Reliable AI-Driven Climate Modeling for more Trustworthy Climate Projection

Christopher S. Bretherton
Senior Director of Climate Modeling, Ai2
(and the Ai2 CM team)

External partners


Spencer Clark


Brian Henn


Anna Kwa


Jeremy McGibbon


Andre Perkins


Oli Watt-Meyer


Elynn Wu


James Duncan


Lucas Harris


Peter Caldwell

Ai2

Physically-based weather and climate modeling

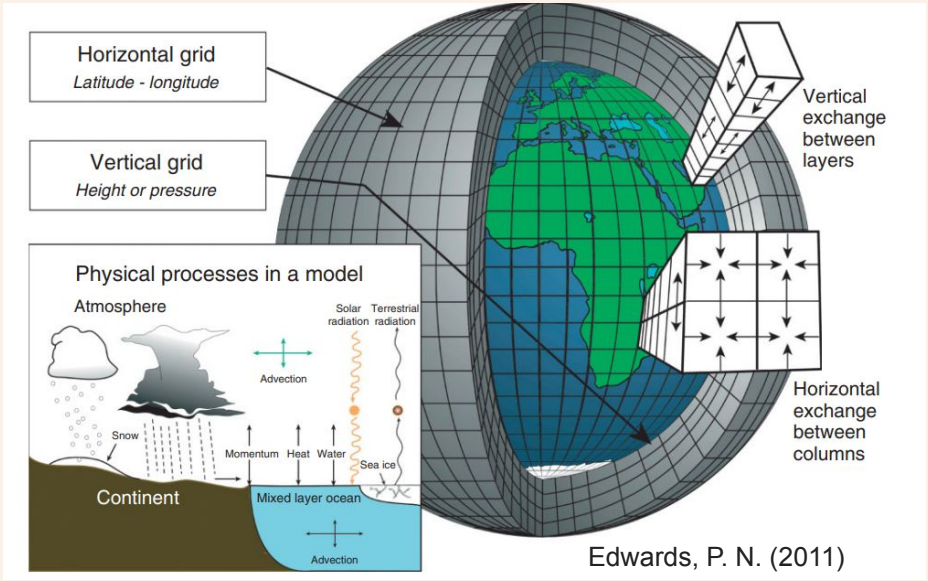
Grounded in plausible equations, improving generalizability

But uncertain parameters affect results

Systematic biases/uncertainties, e.g.

- Climate sensitivity/ cloud feedback
- Regional precipitation trends
- Ecosystem change/ carbon uptake

Computationally heavy



Edwards, P. N. (2011)



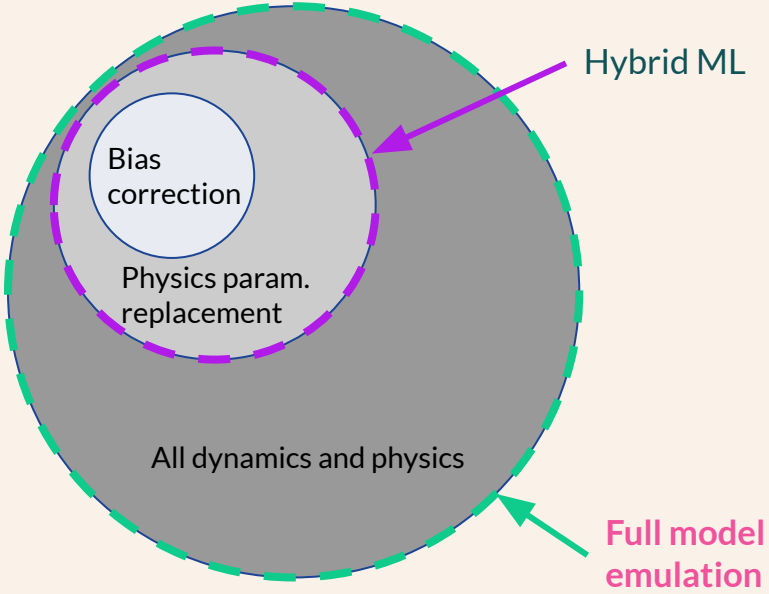


How can ML help weather/climate models?

