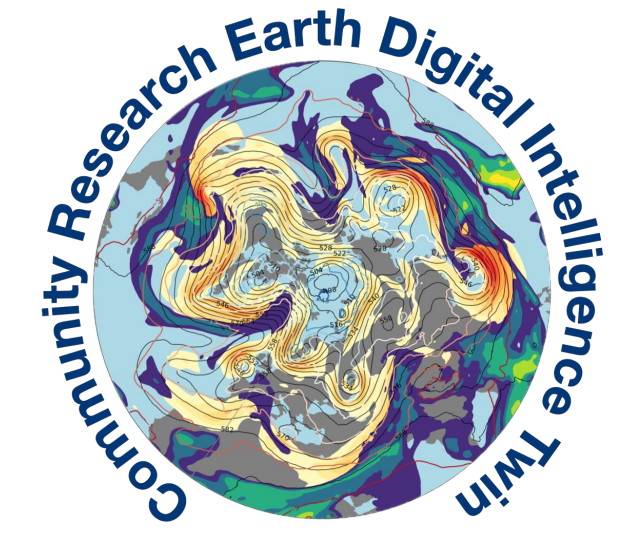


# WIP – CREDIT–MESACLIP: stable emulation of high resolution coupled models



Aidan Janney<sup>1</sup>, John Schreck<sup>2</sup>, David John Gagne<sup>2</sup>, Frederic Castruccio<sup>1</sup>, E. E. Hayden<sup>1</sup>, Dan Amrhein<sup>1,2</sup>

<sup>1</sup>CGD Labratory – NSF NCAR, <sup>2</sup>Computational and Information Systems Laboratory – NSF NCAR



