University of Washington
Atmospheric Sciences

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Coupling the Atmosphere and the Ocean in a Deep Learning Earth System Model

DALE DURRAN, NATHANIEL CRESSWELL-CLAY, MATTHIAS KARLBAUER, ANDY LIU, ZAC ESPINOSA, RAÚL MORENO, THORSTEN KURTH

19 MARCH 2024

Office of Naval Research
Science & Technology

NVIDIA

How Big Should a Deep Learning Weather Prediction Model Be?

- How many *predictable* degrees of freedom does the atmosphere have?
    - The number must decrease with increasing forecast lead time.
- How do we predict them?
    - *With DLWP we can choose prognostic variables (& spatial resolution) for completely different reasons than in NWP.*

Bridging the Compute Divide

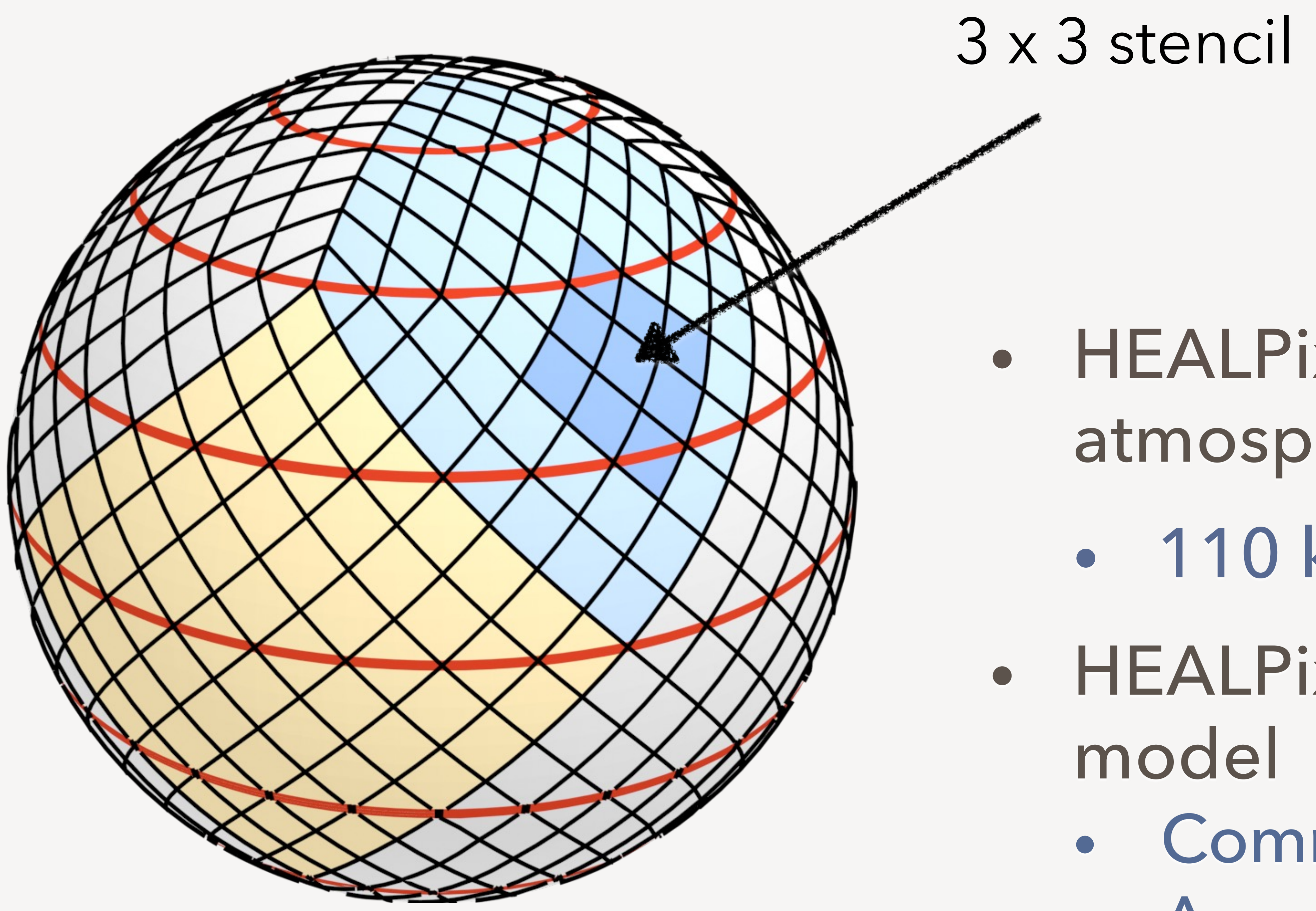
DGX Station: four A100 GPUs 2021

Model	2D Spherical Shells	Parameters
ECMWF IFS (S2S)	911	
GraphCast	227	37 M
SFNO	73	433 M
AIFS	70	
Pangu Weather	69	256 M
DLWP-HPX	8/9	3.5 M
DLOM-HPX	1	2.5 M

Generative Model

Our Parsimonious CNN Model Using 2D Spherical Shells

- Atmosphere
  - 8/9 prognostic variables
  - To ocean:
    - Z_{1000}, 10-m wind speed,
    - OLR
- Ocean
  - 1 prognostic field
  - To atmosphere
    - SST
- Prescribed fields
  - TOA solar radiation
  - Land-sea mask
  - Topographic height
  - *No latitude or longitude!*

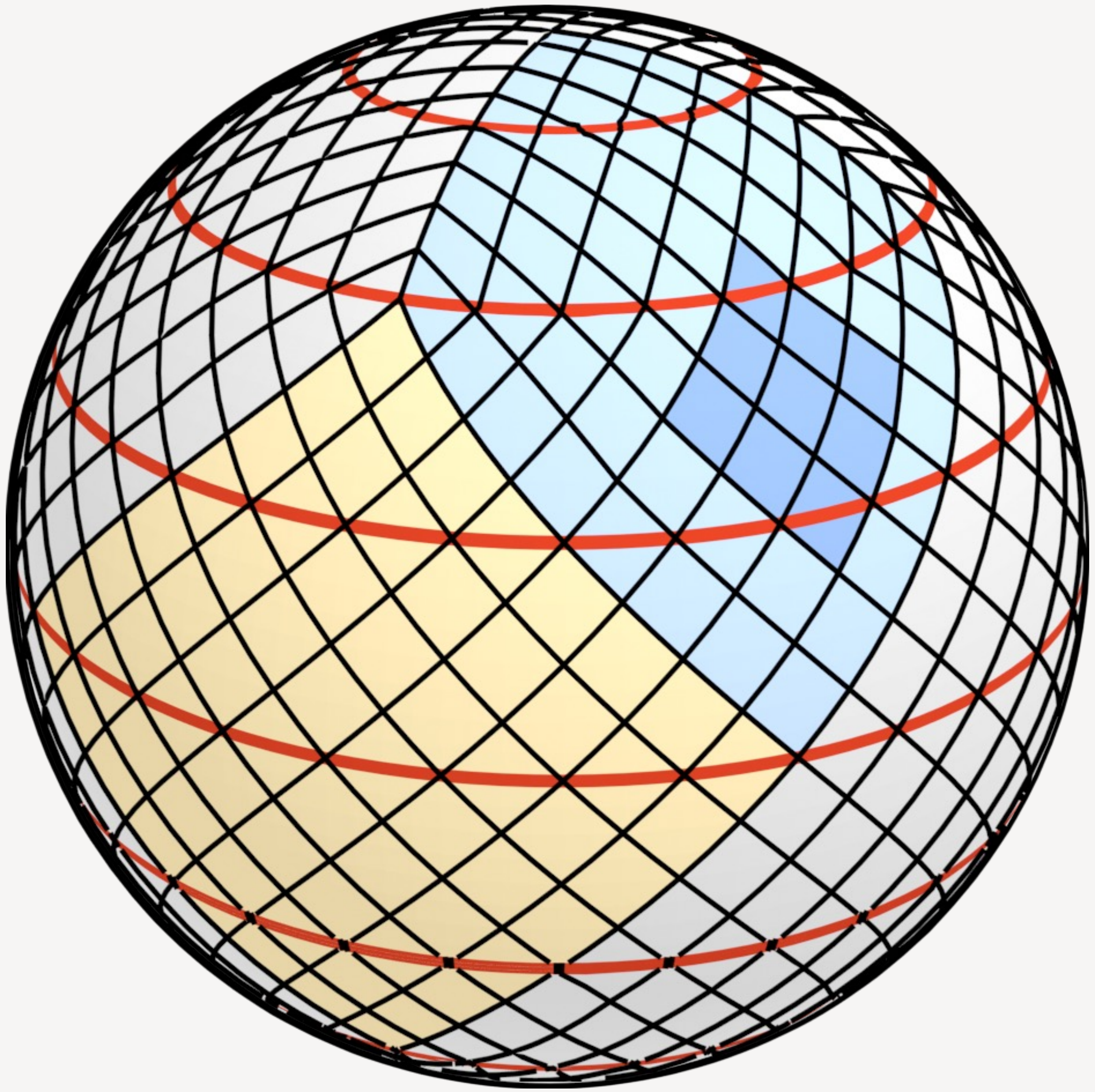


- HEALPix 64 mesh atmosphere-only model
  - 110 km grid spacing
- HEALPix 32: coupled model
  - Common in astronomy
  - Among first use in atmospheric science
  - East-to-the-Right mesh

Hierarchial Equal Area isoLatitude Pixelization

- 12 equal-area faces
- Divide into 2^n pixels along each edge
- Pixels equally spaced along latitude lines
- The “right” way to do quads on the sphere
- “East to the right” mesh