Make global weather and climate models:

- more accurate (less systematic bias)
- faster
- more affordable/accessible

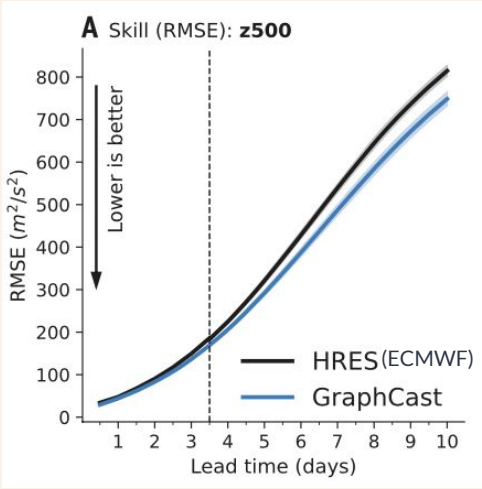
What is the ML responsible for?



ML-based weather forecasts are state of the art

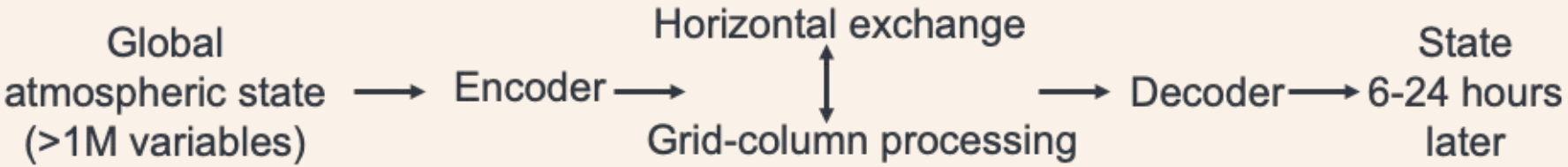
- **2019–2022:** pioneering work showing potential for ML-based global weather forecasting trained on reanalysis (e.g. *Dueben and Bauer*, *Weyn et al.*)
- **2022–2023:** Pangu, GraphCast, etc. demonstrate improved 3-10 day deterministic weather prediction over operational benchmarks
- **2024–:** development of probabilistic (e.g. GenCast) and coupled ocean-atmosphere forecast models (e.g. ORCA, OLA)





GraphCast: Lam et al., 2023 (*Science*)

Diverse ML architectures used for weather forecasts



- 10^6 - 10^9 learnable weights; almost no baked-in physical priors
- Train on 40 years of ERA5 in days to weeks
- Little or no representation of land, ocean, ice evolution
- 10-day forecast on a single GPU in seconds
- A few architectures can produce stable forecasts of a year or longer

AI-based climate vs. weather simulation

1. Future climates are unseen
2. Climate is a coupled earth system process with diverse time scales
3. AI needn't guarantee conservation principles or correct information flow
4. Is there any framework for gaining confidence in AI-based climate projections?

Gain perspective by simplifying to AI emulation of an existing global climate model from limited training data that spans relevant range of climates, using simplified CMIP-style tests [See also *Ullrich et al. 2024*].

An ideal emulator would give results indistinguishable from reference GCM for all CMIP cases and emission scenarios.

Our experience with ACE (Ai2 Climate Emulator)

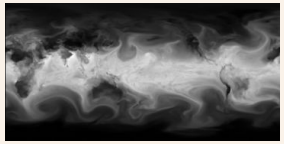
- ACE: ML emulator of an excellent physically-based global climate model for 100-10000x speedup of routine simulations, ensembles, downscaling…
- Derive climate as statistics of simulated weather, like physics-based GCMs
- Train on reference climate model output from diverse range of climates
- Boundary conditions (insolation, SSTs) provided as inputs to ML model
- Show results including coupling to a slab ocean model with perturbed CO_2
- Focus on climate accuracy and lessons from out-of-sample testing
- More on version 2 of ACE: Henn et al. 11A.2 Wed 2 pm

Training Setup

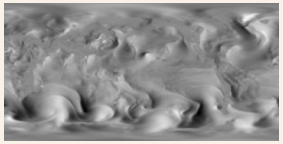
Loss function: mean-squared error of 6-hour forecast computed over all the output variables (for ACE2, accumulated over 2 forward steps)