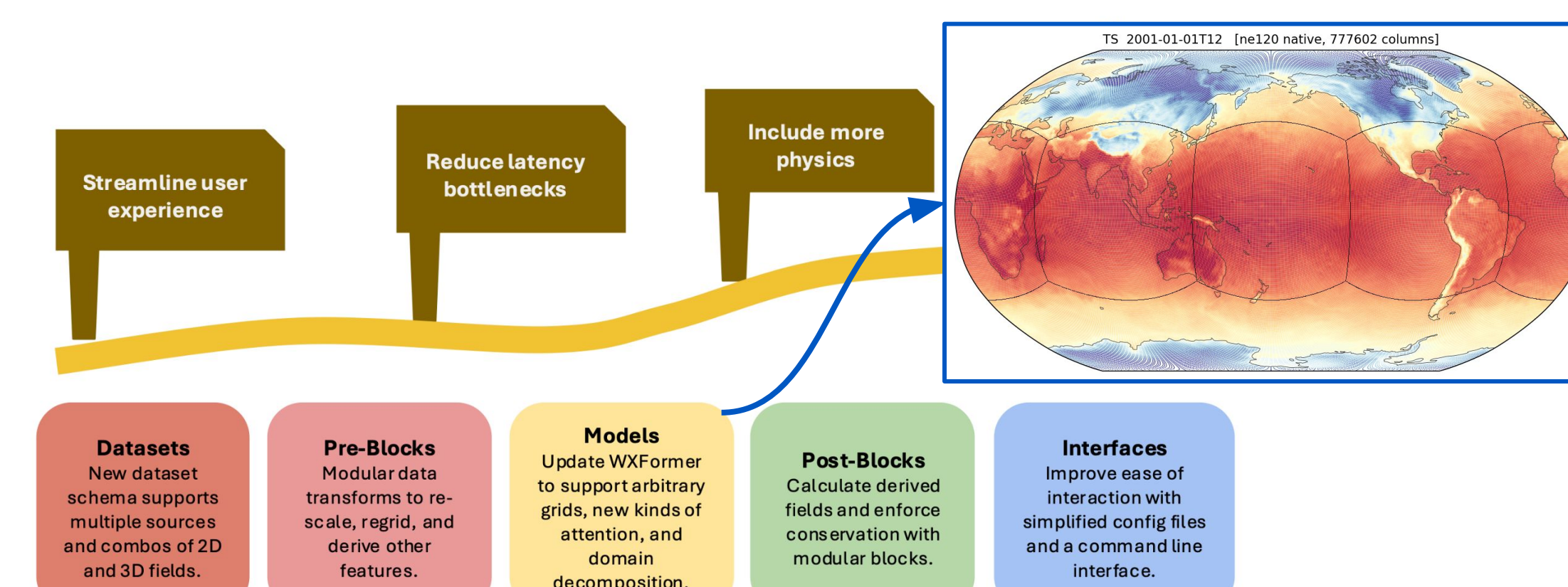
## Background

### MESACLIP iHESP m010 – GDEX Dataset d651007

Training Data: 1920 to 2000  
Validation Data: 2000 to 2005

Attempting 84 Variables

- 7 – 3D variables (all ATM): vvel, spec. humidity, rel. humidity, temp, U, V, geopot. height
- 77 – 2D Variables (ATM/LND/OCN): see additional resources



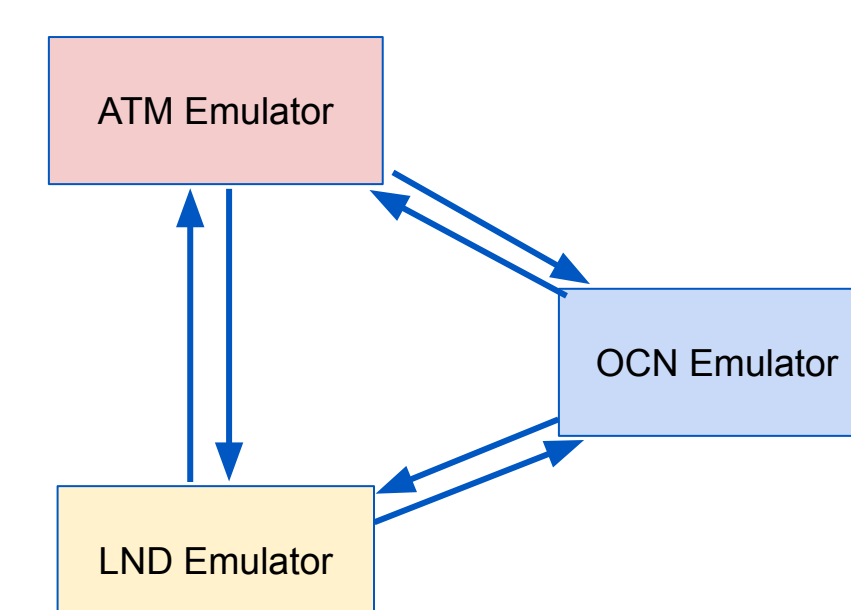
CREDIT Gen2 Architecture with **Cube Sphere Input** – Schreck 2026

## Battle of Scales

Current emulator structure predicts **one state at one frequency**, but we have conflicting priorities in the atmosphere and ocean.

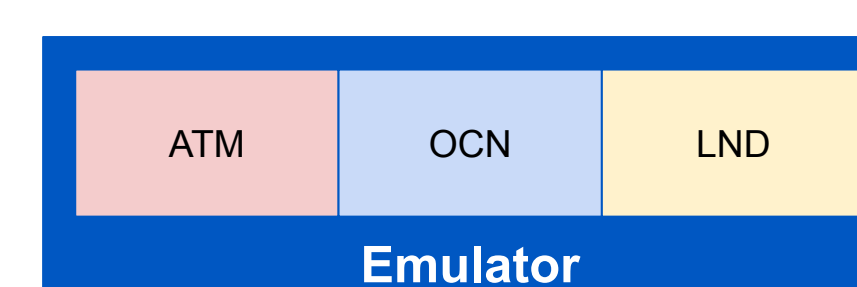
- Atmosphere: large scale, synoptic **O(100–1000km)** on **6h timescales** – highly chaotic, diurnal cycle is important, needs large NN even at lower resolutions
- Ocean: (sub)mesoscale features **O(<100km)** on **daily timescales or longer**. Interactions with bathymetry and LONG timescales are important for stability.

