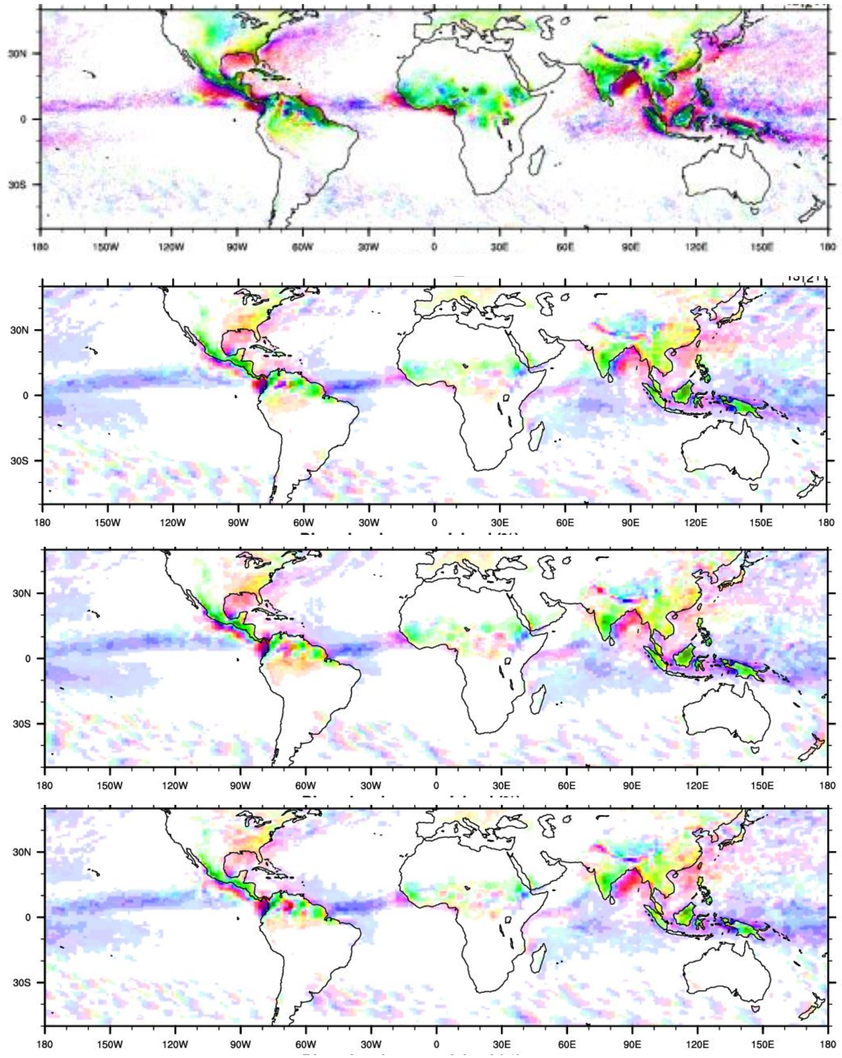
Diurnal cycle

TRMM Observations (1988-2013)

Control

Deterministic ML Trigger

Stochastic ML Trigger



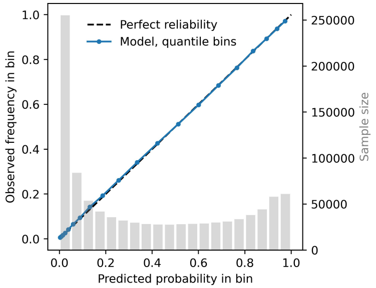
Greta Miller | greta.miller@physics.ox.ac.uk

Next Steps

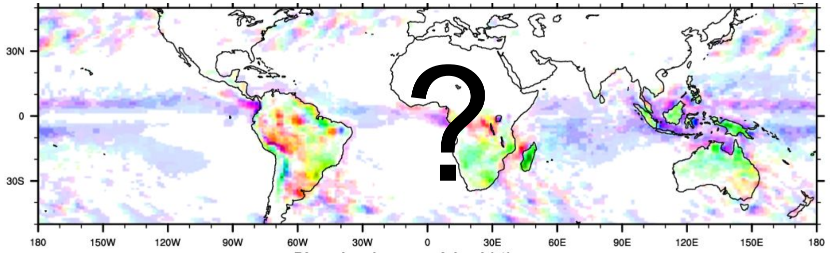
- Refine and test machine learning convection trigger
  - Add dynamics (~dCAPE), land-sea mask, topography
- Assess impact of deterministic vs stochastic model; test space and time correlation of stochastic field
- Assess impact on model mean climate, extremes, modes of variability, convection characteristics
  - Land vs ocean
  - Tropics vs extratropics
  - Summer vs winter hemisphere
- Validate convection in polar regions

Summary

The machine learning convection trigger predicts **reliable probabilities of convection occurrence** across latitudes.