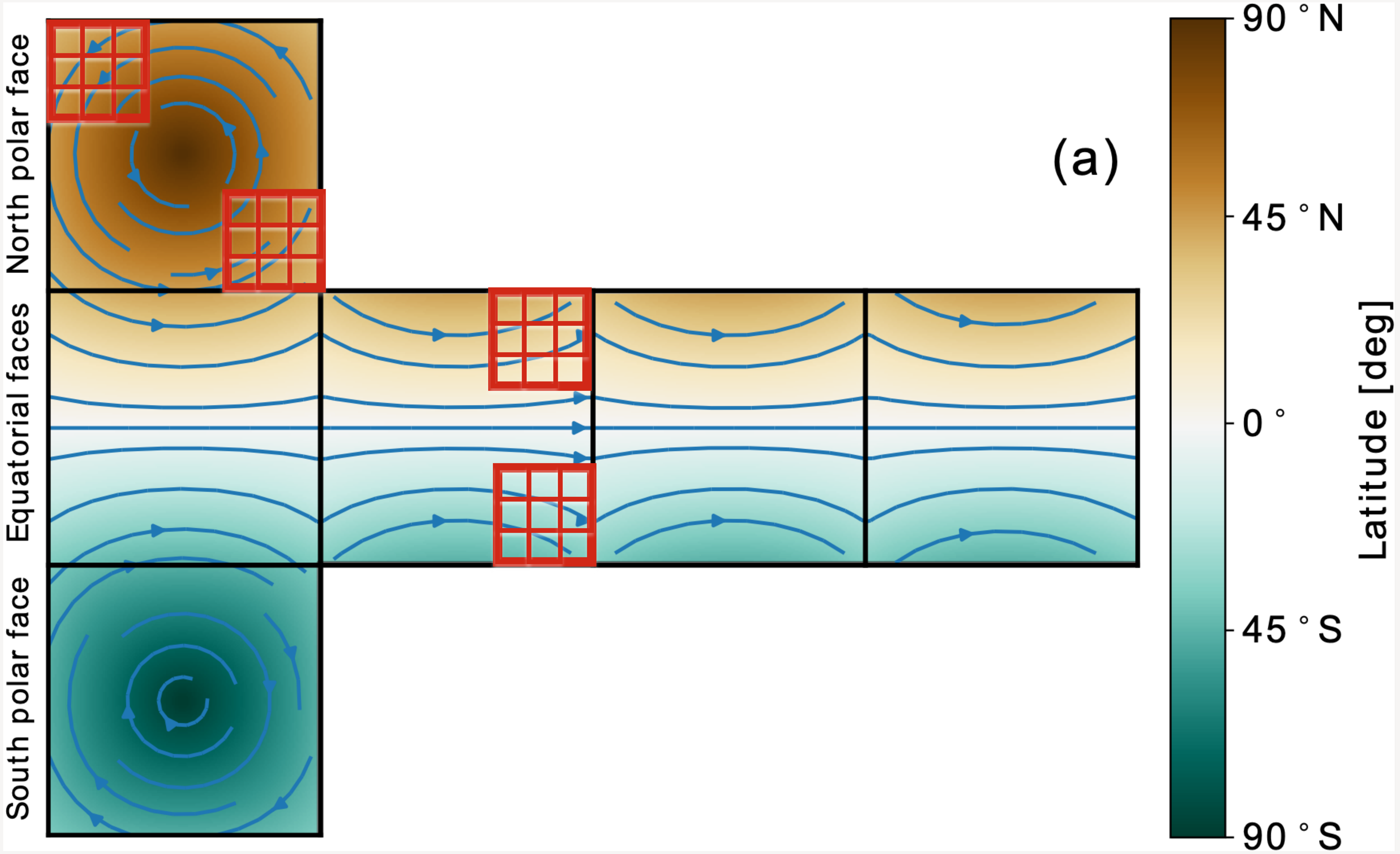
2^n=8

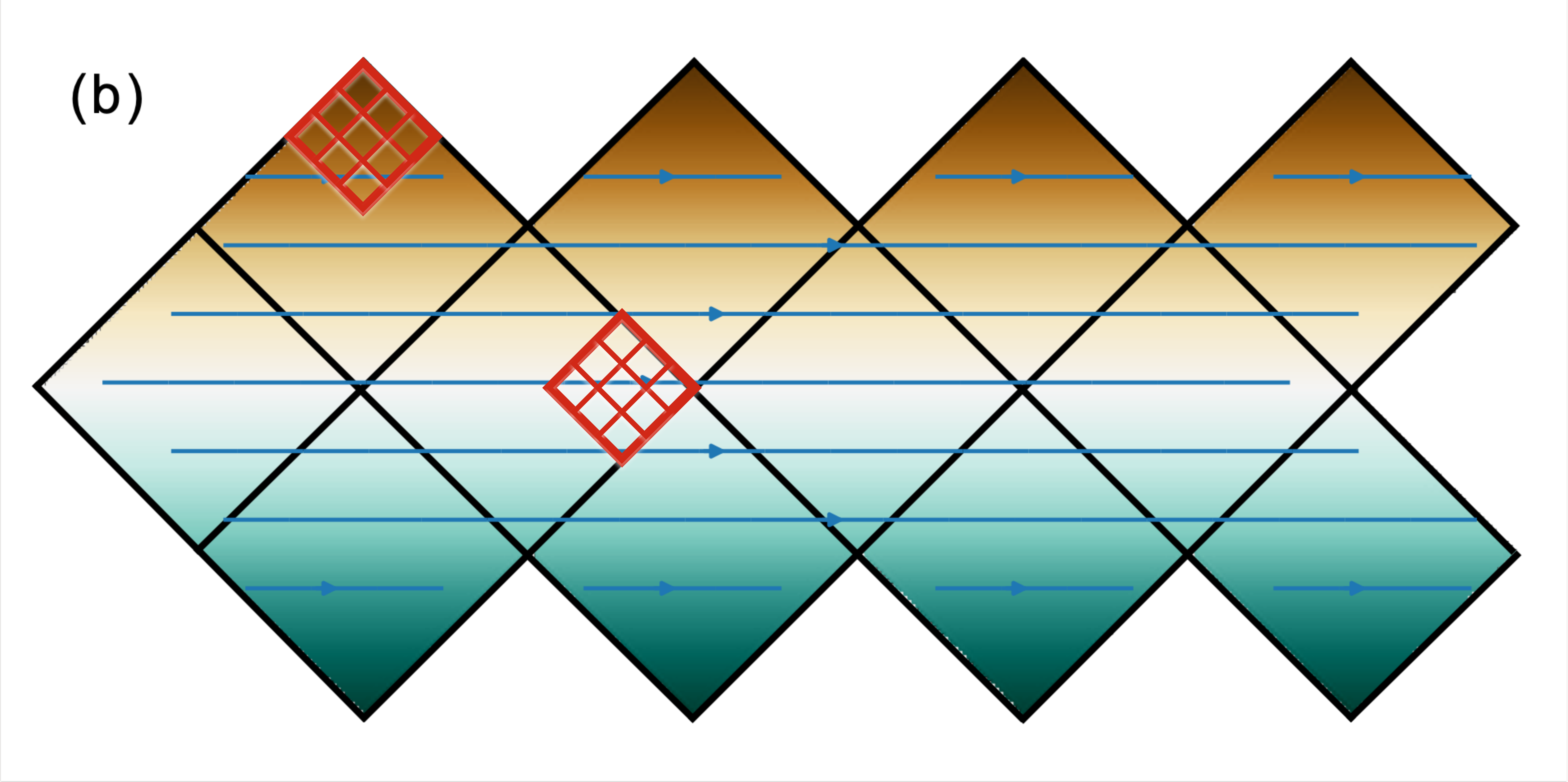


WHY IS THE HEALPIX MESH SO MUCH BETTER?

Cube Sphere

HEALPix

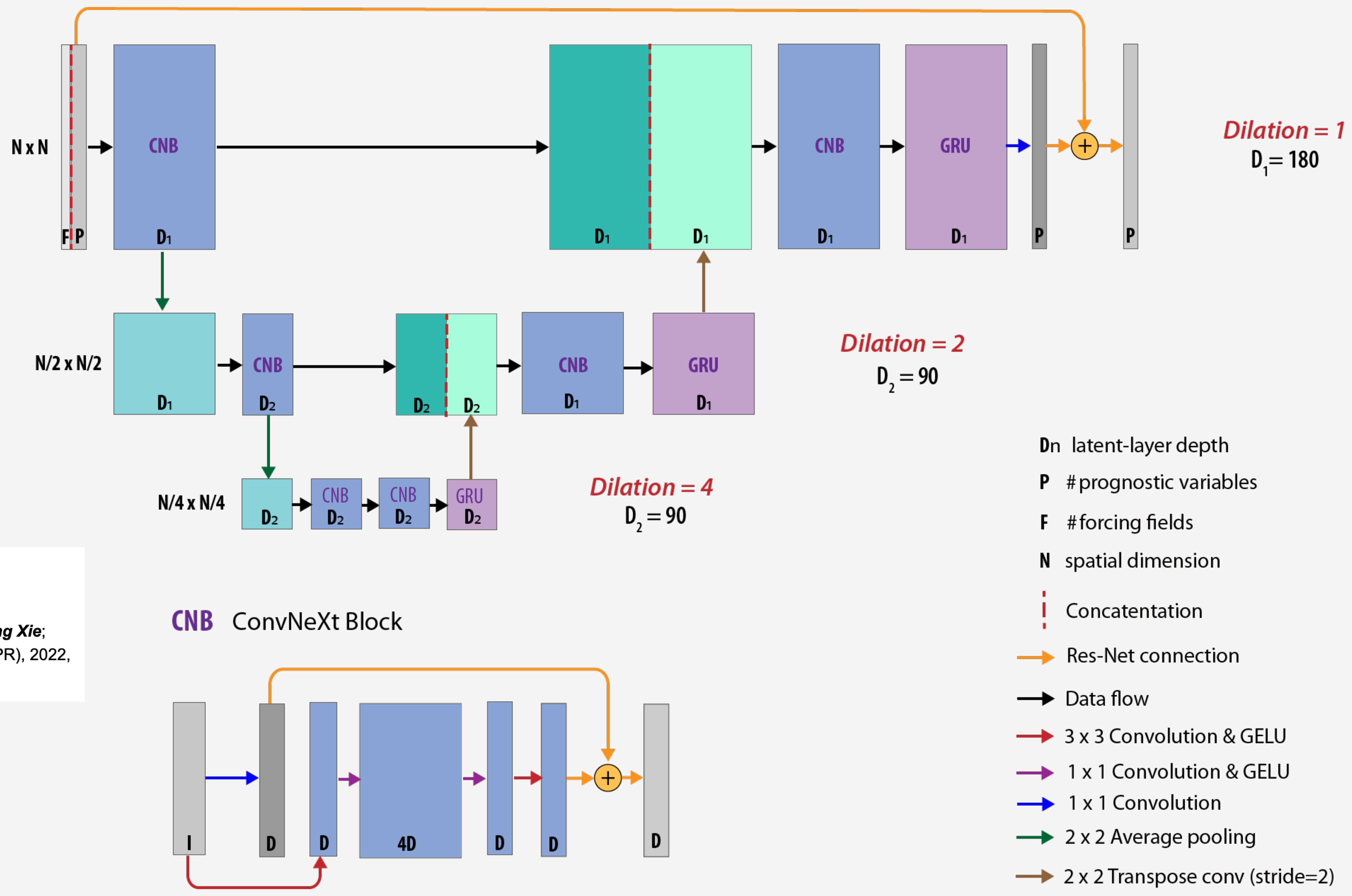




"EAST IS TO THE RIGHT"

MODEL ARCHITECTURE

- Inverted channel depth
- Recurrence in the latent space
- Dilation give large receptive field

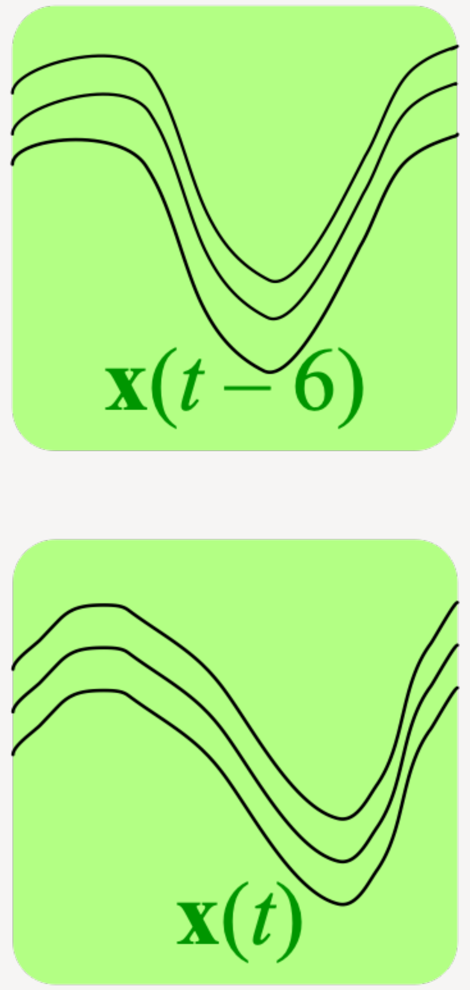


A ConvNet for the 2020s

Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, Saining Xie; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 11976-11986

"We demonstrate that a standard ConvNet model can achieve the same level of scalability as hierarchical vision Transformers while being much simpler in design."

TIME STEPPING



Observed state at two times — *12 hr time step* — *Forecast state at two times* — *12 hr time step* — *Forecast state at two times*

$\tilde{\mathbf{x}}(t-6, t) \rightarrow$ Algorithm $\rightarrow \tilde{\mathbf{y}}(t+6, t+12) \rightarrow$ Algorithm $\rightarrow \tilde{\mathbf{y}}(t+18, t+24)$ $\cdots$

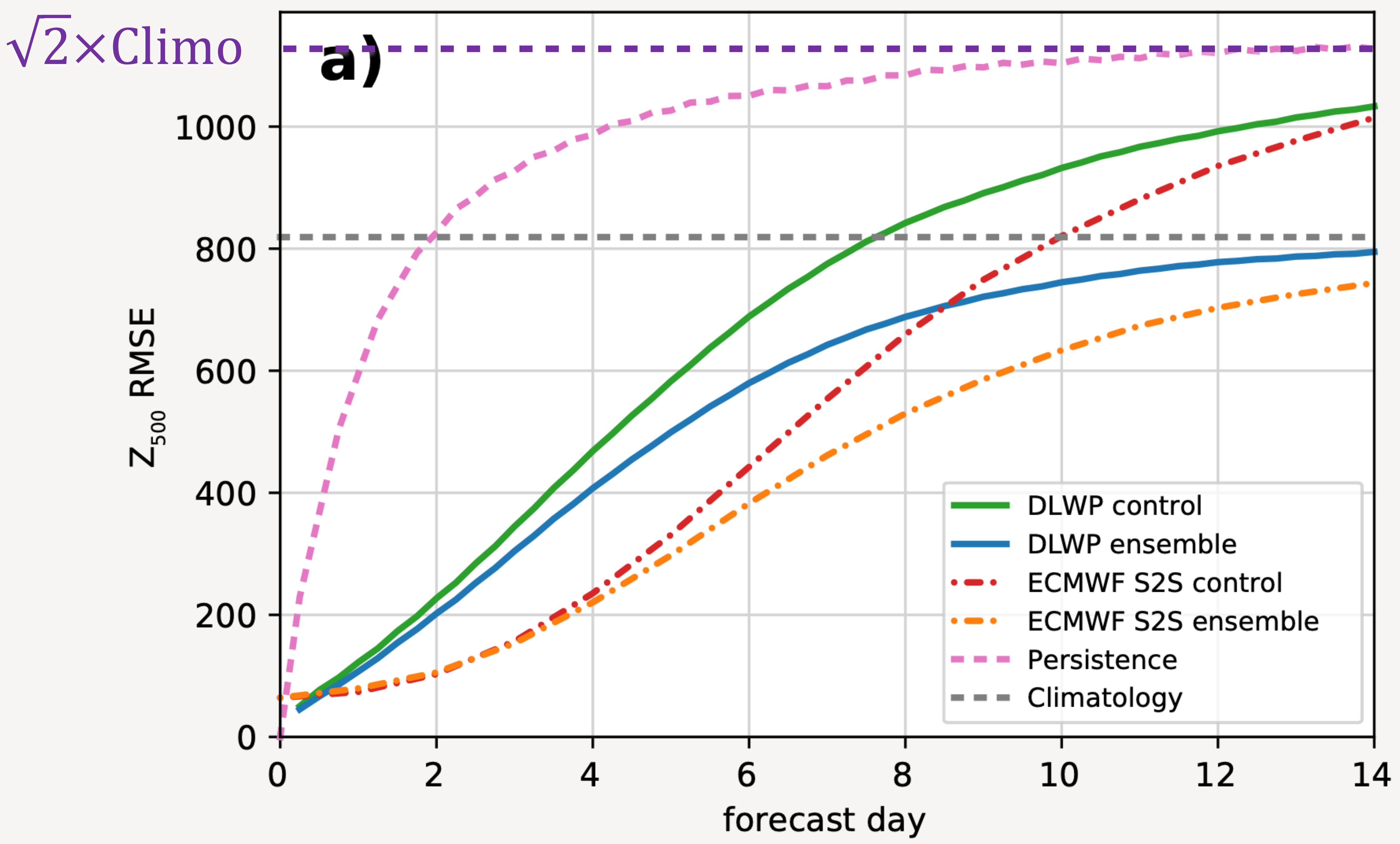
$$J_1 = \left\| \mathbf{y}(t+6) - \mathbf{x}(t+6) \right\|_2^2 + \left\| \mathbf{y}(t+12) - \mathbf{x}(t+12) \right\|_2^2 \qquad J_2 = \left\| \mathbf{y}(t+18) - \mathbf{x}(t+18) \right\|_2^2 + \left\| \mathbf{y}(t+24) - \mathbf{x}(t+24) \right\|_2^2$$