### Traditional Coupling



- Variable frequency
- Selected variables
- Learned or Deterministic?
- Biased by existing model development paradigms

### Implicit Coupling



Give the model all the data at once  
Learn relationships of the **continuous system**

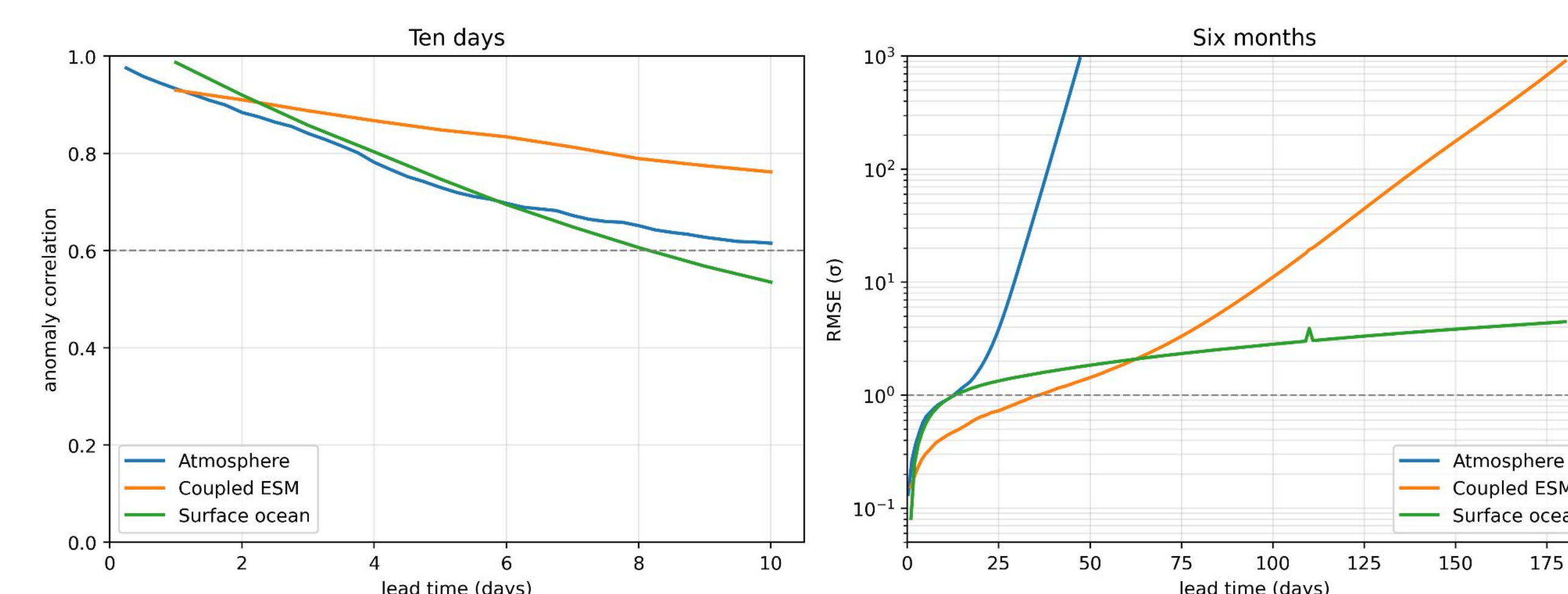
- Shared timescale
  - Aggregate and Interpolate
- Flexible and novel coupling terms
- Differing spatial scales and grids: an architecture issue
  - Band-aid: interpolate and hope hi-res info persists in large scale dynamics

## Component Comparison

Training defaults – singlestep, levelwise normalization

- ATM only – forced by surface OCN/LND
- Surface OCN/LND – forces by lower atmosphere
- Coupled ESM: 3D ATM + 2D LND+ 2D OCN – no forcing, implicit coupling