Implementation in CAM will test how an **observation-based stochastic ML** trigger will impact precipitation biases (in progress).



References

Cui, Z., Zhang, G. J., Wang, Y., & Xie, S. (2021). Understanding the Roles of Convective Trigger Functions in the Diurnal Cycle of Precipitation in the NCAR CAM5. *Journal of Climate*, *34*(15), 6473–6489. https://doi.org/10.1175/JCLI-D-20-0699.1

Preliminary Results: Diurnal Cycle (Dec/Jan/Feb)

Diurnal cycle

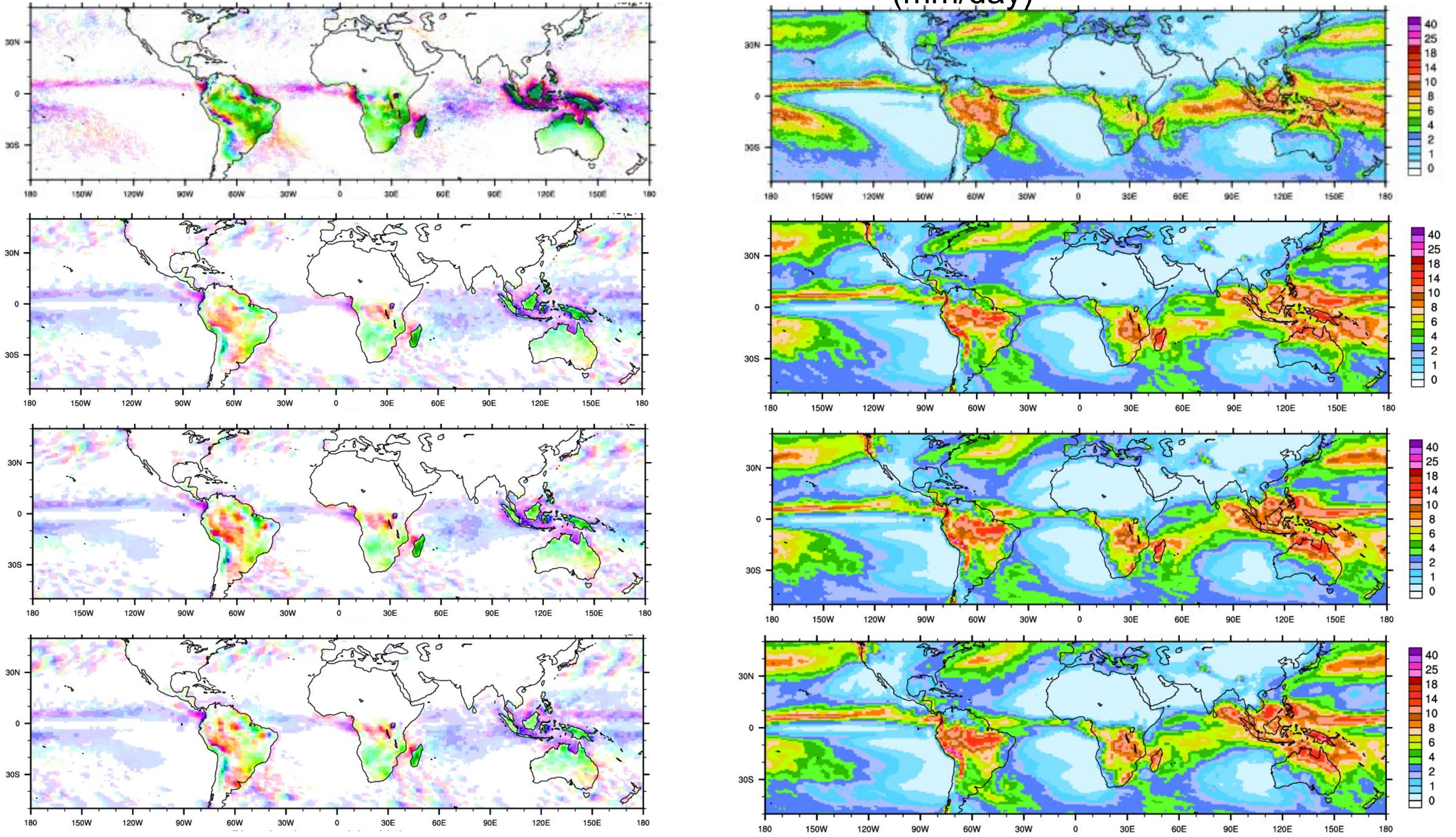
Mean Precipitation (mm/day)

TRMM Observations (1988-2013)

Control

Deterministic ML Trigger

Stochastic ML Trigger



Diurnal cycle