Input at time *t*

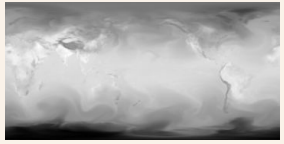
Humidity



Wind



Temperature

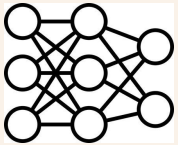


Surface pressure



(**forcing** + **prognostic** vars) x 180 x 360

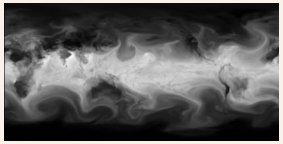
SFNO*



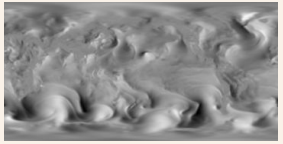
$$f(\mathbf{u};\theta)$$

Prediction of *t* + 6 hours

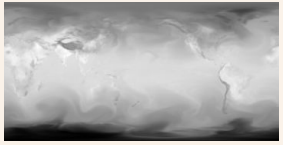
Humidity



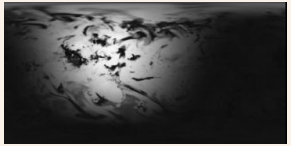
Wind



Temperature



Down SW sfc rad flux



(**prognostic** + **diagnostic** vars) x 180 x 360

Showing subset of inputs/outputs

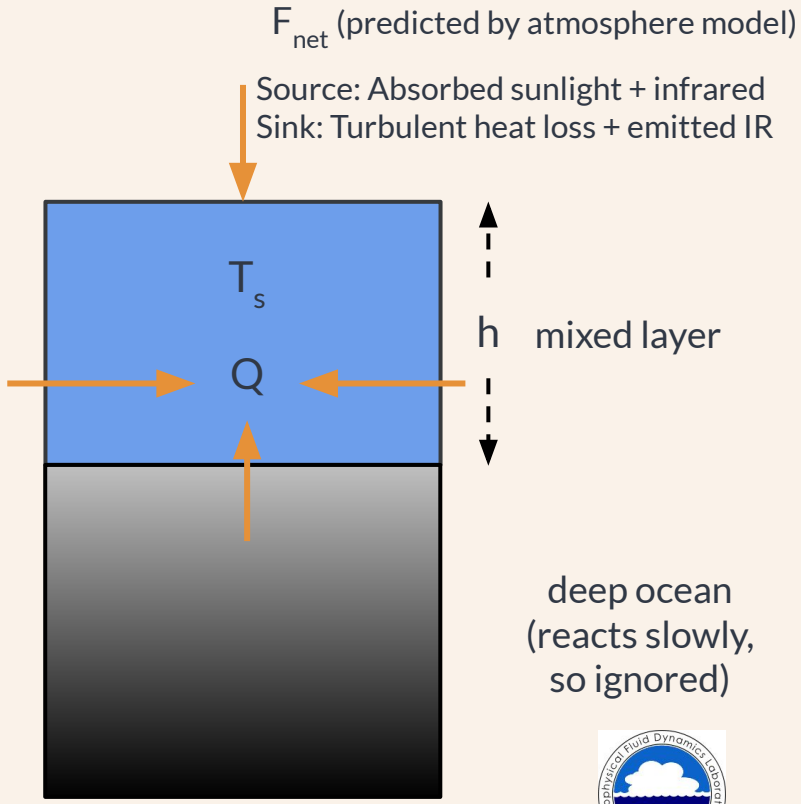
*__SFNO__: Spherical Fourier Neural Operator from Bonev et al., 2023 arxiv.org/abs/2306.03838

Ai2

NVIDIA

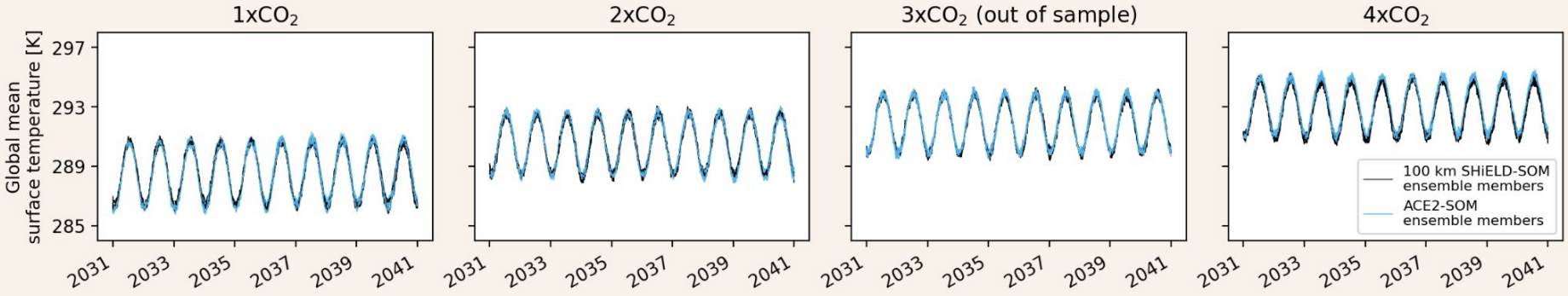
ACE2 + interactive slab ocean (ACE2-SOM)

