Why are we doing this?

Western North America temperature

IPCC AR6 Atlas (CMIP6 models)
Why are we doing this?

Western North America **temperature** and **precipitation**

Models have **less agreement** about future local precipitation trends compared to temperature. This matters!
ML Goal: Improve coarse-model simulations

High fidelity reference
reanalysis or
fine-grid (~3km) simulation

Use machine learning to make coarse model behave more like reference

Climate model (25-200 km)
Challenge of 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
Previous work: replacing parameterizations

Coarse-grain a global storm-resolving model and learn apparent sources:

- Conceptually appealing, but there are challenges with coarse-graining especially in presence of topography

Emulate embedded cloud-resolving models (SP-CAM):

- Clean interface between large and small scales—a very well posed ML problem
- **Aquaplanet**: Gentine et al. (2018), Rasp et al. (2018)
- **Realistic geography**: Han et al. (2020), Wang et al. (2022)

These approaches have the ML replace existing parameterizations and/or fully represent convection+turbulence (and in some cases radiation).
Corrective machine learning approach

- Online bias correction has a long history (e.g. Leith 1978, DelSole et al. 2008)
- Our approach:
  1. Nudge coarse-resolution model towards some reference dataset
  2. Train ML to predict nudging tendencies using coarse-res model state as input
  3. Run coarse-res model again, with ML predicting corrective tendency at every time step

\[ \Delta Q_a = -\frac{a - a_{\text{ref}}}{\tau} \]  
\( \Delta Q_a \) := ML target

Important parameter!

See Arcomano et al. (2022) for a corrective ML approach using reservoir computing
Corrective machine learning approach

- Online bias correction has a long history (e.g. Leith 1978, DelSole et al. 2008)
- Our approach:
  1. Nudge coarse-resolution model towards some reference dataset
  2. Train ML to predict nudging tendencies using coarse-res model state as input
  3. Run coarse-res model again, with ML predicting corrective tendency at every time step

Possible reference datasets

1. **Reanalysis**
   - The real world! (mostly)
   - Present-climate data readily available
   - State is well-constrained, but tendencies less-so

2. **Coarsened hi-res simulation**
   - Has its own biases
   - Can generate data from warmer/cooler climates
   - We know apparent sources

\[ \Delta Q_a = \frac{a - a_{ref}}{\tau} \implies \text{ML target} \]
Corrective machine learning approach

- Online bias correction has a long history (e.g. Leith 1978, DelSole et al. 2008)
- Our approach:
  1. Nudge coarse-resolution model towards some reference dataset
  2. Train ML to predict nudging tendencies using coarse-res model state as input
  3. Run coarse-res model again, with ML predicting corrective tendency at every time step

Possible reference datasets

1. Reanalysis
   - The real world! (mostly)
   - Present-climate data readily available
   - State is well-constrained, but tendencies less-so

2. Coarsened hi-res simulation
   - Has its own biases
   - Can generate data from warmer/cooler climates
   - We know apparent sources

\[ \Delta Q_a = \frac{a - a_{\text{ref}}}{\tau} \quad := \text{ML target} \]

\( \tau = 6 \text{ hours} \)
We use single-column approximation: predict columns of nudging tendencies using vertical profiles of inputs.

Features:
- air temp vertical profile
- specific humidity vertical profile
- land/sea mask
- cos(zenith angle)
- surface elevation

Predictions:
- Heating vertical profile
- Moistening vertical profile

In some cases also predict horizontal wind nudging tendencies
What do these tendencies look like?

2-year nudged run starting 1 Jan 2015

2015: training data
2016: test data

\[ \Delta Q_a = \frac{a - a_{\text{ref}}}{\tau} \Rightarrow \text{ML target} \]
What do these tendencies look like?

2-year nudged run starting 1 Jan 2015

2015: training data
2016: test data

Random Forest for ML

$R^2$ ranges from 0.1 - 0.3 depending on vertical level and variable

$$\Delta Q_a = \frac{a - a_{ref}}{\tau} \equiv \text{ML target}$$

Watt-Meyer et al. 2021, GRL
Online results for ML-corrected simulations

- Adds a full day to medium-range weather forecast skill
- Annual precipitation error reduced 23% vs. no-ML baseline 200 km model.

Watt-Meyer et al. 2021, GRL
Apply similar technique with fine-res reference

- Use coarsened 3-km SHiELD run (GFDL storm-resolving model) as reference
- Different features of this application:
  - Only 40 days of training data due to expense of fine-res simulation
  - 3-hour nudging timescale instead of 6-hour
  - Use neural network for ML prediction
    - Higher offline skill, but substantial regularization required
  - Train additional model for correction of surface radiative fluxes
Reduction in precipitation biases

- 30% improvement in time-mean precipitation RMSE
  - 3.66 mm/day in no-ML baseline to 2.56 mm/day in ML-corrected run
- Land-mean precipitation bias is almost entirely removed
  - This comes from the surface radiative flux correction
- Improvements in diurnal cycle over land

Bretherton et al. 2022, JAMES
Assessment of the nudging approach

The good

- Straightforward to implement with a wide range of reference datasets
- Existing physical parameterizations encode feedbacks that can keep model on the rails
- Builds on existing skillful parameterizations

The not-so-good

- Physical interpretation can be challenging
- Using difference in states to define error leads to smeared out signal, especially for temperature (weak-temperature gradient)

\[
\Delta Q_a = -\frac{a - a_{\text{ref}}}{\tau} \quad \text{ML target}
\]
Conclusions

• Nudging is useful and straightforward way to diagnose and correct model biases

• A corrective-ML approach leads to significant improvements to a coarse-resolution model’s forecast lead time and precipitation pattern
  ○ Some outstanding issues related to zonal mean circulation (lower stratosphere temperature drifts; Hadley cell slow-down)

• Lots of other exciting work happening in our group:
  ○ Parameterization replacement!
  ○ Use higher (25km) resolution for baseline model to be corrected
  ○ Correction of climate-change simulations (Chris will present Thursday)

• See papers:
  ○ Watt-Meyer et al. (2021) doi:10.1029/2021GL092555
  ○ Bretherton et al. (2022) doi:10.1029/2021MS002794

Code: github.com/ai2cm
Website: allenai.org/climate-modeling