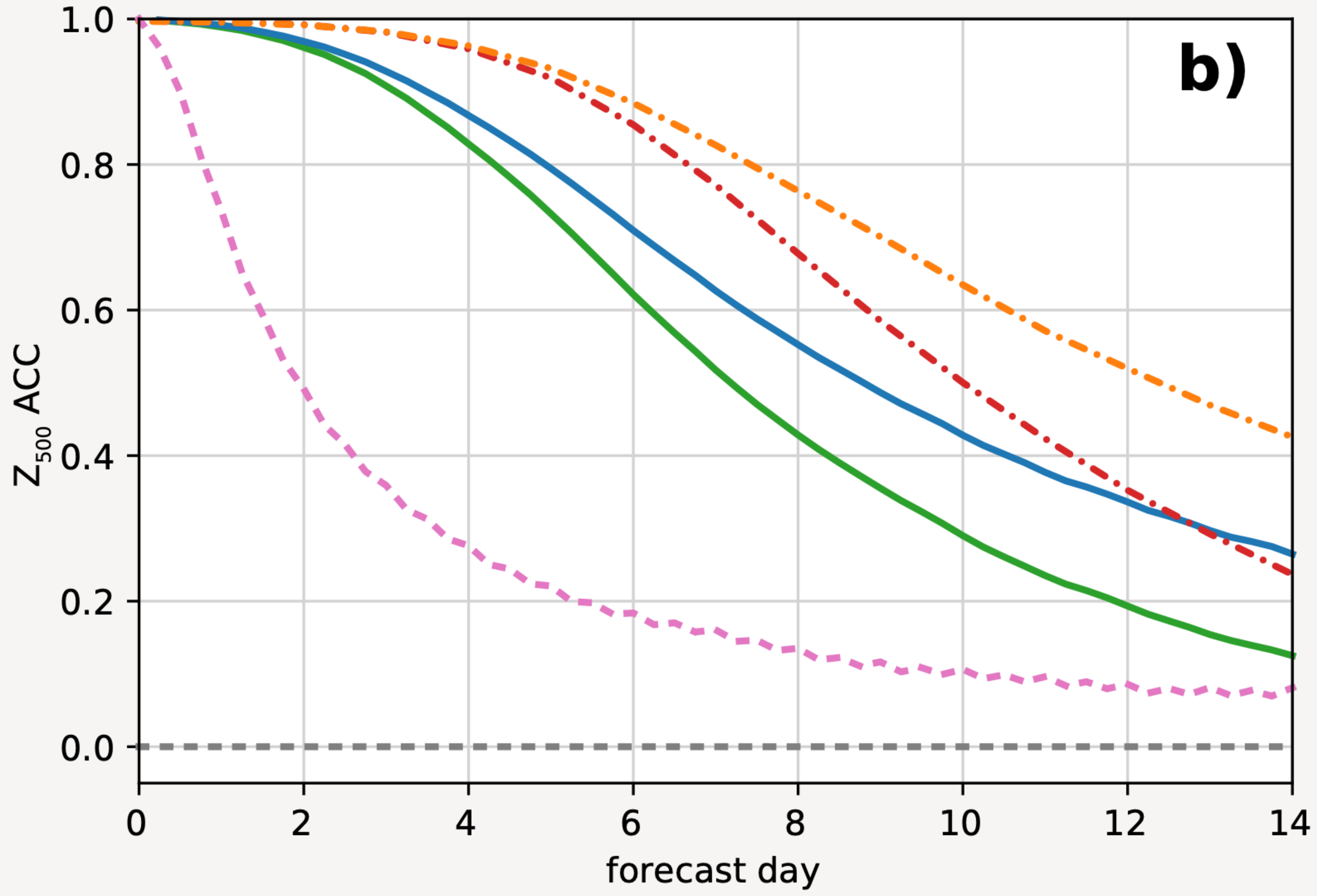
Diurnal cycle drives many weather phenomena

$$J = J_1 + J_2$$

*Loss function is the sum of the MSE over **24 hrs***

PHYSICAL INTERPRETATION OF MULTI-DAY LOSS FUNCTIONS: ENSEMBLES VS SINGLE STATES

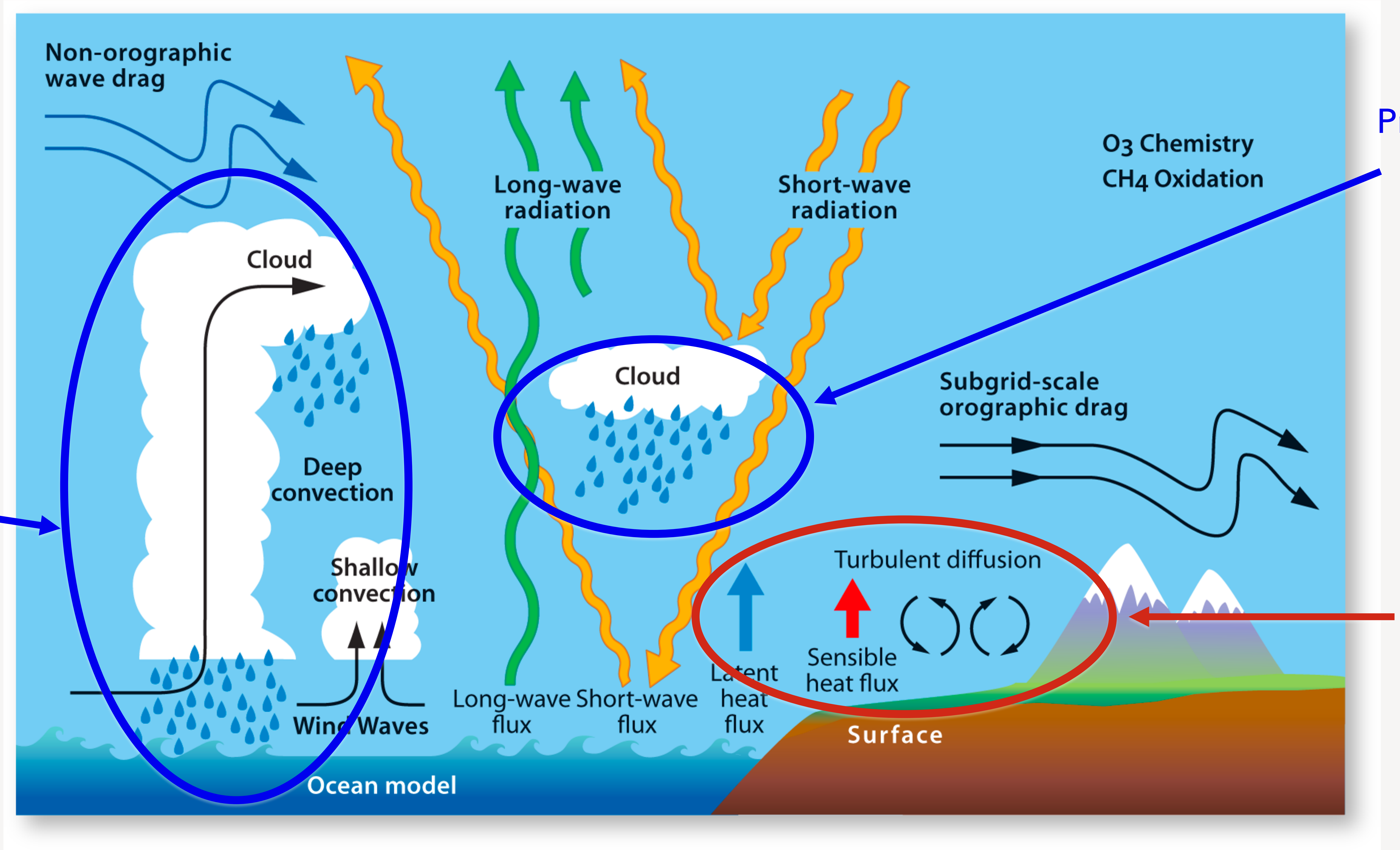




- DLWP (Weyn et al., 2021 shown) and IFS show difference between single run and Monte-Carlo ensemble
- AI forecast models with long multi-day loss functions train toward an ensemble forecast
  - GraphCast (Lam et al., 2022): 3-day loss; FengWu (Chen et al., 2023) 10-day loss!
  - Brenowitz et al., 2024 A Practical Probabilistic Benchmark for AI Weather Models

Avoiding Parameterized Physics

ECMWF Parameterized Processes



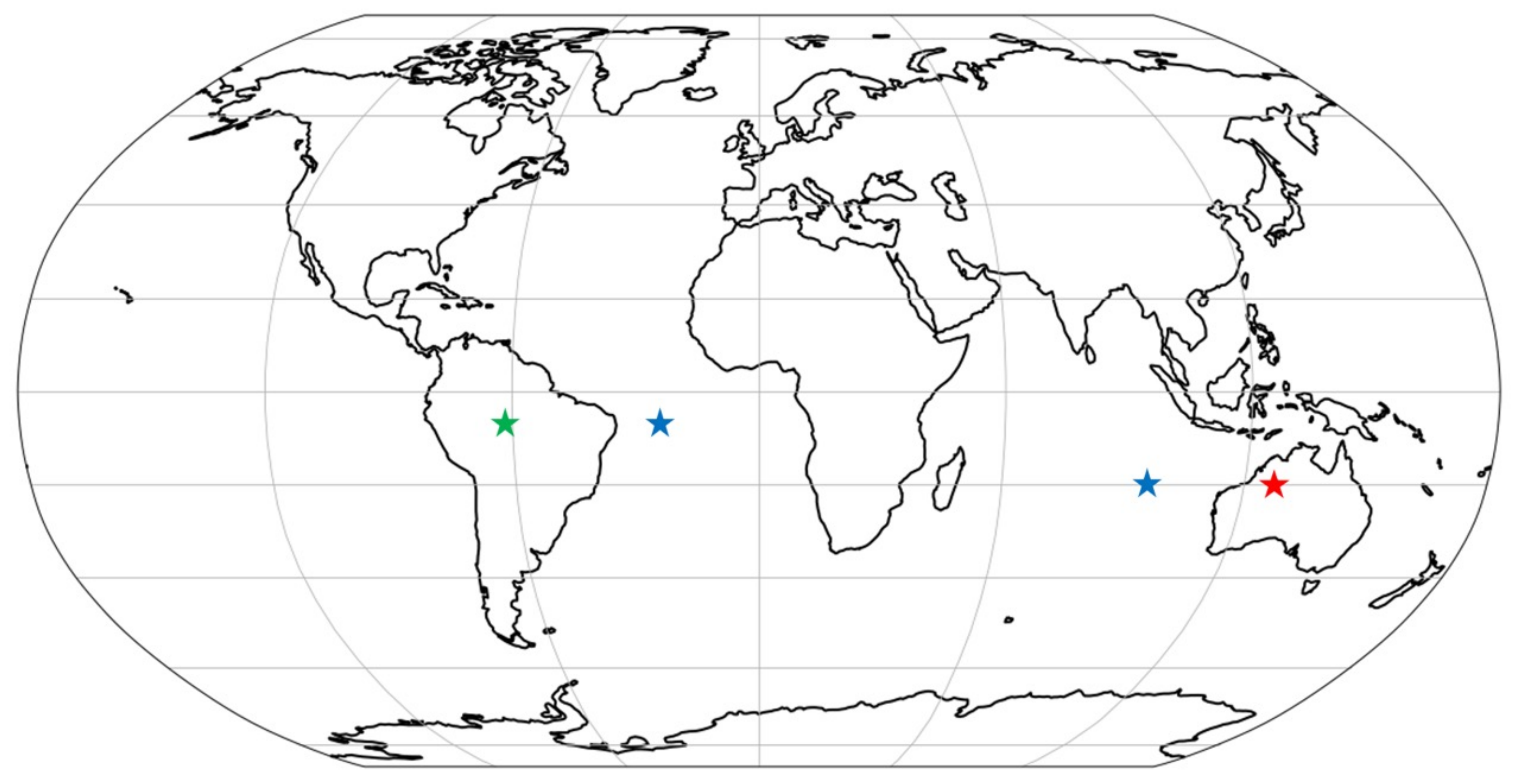
Precipitation from resolved-scale ascent

