Improving climate models using corrective machine learning

Anna Kwa, Spencer Clark, Oli Watt-Meyer, Brian Henn, Jeremy McGibbon, Andre Perkins, Chris Bretherton, AI2 Climate Modeling, Seattle, WA
Lucas Harris, GFDL

Webpage: https://allenai.org/climate-modeling
Philanthropic project of the Allen estate, at Vulcan (2019-2021) and now at AI2.

- **Goal:** Halve climate model uncertainty about 21C regional precipitation trends
- **Strategy:** Make coarse-grid climate models better using fine-grid models as reference

Partnered with NOAA/GFDL, developers of a 3-km version of FV3GFS global weather model

Ongoing collaborations with other R&D groups (e.g. NVIDIA, LLNL) and summer interns
How can ML help weather and climate models?

Make global weather and climate models:
- accurate
- faster
- affordable

Three general strategies:
- Hybrid: Replace or correct parts of the climate model, e.g., physics parameterizations
- Full model emulation (FME): ML of entire global atmospheric evolution
- Flexible, nonlinear bias correction
Corrective ML to improve coarse-model simulations

Climate model (25-200 km)

Train corrective ML to make temperature, humidity, winds of the coarse model track reference data.

High fidelity reference: observations or fine-grid (3 km) simulation

3-hrly nudging tendencies are interpreted as a needed correction to the physical parameterizations and coarse-grid dynamics.
Fine-grid reference model: accurate simulations across a range of climates

3 km grid gives a better rainfall simulation over land than 200 km:
- Enabled by explicit simulation of cumulonimbus clouds & well-resolved mountains
- The 3 km model resolves variability that requires subgrid parameterization in GCMs

3 km model is expensive & imperfect but enables 1+ yr simulations in multiple climates
‘Nudge to fine’ ML methodology

Features:
- Air temp
- Vertical profile
- Specific humidity
- Vertical profile
- Additional potential features, ex. insolation, land/sea mask, surface elevation

Predictions:
- Heating vertical profile
- Moistening vertical profile
- Surface downward radiation

Diagram showing input, hidden, and output layers.
‘Nudge to fine’: ML offline evaluation

Nudging tendencies show systematic errors of coarse model vs. reference.

The ML schemes produce a smoothed, lower-amplitude, unbiased version of the noisy nudging tendencies.

Bretherton et al. 2022

Watt-Meyer et al. 2021
Challenge of hybrid ML coupled to other components

Coupled to fluid dynamics and parameterized physics

- ML
- Fluid Dynamics

Training ≠ Testing (offline) (online)

Online ML Skill

- Low Climate Bias
- Accurate weather forecasts
- Stable Simulations
- Offline Skill

Accurate weather forecasts
Corrective ML results (200 km coarse, 3 km fine)

Presented to POS panel in 10/2022 by Oli Watt-Meyer:

- ML corrections trained using reanalysis reference (Watt-Meyer et al. 2021) add a day of forecast skill and improve mean surface precipitation patterns over land 20%

- ML corrections trained with 3 km GSRM reference improve land surface precipitation 20-30% (Bretherton et al. 2022, Kwa et al. 2023) and retain multiyear stability

- ML corrections to 200 km coarse model trained using 25 km AGCM reference across a range of climates specified using -4K, 0, 4K and 8K SST increments improve surface temperature and precip patterns by 10-30% in all climates (Clark et al. 2022)
Further climate bias reductions are ongoing

- With ML correction of u&v as well as T&q, annual-mean pattern biases of baseline are reduced 50% for surface temperature and 30% for surface precipitation.

**ML-corrected bias**
- Mean: -0.027 K, RMSE: 1.196 K
- Mean: 0.056 mm/day, RMSE: 1.265 mm/day

**No-ML baseline bias**
- Mean: -0.329 K, RMSE: 2.458 K
- Mean: 0.123 mm/day, RMSE: 1.781 mm/day

Further improvements attainable using out-of-sample detection (Sanford et al. 2023) and better coarsening algorithms.
Can we improve on nudging-based corrective ML for climate?

- Corrective ML trained on nudging tendencies improve forecast skill and land surface precipitation/temperature climatology of a coarse-grid global climate model by O(50%)
- It is physically interpretable and obeys mass, heat, and moisture conservation
- To make a really attractive emulator of a reference model, we must further reduce its climatological biases vs. reference target. This is challenging using our corrective ML
- We are thus investigating promising new approaches:
  - FME (e.g. FourCastNet, Pangu, GraphCast, ClimaX) - we have stable 10 year simulations
  - Hybrid reservoir computing (Arcomano et al. 2022, 2023)
  - CycleGAN post-processing of coarse model precipitation fields (McGibbon et al. 2023)
CycleGAN precipitation morphing (200 km -> 25 km)

Mean annual precipitation bias almost perfectly removed across range of climates

Zhu et al. 2017

McGibbon et al. 2023
Outlook

Hybrid:
- Can substantially improve climate relative to a reference
- Can be made fast
- Harder to efficiently train
- Easier to diagnose and enforce conservation laws

FME:
- Transformationally fast; better use of ML
- Skillful for weather
- Climate accuracy/stability in progress

ML (e.g. CycleGAN) is also a powerful tool for post-hoc bias correction

These tools may revolutionize applications-oriented climate modeling