- Training/validation data from three 50-year 100 km SHiELD simulations:
  - 1xCO_2
  - 2xCO_2
  - 4xCO_2
- Train 1°x1° ACE-SOM with CO_2 as input.
- Training: 3 days on 8 H100 GPUs
- Inference: 1600 years/day on 1 H100 GPU (~100x faster than SHiELD)
- *Clark et al. 2024*, *JGR-ML*, submitted



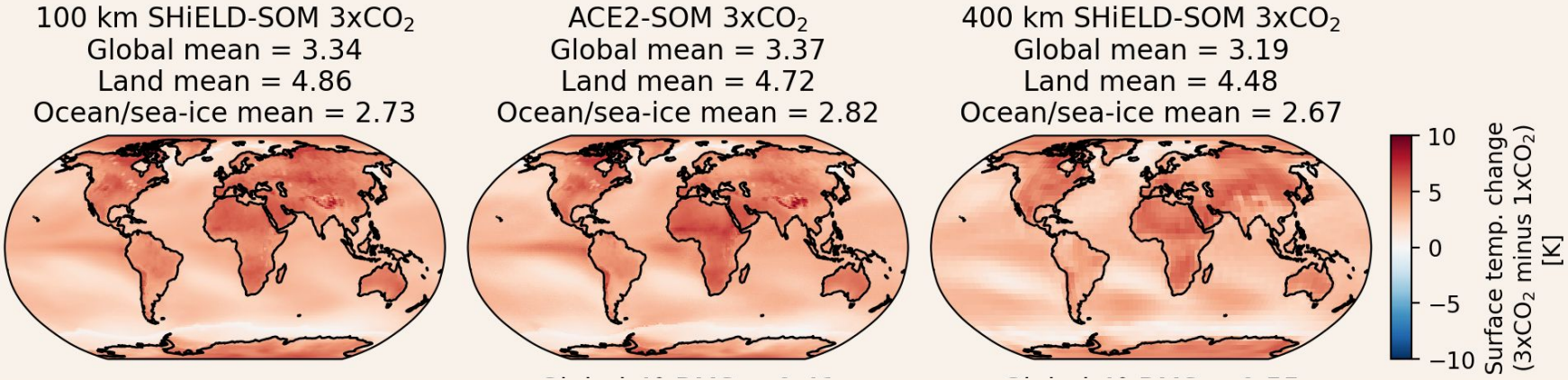
Ai2

ACE2-SOM is stable in both in-sample and out-of-sample equilibrium climates



- Results are from five-member initial condition ensembles of 10-year 100 km SHiELD-SOM (target model) and ACE2-SOM simulations in each climate.
- ACE2-SOM did not see data from the 3xCO_2 climate during training.

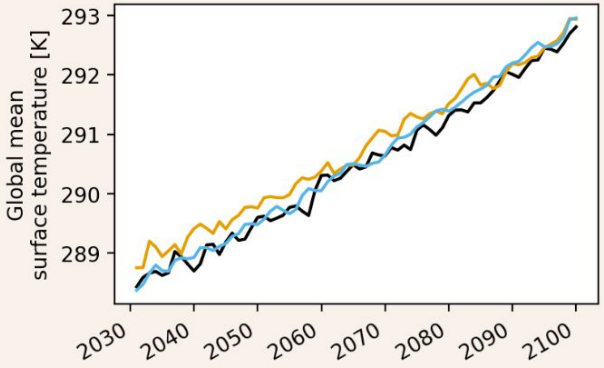
ACE2-SOM is stable and accurate in multiple climates