Precipitation from deep convection

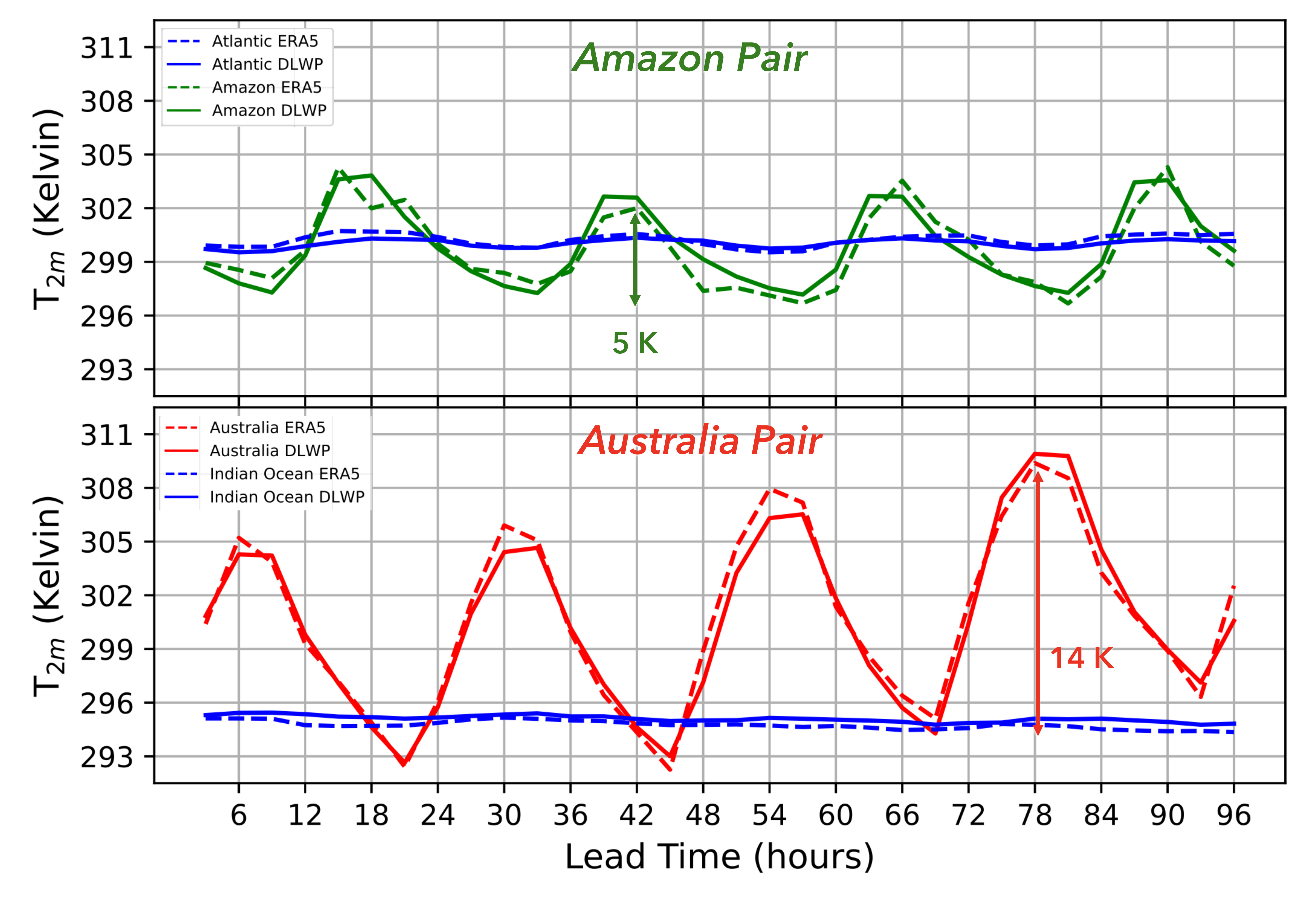
Boundary layer: temperature & surface wind

SURFACE TEMPERATURES *WITHOUT* A BOUNDARY-LAYER PARAMETERIZATION

- 2-m temperature
- 2 paired sites
  - Amazon & ocean
  - Australia & ocean
- 4-day forecast
  - Initialized March 12, 2018 at 00 UTC



2-m Temperature Forecasts (8 points per day)



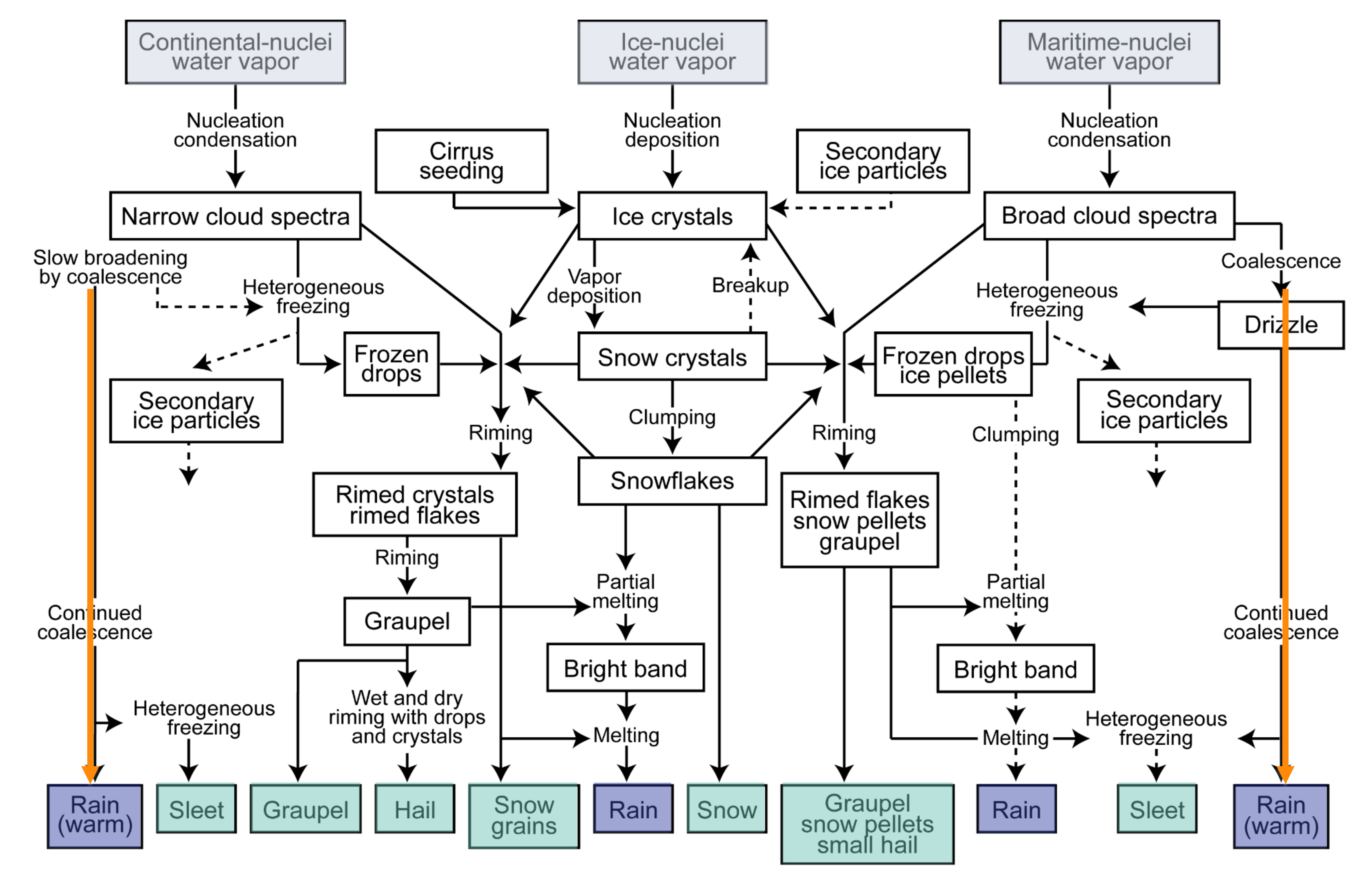
- Little temperature variation over oceans (land-sea mask)
- **Larger diurnal variations over Australia than the Amazon**
  - Total column water vapor?

- ***No geo-specific NN weights***

  (CNN is translation invariant)
- ***No lat-lon input data***

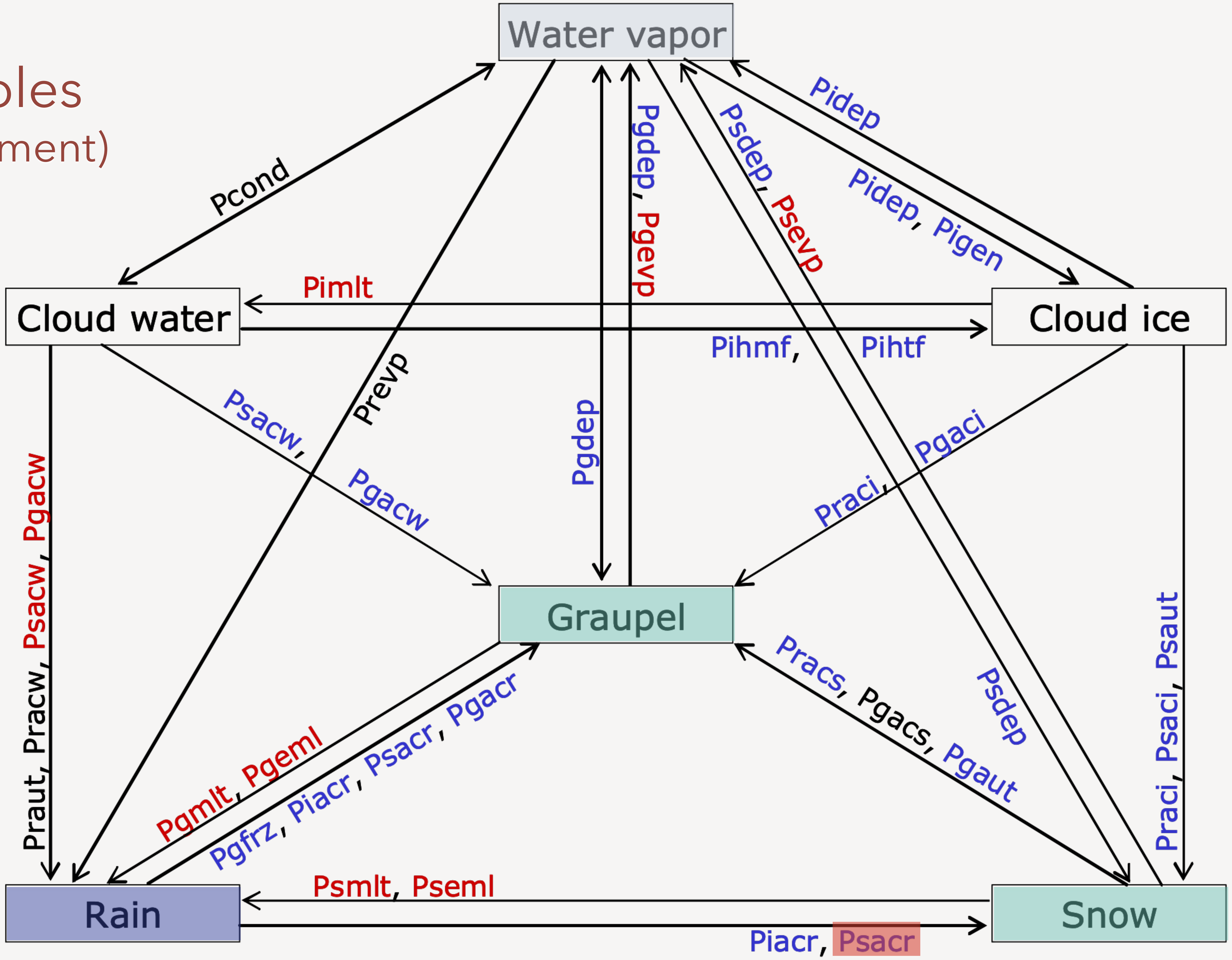
Cloud and Precipitation Processes: Underlying Physics

Rain without ice processes



WSM6 Parameterization of Cloud and Precipitation Processes in NWP

6 variables
(single moment)



Snow accreting rain

$$\mathrm{Psacr}[\mathrm{kgkg^{-1}s^{-1}}] = \pi^2 E_{SR} n_{0R} n_{0S} \left(\frac{\rho_W}{\rho}\right) \times |V_S - V_R| \left(\frac{5}{\lambda_R^6 \lambda_S} + \frac{2}{\lambda_R^5 \lambda_S^2} + \frac{0.5}{\lambda_R^4 \lambda_S^3}\right),$$