**Undertrained and in-need of hyperparameter changes, some 10-day skill!**

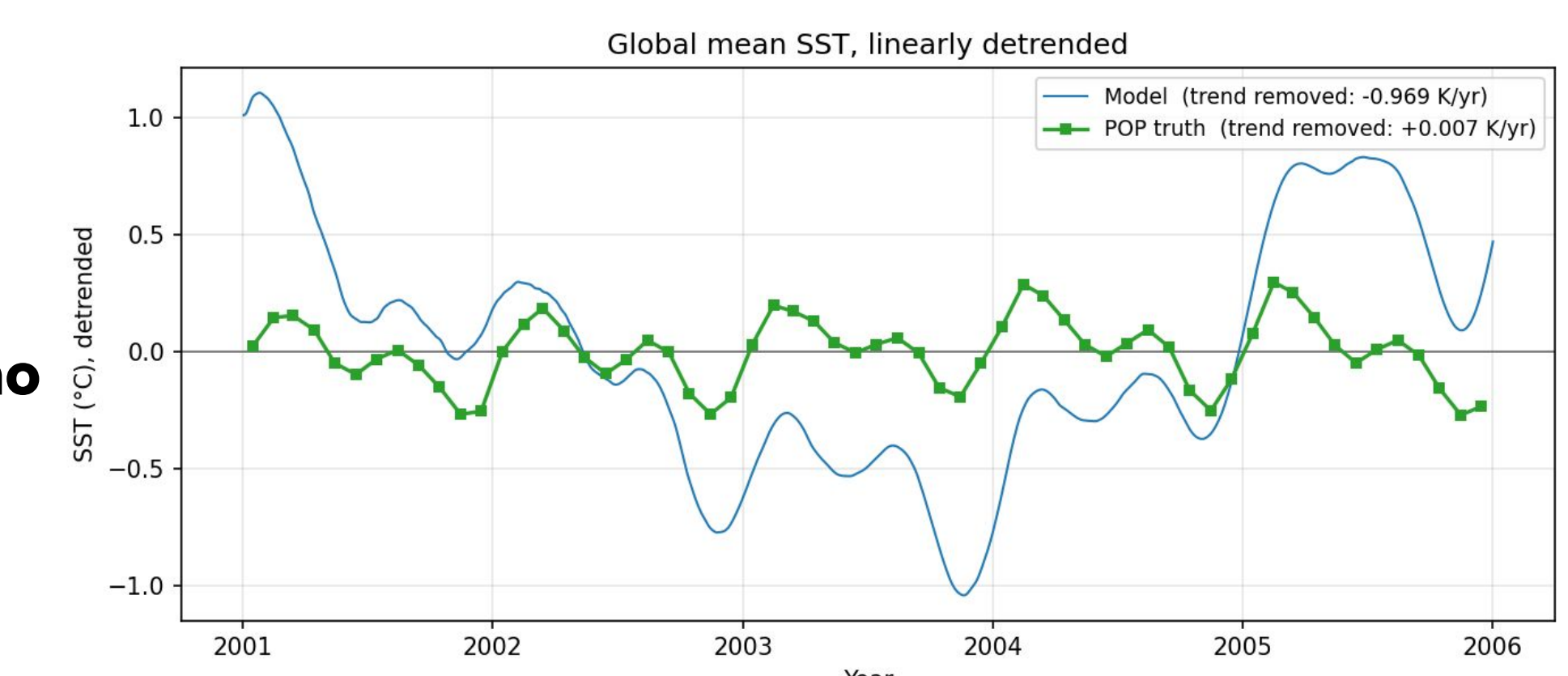


## Teaser: Hi-Res Ocean Emulation

Stable Rollouts at >5 years at 1/10° – highest resolution global ocean emulation to date.

Drifting (linearly) towards it's own attractor. We could constrain with climatology (Samudra)?