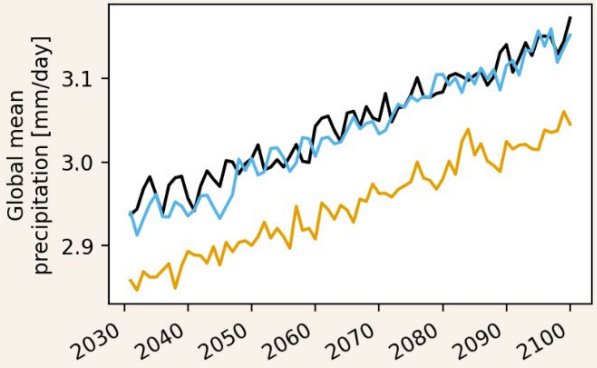


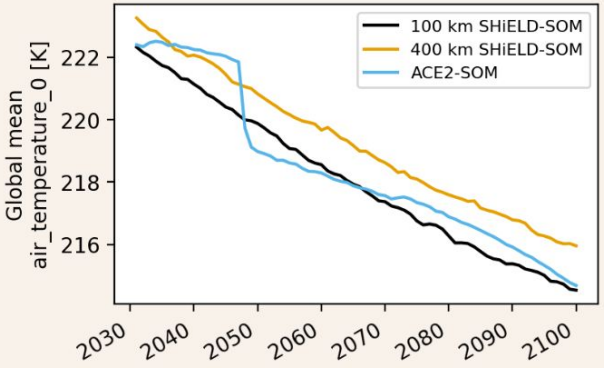
ACE2-SOM global warming pattern matches the physics-based model, capturing robust features like amplified warming over land.

What about some harder generalization tests?

Out of sample test: 2% CO_2 increase per year



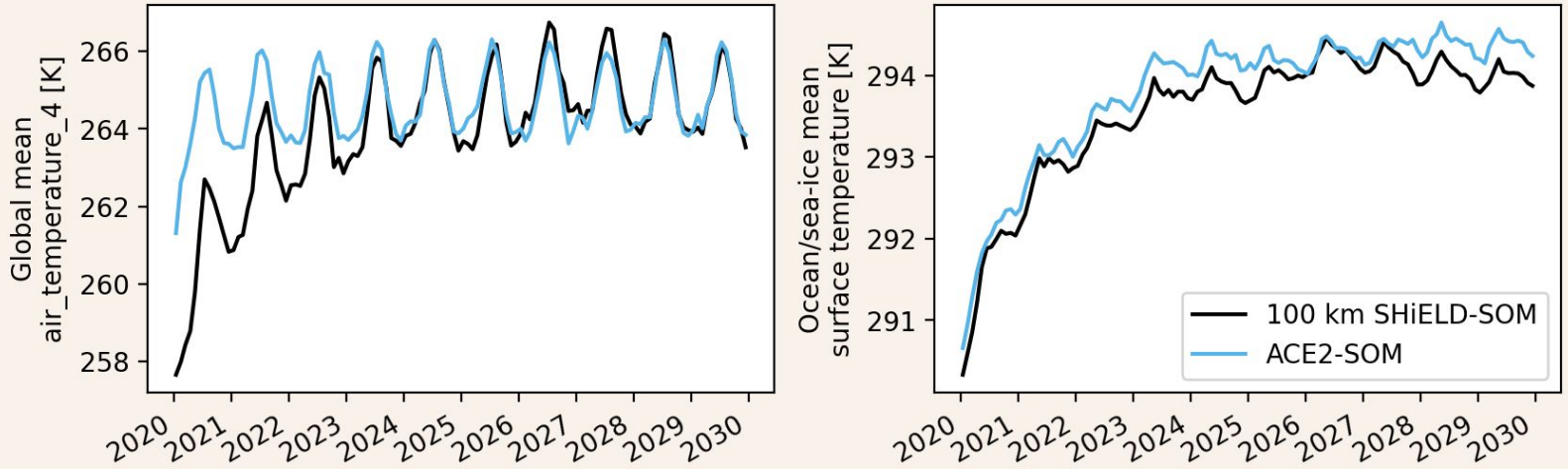




- 70-year rollout ramping from 1xCO_2 to 4xCO_2 concentration
- Surface variables good, but stratospheric temperature has a spurious climate jump
- ML with 6-hr steps trains unreliably since stratosphere evolves slower than troposphere

Ai2

Out of sample test: Abrupt 4xCO_2 increase



- ML-controlled fields (except ocean temperature) unrealistically jump to 4xCO_2 regime. Global atmospheric and land surface energy conservation are violated.
- Ocean temperature, aided by prescribed thermal inertia, warms realistically slowly.
- ACE2-SOM still finds the right steady state, helped by its equilibrium 4xCO_2 training.

AI-based climate vs. weather simulation

1. Future climates are unseen
2. Climate is a coupled earth system process with diverse time scales
3. AI needn’t guarantee conservation principles or correct information flow
4. Is there any framework for gaining confidence in AI-based climate projections?

AI will not magically overcome these issues, but an iterative design process centered around CMIP cases can expose issues and allow architectural modifications that can overcome them.