Snow size distribution

Intercept parameter

$$n_{0S}\ (\mathrm{m^{-4}}) = 2 \times 10^6 \exp[0.12(T_0 - T)]$$

Slope parameter

$$\lambda_S = (\pi \rho_S n_{0S} / \rho q_S)^{1/4}$$

Hong and Lim, 2006

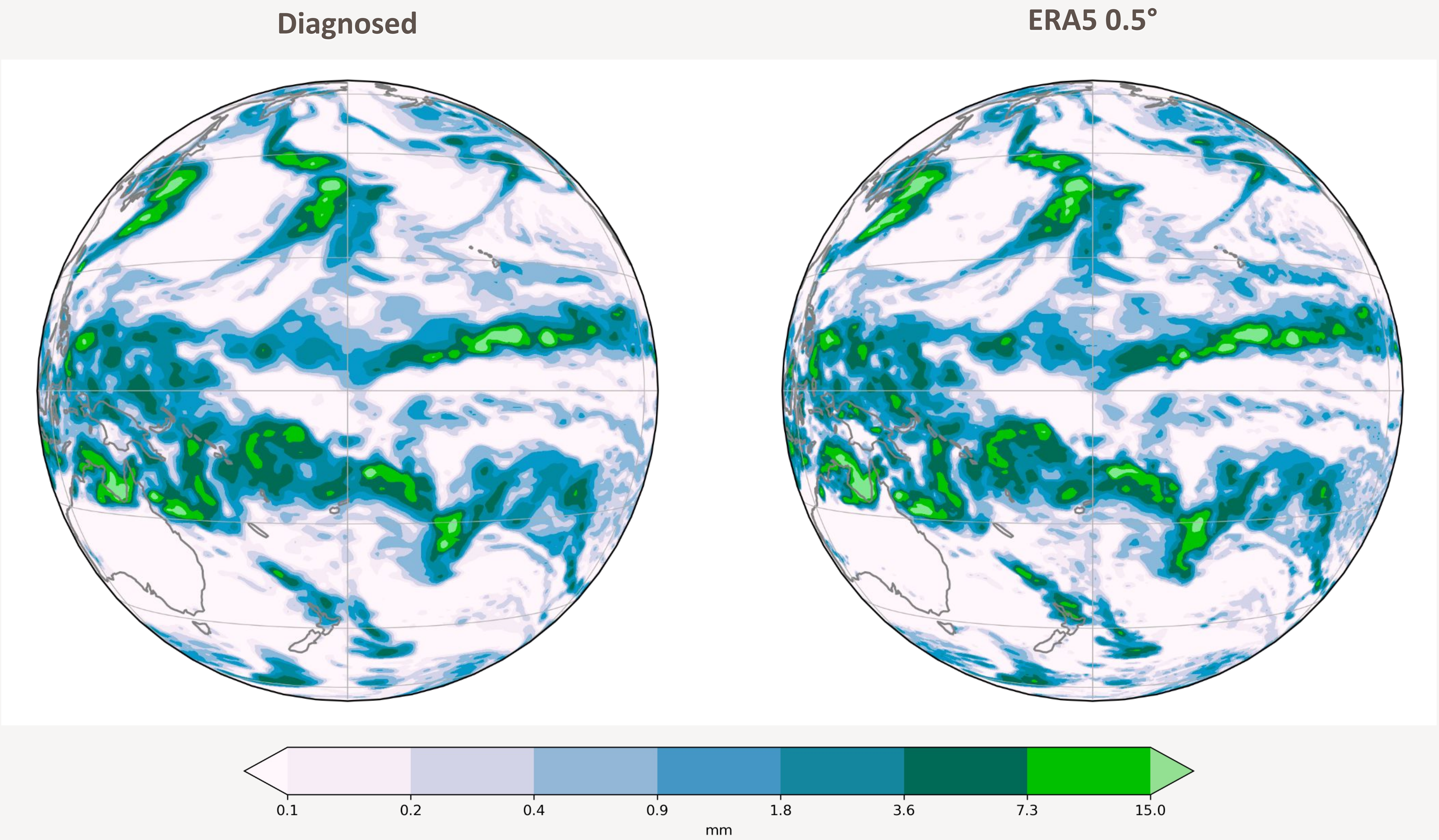
Can We Determine Precipitation From Other Physical Fields?

- Not a microphysics emulator
  - WSM6 uses $T(z)$, $q_v(z)$, $q_c(z)$, $q_i(z)$, $q_r(z)$, $q_s(z)$, $q_g(z)$, …
- Determine precipitation from the model fields
  - Prognostic fields
    - T_{2m}, T_{850}
    - Z_{1000}, Z_{500}, Z_{250}, $(Z_{300}$-$Z_{700})$
    - 10 wind speed
    - Total-column water vapor
  - Specified fields
    - Surface elevation
    - Land-sea mask
    - Top of atmosphere solar

PRECIPITATION *DIAGNOSIS* FROM *OBSERVABLE* FIELDS

- *55 km grid spacing (HEALPix 128)*
- *No*
  - Microphysics
  - Convective parameterization
  - Resolved convection

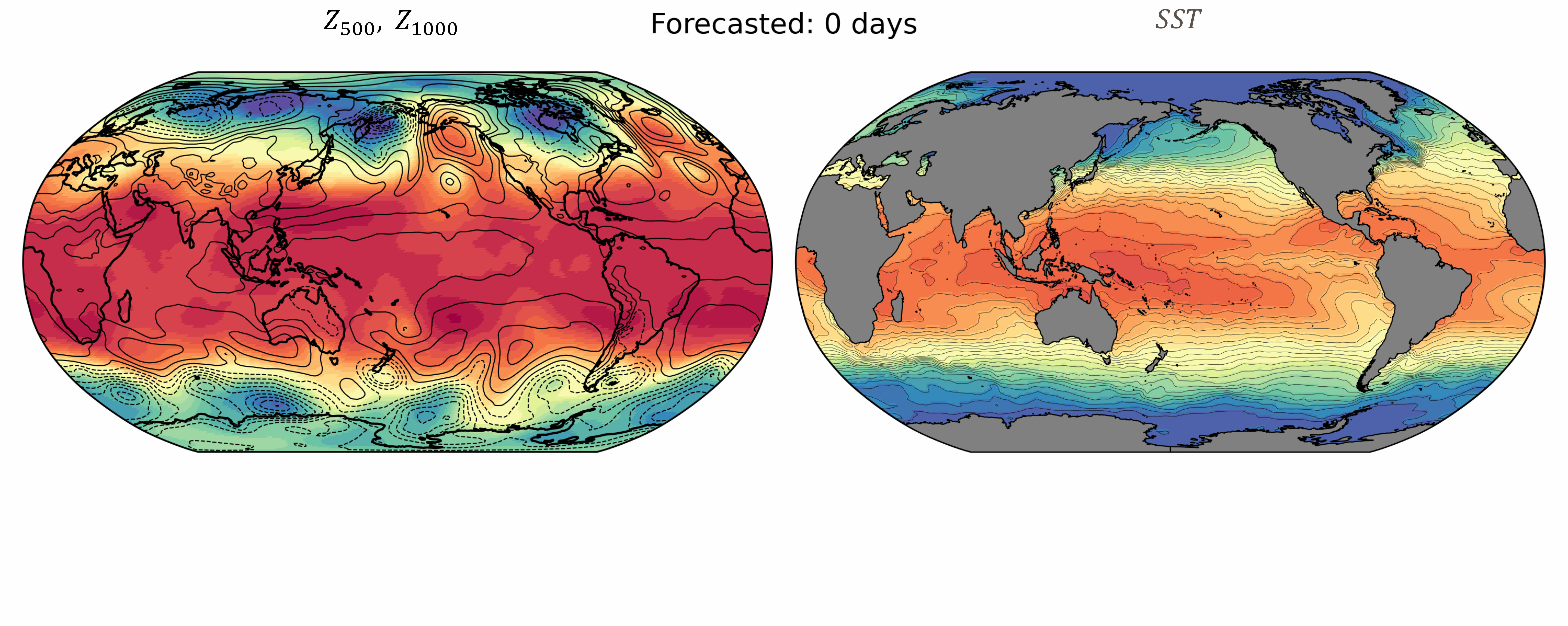
6-h Precipitation: 12 UTC 22 January 2018



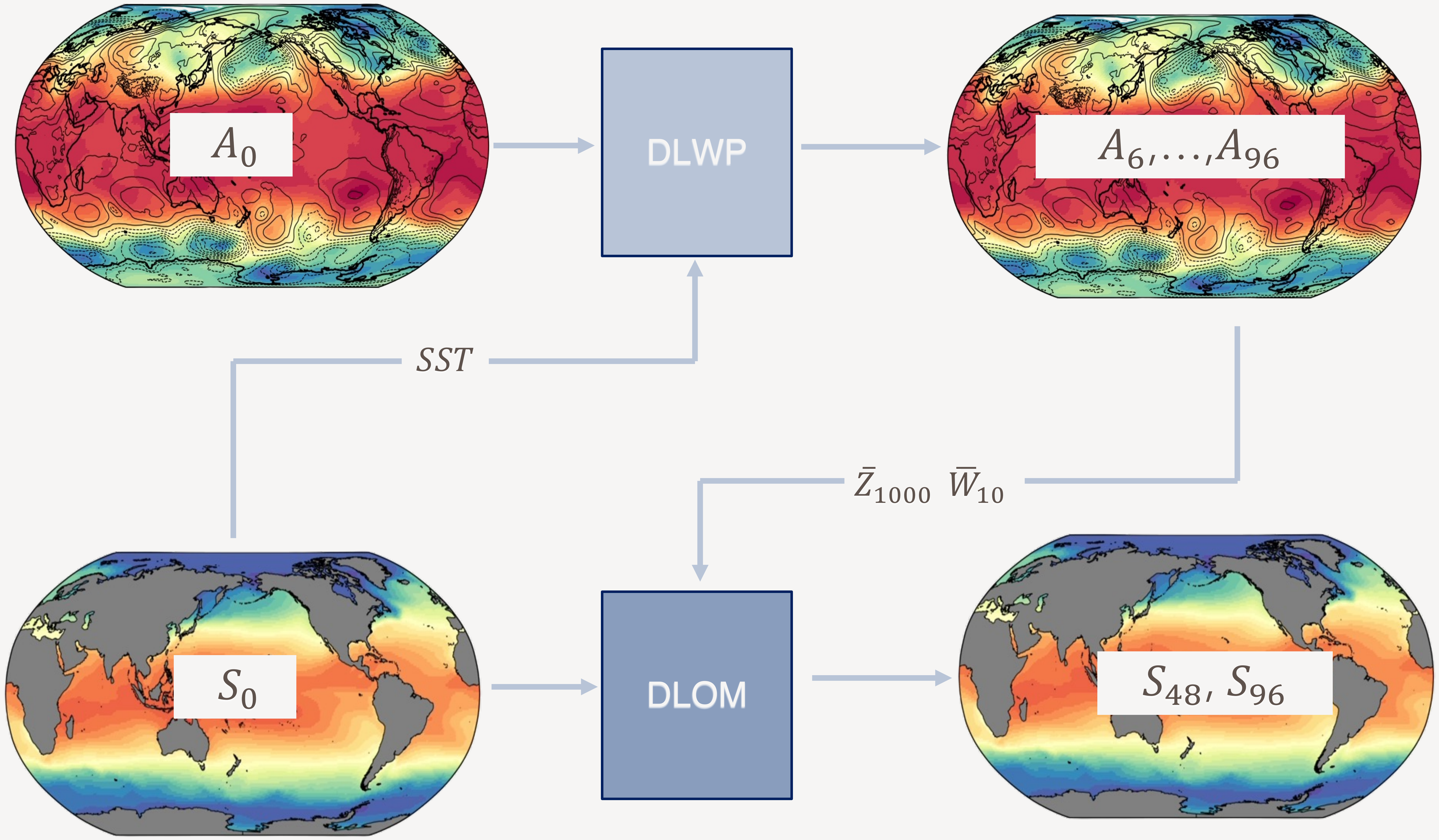
Basic Deep Learning Earth-System Model

Couple the atmosphere and the ocean.

Atmosphere-Ocean Coupling: Time Scales



Atmosphere-Ocean Coupling: Architecture



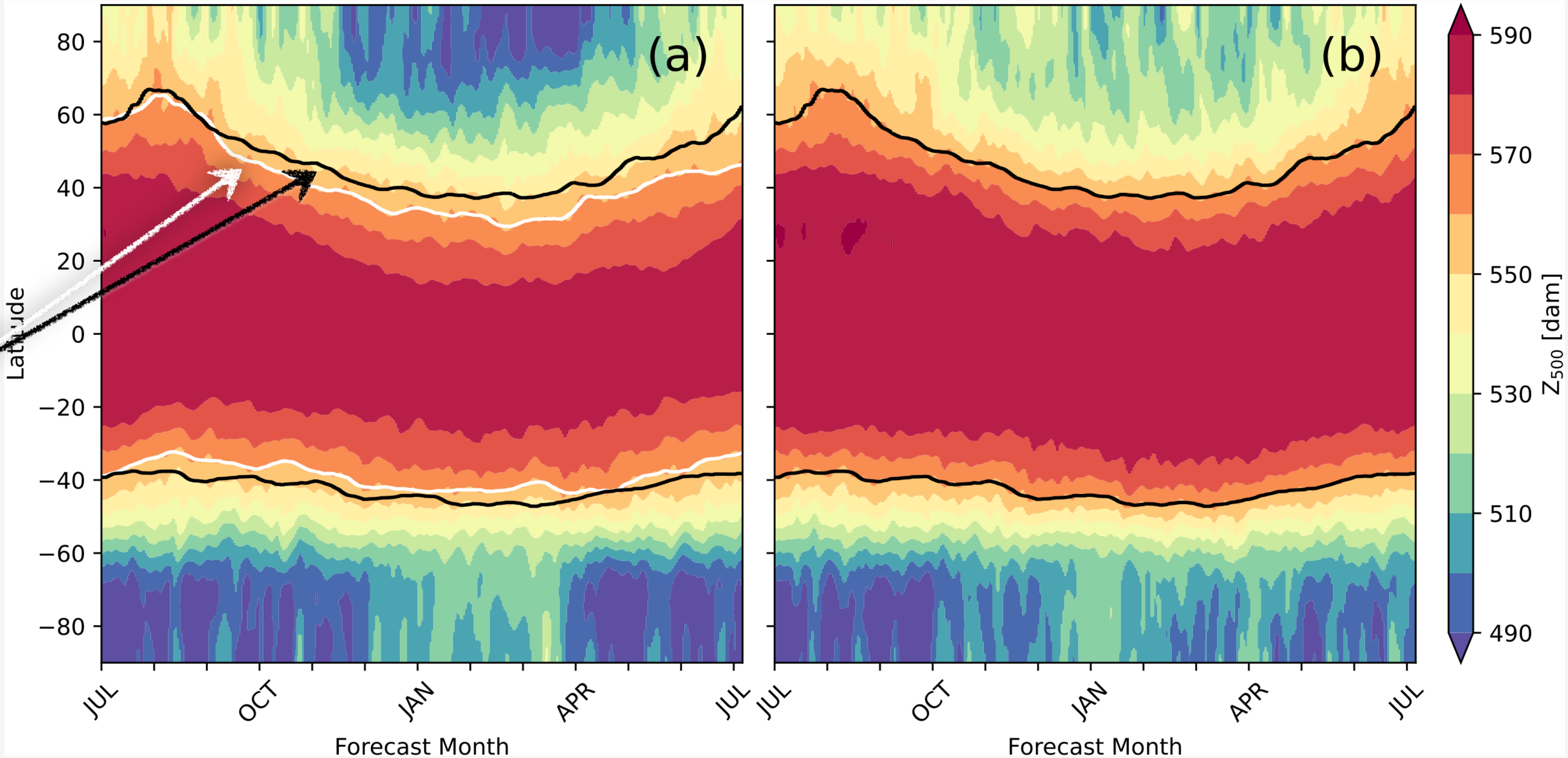
1-Year Rollout: *Atmosphere Only* (1460 steps)