Free running ocean **no forcing!**



## Next Steps – More Training!!

**Model Architecture:** refine cube sphere processing, include multi-grid handling – significant

Shared latent space

Flux constraints

**Compressing Data:** reduce volume while retaining info

PCA/EOF, SOM, autoencoding

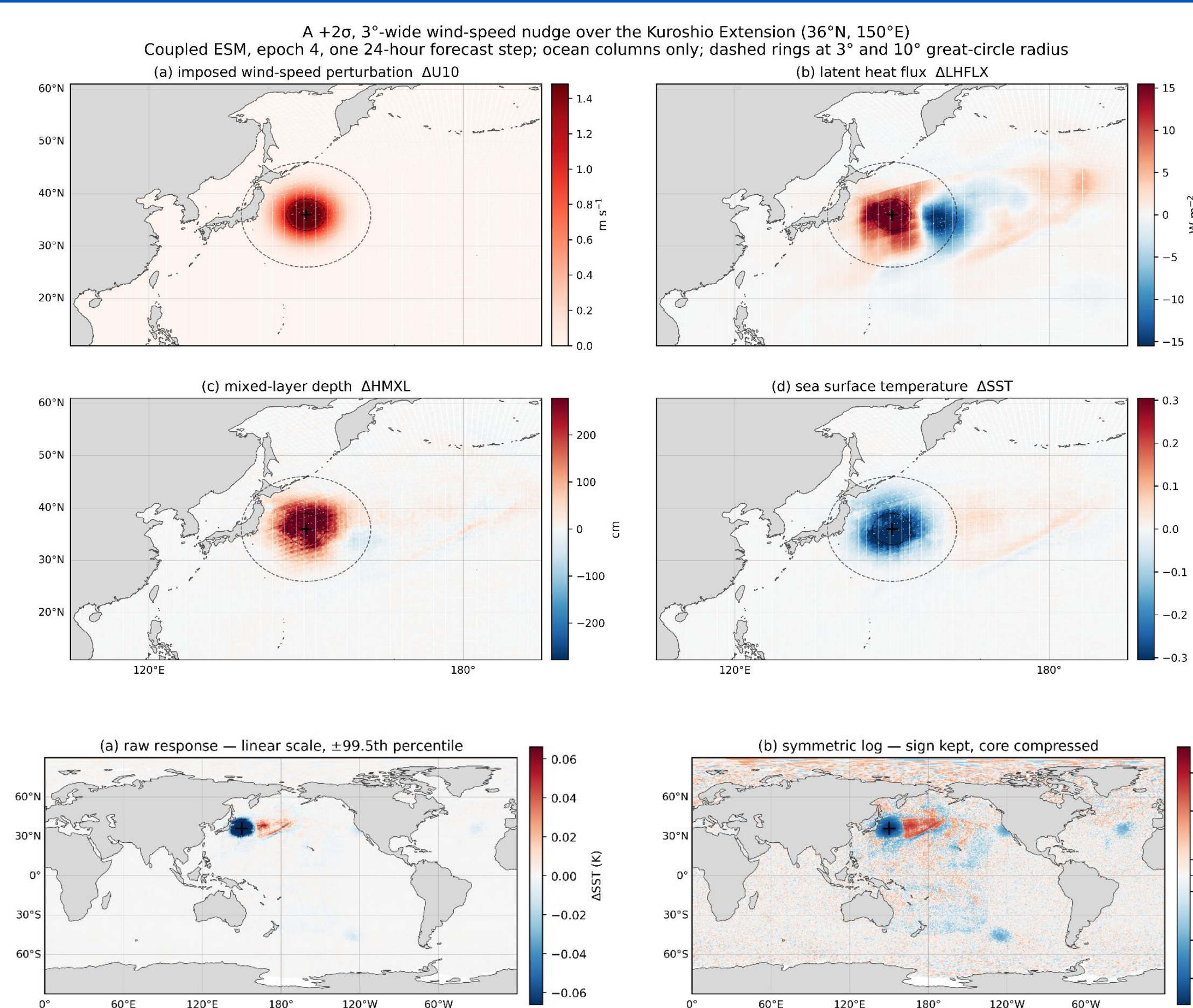
**Tiered model hierarchy:** reflect innate timescale variability of the Earth System, traditional coupling

Monthly ocn, daily sfc, hourly atm

MLe coupling extraction, ensembles in latent space, **data**

Diffusion for interpolation and DA?

## Poke at the Model – Test for Physical Responses



Physically realistic response from variables, but **generic**

Primarily **diffusive spreading**.

Global noise and teleconnections after only 1 time-step. Follows cube faces.

## Additional Resources

This work was funded by the **CGD Strategic Initiative Fund 1090941**.