Zonally averaged 500 hPa height; 3-d mean

DLWP

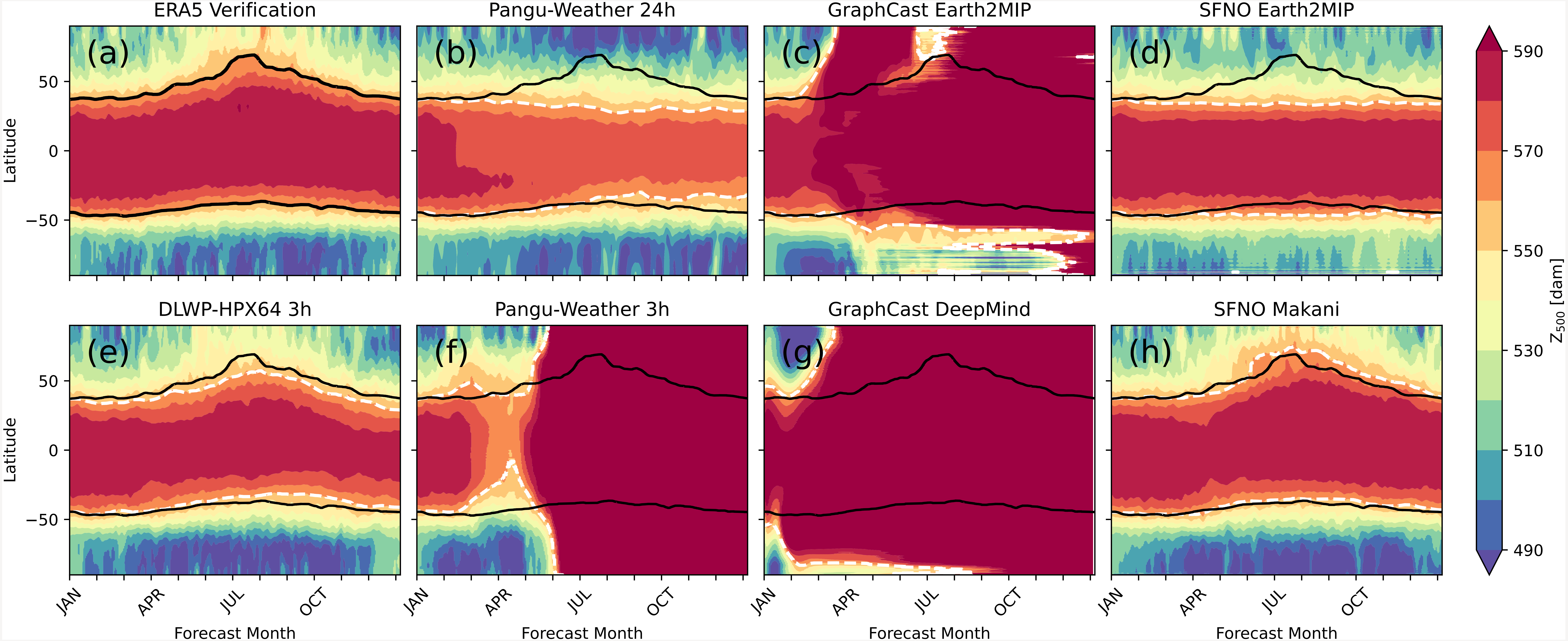
ERA5

560 dam contour
15-d mean

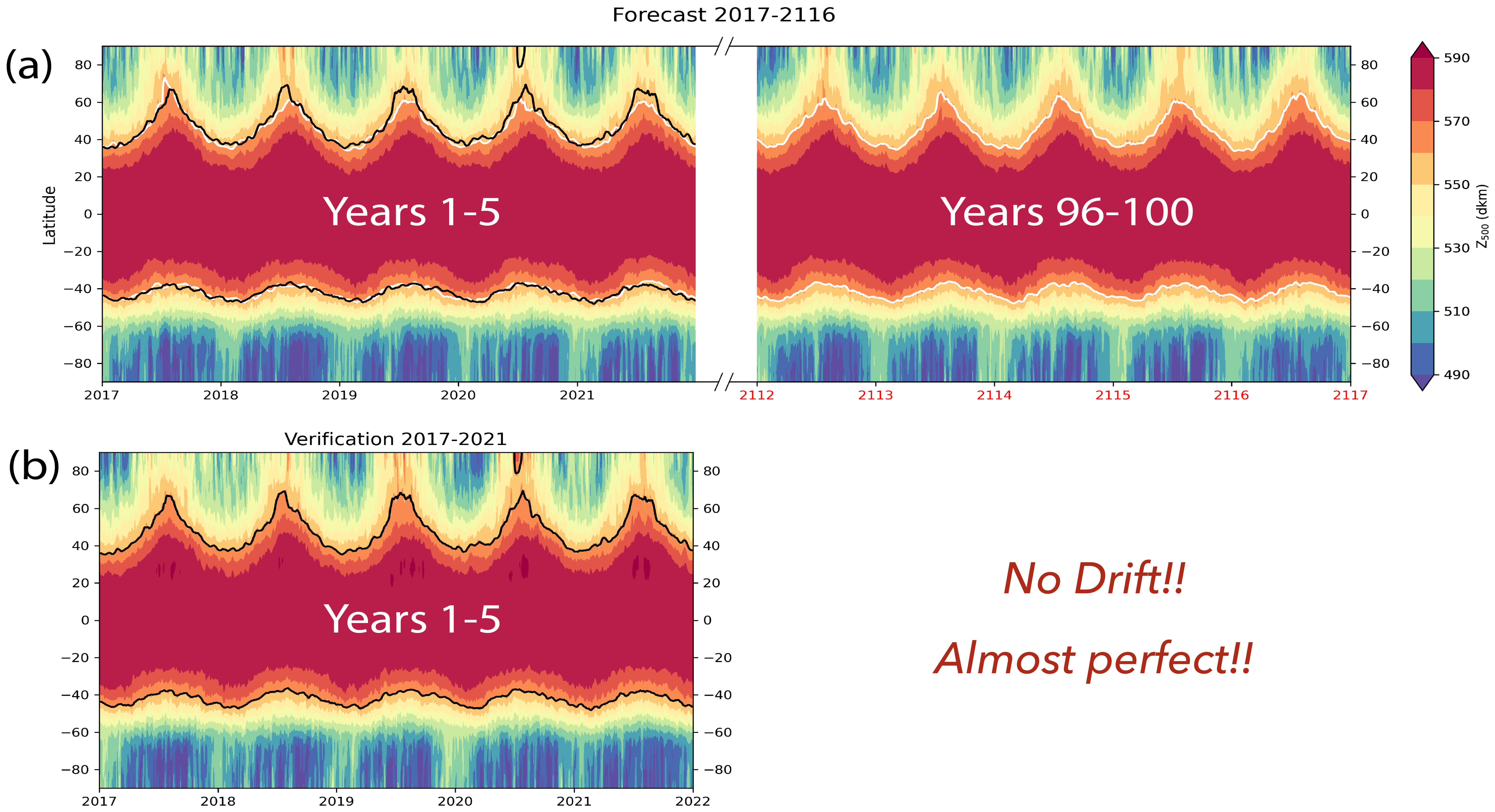


1-Year Rollout: *Atmosphere Only*

Zonally averaged 500 hPa height; 3-d mean



100-Year Rollout: *Coupled Atmosphere-Ocean*



No Drift!!

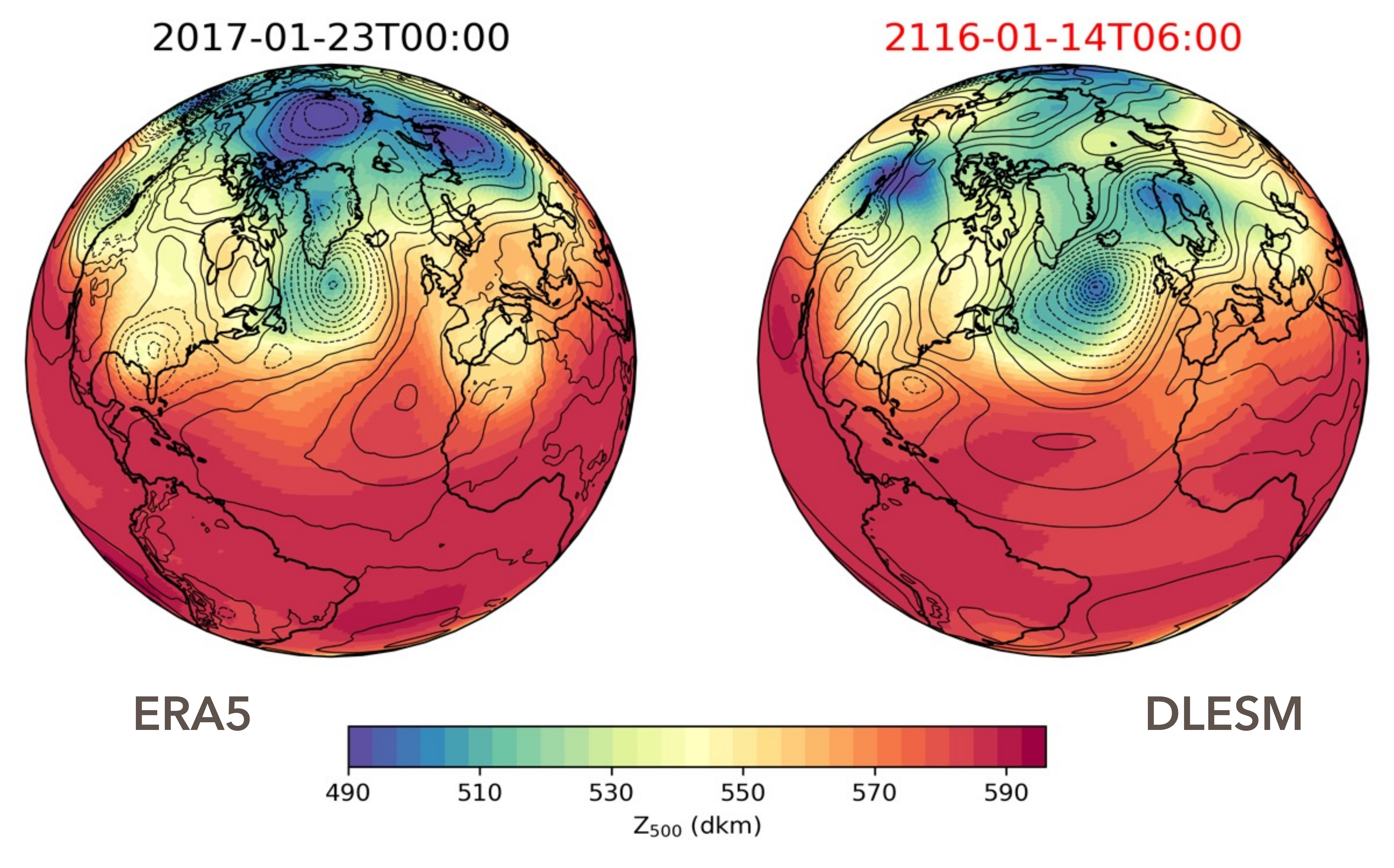
Almost perfect!!

99-Year Rollout: January Cyclone

500-hPa height, SLP

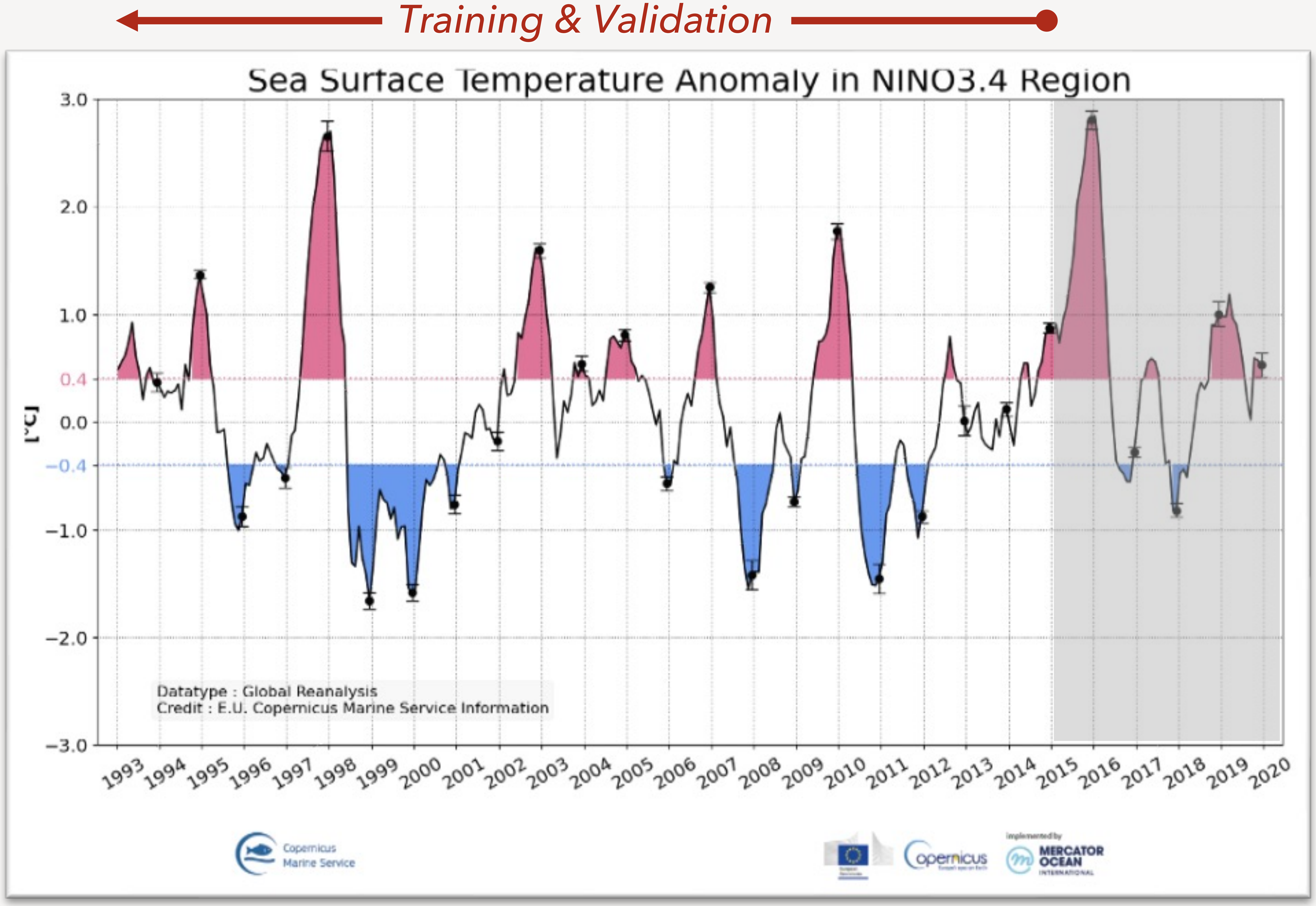
220 km HPX 32

1°×1° ERA5 example



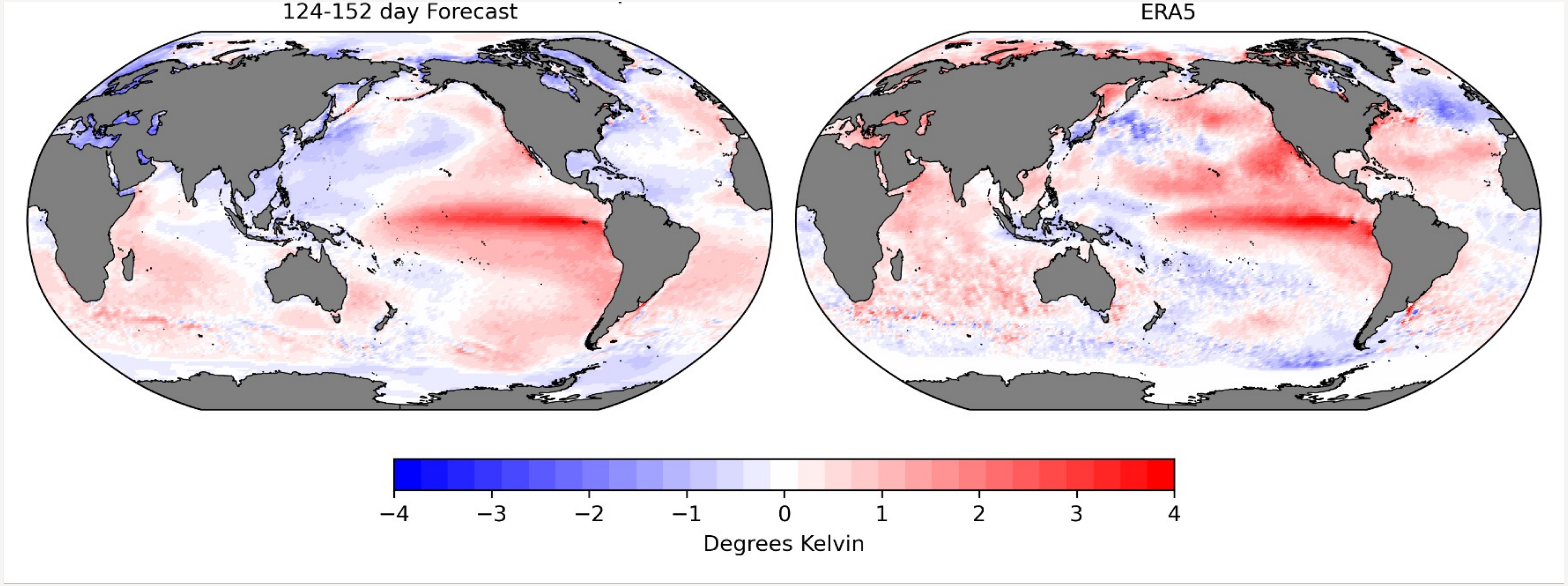
2015-2016 El Niño Forecast

Training loss:
8-day RMSE



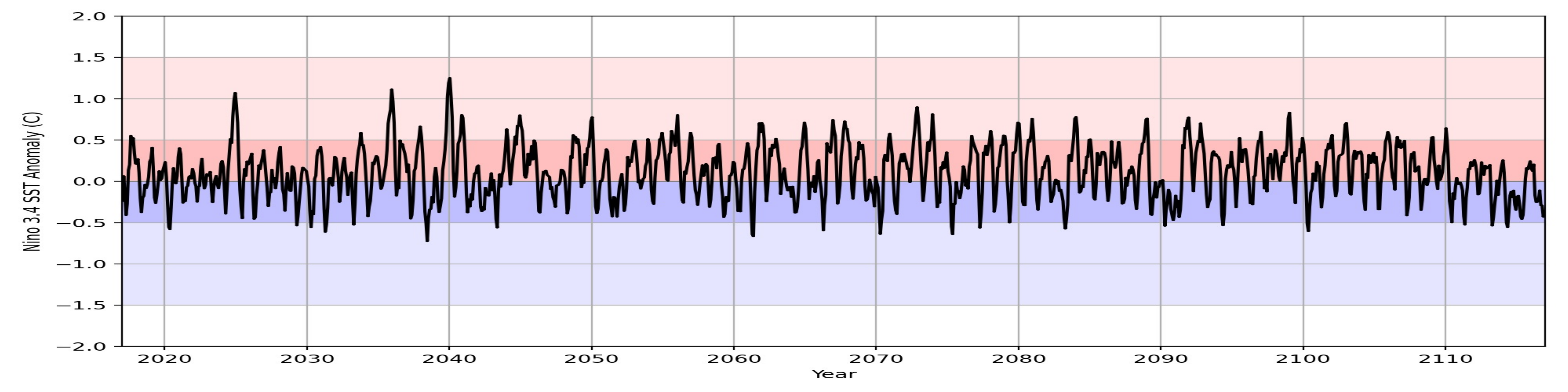
May 1st Rollout: Best Ensemble Member

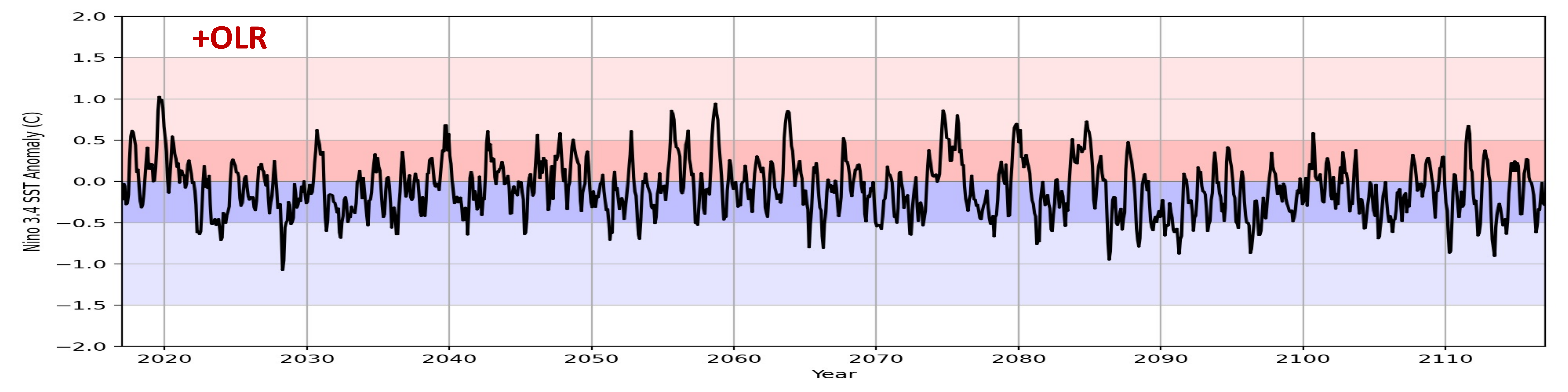
September 2015 Averaged SST Anomalies



SST Anomalies,100-year Rollout: Niño 3.4 Index

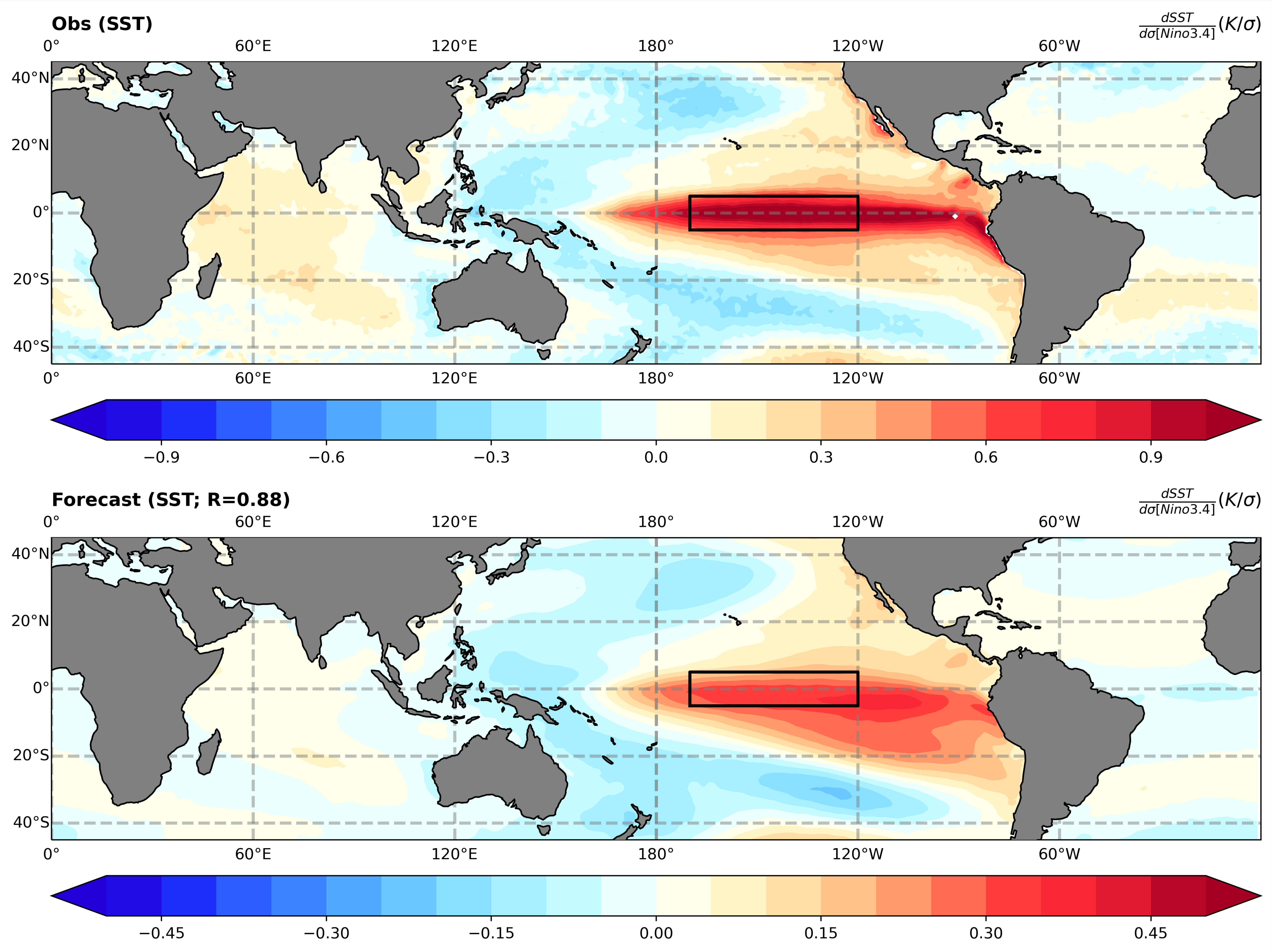
- *No Drift!*
- *Approximately correct period*
- *Amplitude too low (220 km cell spacing)*





SST Anomalies Regressed on Niño 3.4 Index

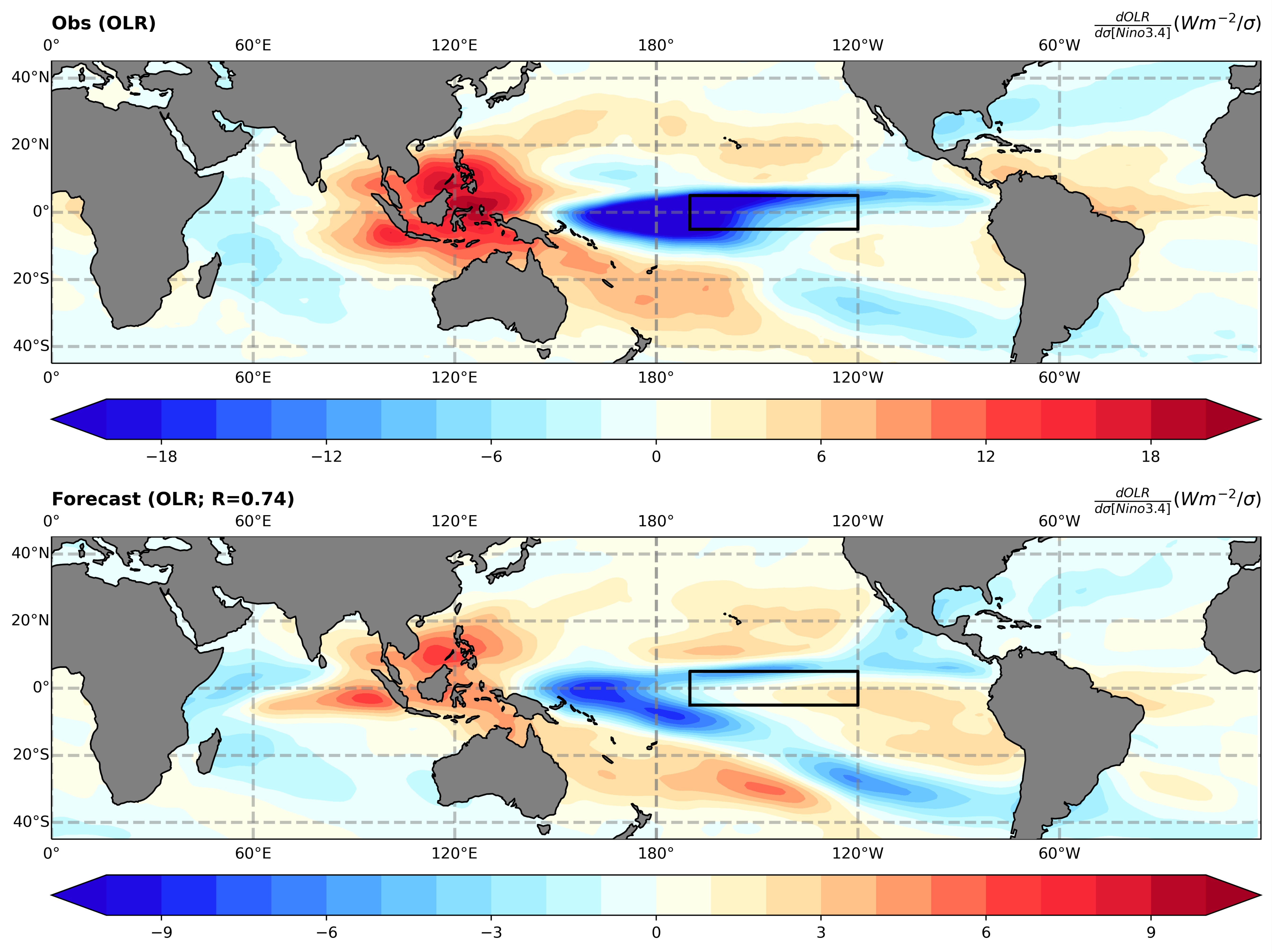
- 100-year rollout
- Includes OLR forcing
- R=0.88



OLR Anomalies Regressed on Niño 3.4 Index

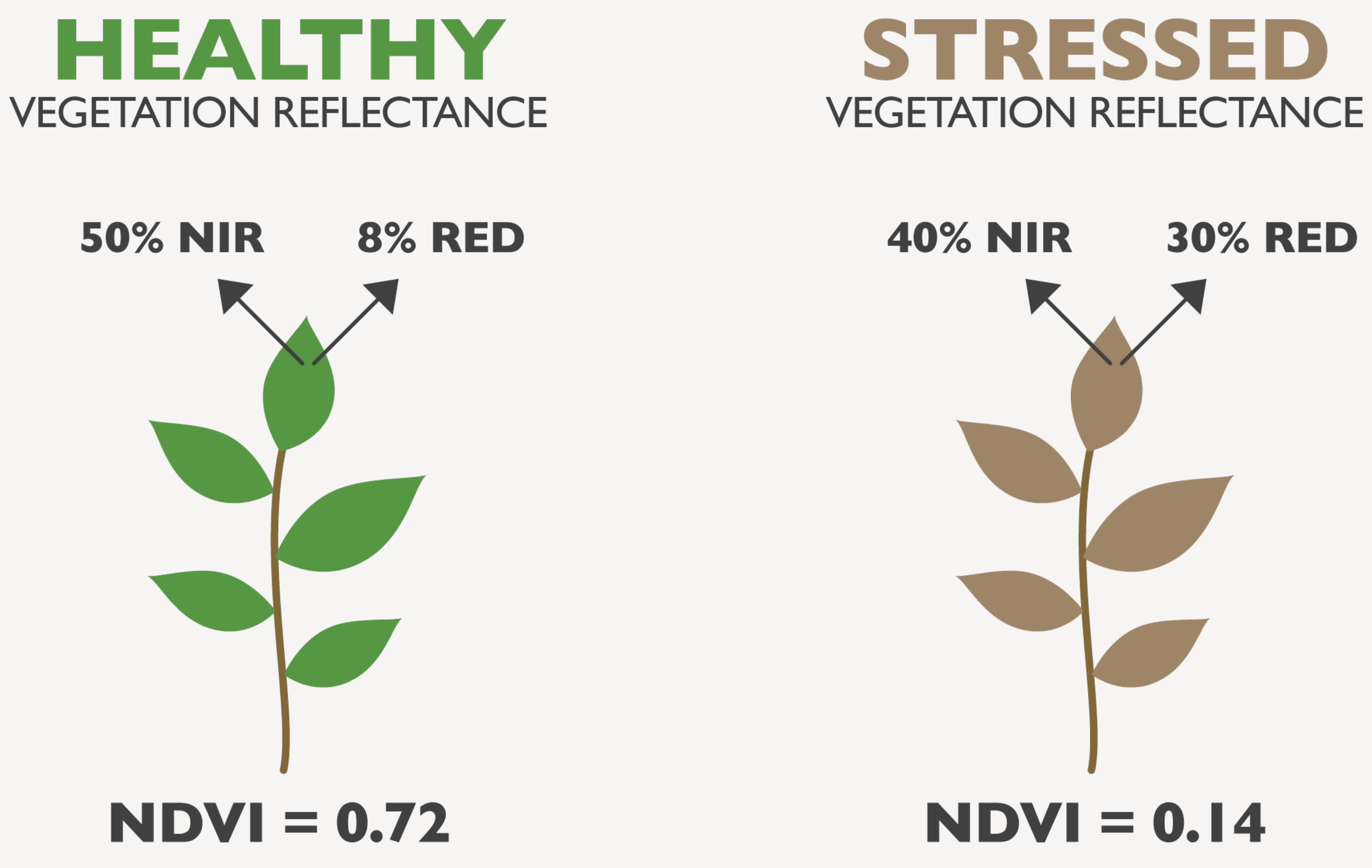
- 100-year rollout
- Includes OLR forcing
- R=0.74

Satellite data, not ERA5



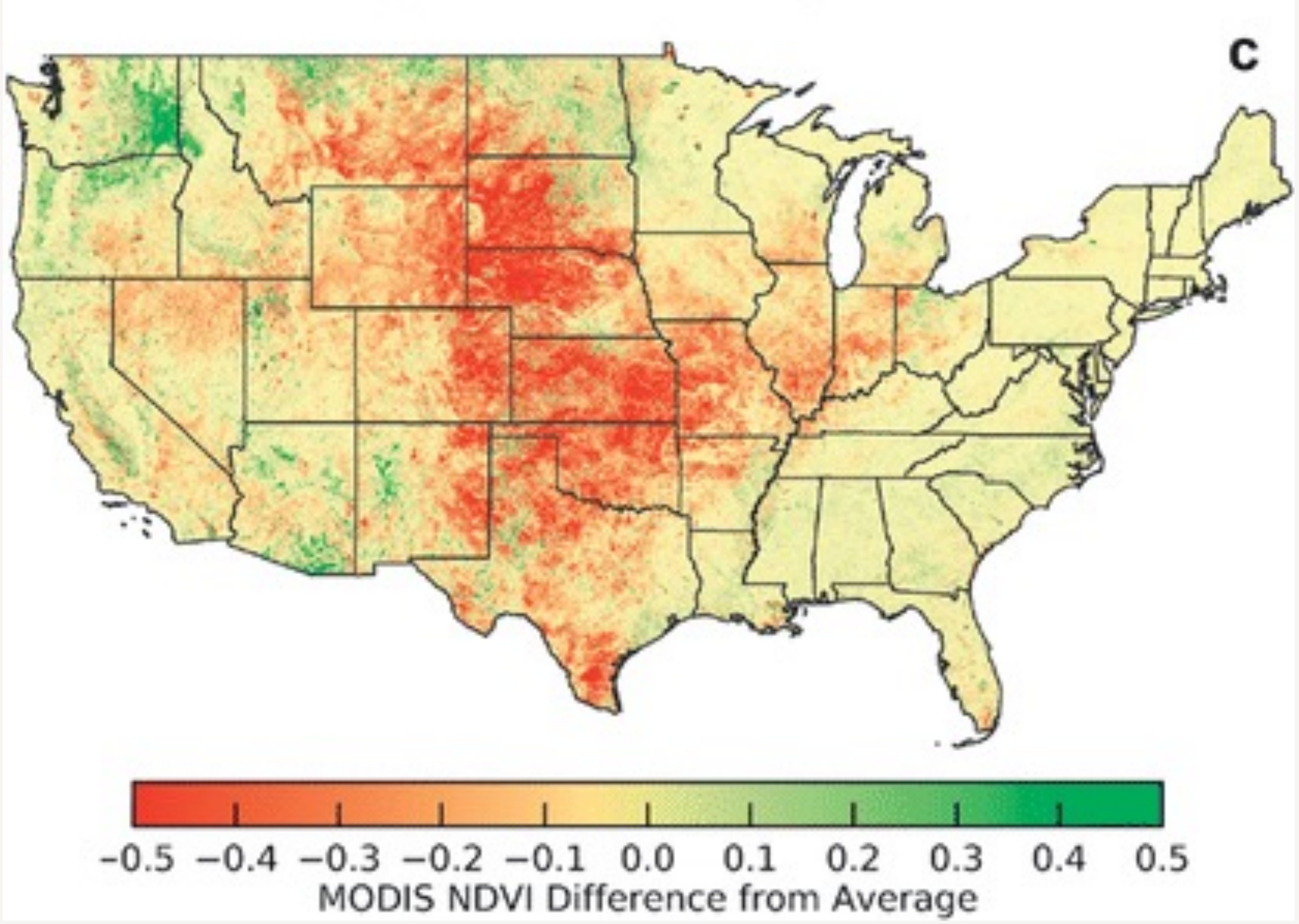
Novel satellite data for DL Earth System Model

Normalized differential vegetation index



$$\text{NDVI} = \frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED}}$$

2012 Great Plains drought



Satellite remote sensing is the future!

Conclusions

- **Deep learning weather prediction models**
  - Can sidestep parameterizations crucial for NWP
  - Yield stable 1-year rollouts at 3-h time resolution with minimal drift
- **Coupling to a deep learning ocean model**
  - Stabilizes iterative rollouts to at least 100 years
  - Removes drift and improves climatology
  - Captures many features of El Niño/ENSO
    - Further improved by adding OLR field from ***satellite observations***
- **Suggests optimism for expanding to a full earth-system model**
- **Parsimonious DLWP model trains for 4 A100s in 3.5 days**

HEALPix DLWP Reference

manuscript submitted to *Journal of Advances in Modeling Earth Systems (JAMES)*

Advancing Parsimonious Deep Learning Weather Prediction using the HEALPix Mesh

Matthias Karlbauer[1], Nathaniel Cresswell-Clay[2], Dale R. Durran[2], Raul A. Moreno[2], Thorsten Kurth[3], Boris Bonev[3], Noah Brenowitz[4], and Martin V. Butz[1]

[1]Neuro-Cognitive Modeling Group, Department of Computer Science, University of Tübingen, Tübingen, Germany
[2]Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA
[3]NVIDIA Switzerland AG, Zürich, Switzerland
[4]NVIDIA Corporation, Seattle, USA

essopenarchive.org/doi/full/10.22541/essoar.169603505.58030377