THE CLOUD BRAIN V2

Achieving conservation and probing interpretability using a refined deep learning emulator of global cloud superparameterization

Mike Pritchard
Associate Professor
University of California, Irvine
Road Map

I. Emulating classical superparameterization on a simple aquaplanet.
   Summary of results, challenges & surprises from Rasp, Pritchard & Gentine (2018).

II. New tests towards adding land and real geography.
   Successful one-way land coupling, fit quality w/ geography & seasons, new challenges.

III. Respecting physics, probing interpretability.
   Achieving energy conservation, summarizing convective dynamics.
How will low cloud organization interact with climate dynamics?

Shallow cloud organization is ubiquitous in observations. The relevant turbulent processes can be simulated in LES...

...which suggest future changes in cloud organization may impact TOA radiation. Unless buffered, this will demand ITCZ shifts & ocean circulation changes.

Why does vegetation dry itself out over the Amazon?

At high CO₂, more efficient plant water use changes surface energy partitioning. Parameterized turbulence responds by lofting more vapor to altitudes where it can be flushed by Andean mountain jets. Thus starving the Amazon of rainfall through column energetics.

Turbulence matters

MOTIVATION
MOTIVATION

Turbulence matters

To planetary climate dynamics

To the regional water cycle
Unsatisfying approximations of turbulence in global climate models seem inescapable.
Deep Learning emulation might allow high definition 3D turbulence ahead of schedule!

If the job is hard, e.g. simulating the whole atmosphere for decades...

...satisfying 3D turbulence calculations can seem too much even for powerful computers.
Deep Learning emulation might allow high definition 3D turbulence ahead of schedule!

If the job changes to making short simulations just for training machine learning emulators...

...we can do much more justice to turbulence physics.
I. Emulating classical superparameterization on a simple aquaplanet.
   Summary of results, challenges & surprises from Rasp, Pritchard & Gentine (2018).

II. New tests towards adding land and real geography.
   Successful one-way land coupling, fit quality w/ geography & seasons, new challenges.

III. Respecting physics, probing interpretability.
   Achieving energy conservation, summarizing convective dynamics.
I. Emulating classical superparameterization on a simple aquaplanet.

Summary of results, challenges & surprises from Rasp, Pritchard & Gentine (2018).

II. New tests towards adding land and real geography.

Successful one-way land coupling, fit quality w/ geography & seasons, new challenges.

III. Respecting physics, probing interpretability.

Achieving energy conservation, summarizing convective dynamics.
Cloud SuperParameterization (SP)

Strategically undersampling horizontal space to explicitly represent important sub-grid processes.
In the past decade the 1-4 km resolution regime was a frontier that SuperParameterization helped explore.

Now Global Cloud Resolving Models handle this more elegantly.
This decade, superparameterization could help penetrate the turbulence-permitting frontier.

But it takes a heck of a lot of computing power.

“Cori” at NERSC in Berkeley - 30 petaflops
~ 2,000 Intel Haswell nodes
(~ 75,000 2.3 GHz cores)
Training NNs on SP data is easier than coarse-graining.

**Pro:** Like nature, no artificial scale separation in the data.

**Con:** Coarse-graining draws on after the fact. No clean info on what’s needed to correct a coarse-res model.

**Pro:** Convenient scale separation in the training data, well suited to correcting exterior model.

**Con:** That scale separation is utterly artificial & interferes with modes of variability.
Schematic of the sort of NN we will use.

Inputs (large scale)

- T profile (30 values)
- q profile (30 values)
- Sensible heat flux
- Latent heat flux
- Surface pressure
- Incoming sunlight
Schematic of the sort of NN we will use.

Inputs (large scale)
- T profile (30 values)
- q profile (30 values)
- Sensible heat flux
- Latent heat flux
- Surface pressure
- Incoming sunlight

128 node hidden layers

Outputs (SGS tendency)
- Convective + radiative heating rate profile (30 values)
- Convective moistening rate profile (30 values)
- Precipitation rate
- TOA radiative flux
- Sfc radiative flux
Is deep learning viable for emulating superparameterization?

Zonally symmetric aquaplanet testbed with classical superparameterization

1 year for training
1 year for validation

Time-step level output (incl. what is needed to close budgets)
Is deep learning viable for emulating superparameterization?

Global aquaplanet testbed

Can 140,000,000 outputs from 1 year of ~ 10,000 cloud-resolving models...

Be fit by a deep neural network?

Gentine, Pritchard, Rasp et al., GRL, 2019.
Is deep learning viable for emulating superparameterization?

Yes, e.g., $R^2 > 0.7$ for mid-tropospheric heating by convection and radiation.

Quite possibly!

The “Cloud Brain”

Global aquaplanet testbed

Can 140,000,000 outputs from 1 year of ~10,000 cloud-resolving models...

Be fit by a deep neural network?

Can be fit by a deep neural network for mid-tropospheric heating by convection and radiation.
Prognostic tests: Neural Network producing same mean climate 20x faster.

High enough accuracy to correct convective biases in CAM at low computational cost ( < 10% )

Spooky out-of-sample generalizability was found in the NN: Response to +3K warm pool perturbation of zonally symmetric aquaplanet

Adding zonally asymmetric SSTs to the SP benchmark solution... ...anchors a Walker cell as expected.

Same response in the Neural-Network GCM despite no Walker Cell in its training data... ...and despite these SSTs exceeding maximum values in the training data set by +3K.

Quandary: generalizability has limits that are totally empirical

Response to +1K to +4K surface warming

Benchmark:

NN prediction:

Quandary: instabilities abound and stable runs are rare.

Example of the neural network blowing up in prognostic mode.

Figure courtesy of Tom Beucler.
I. Emulating classical superparameterization on a simple aquaplanet.

Summary of results, challenges & surprises from Rasp, Pritchard & Gentine (2018).

II. New tests towards adding land and real geography.

Successful one-way land coupling, fit quality w/ geography & seasons, new challenges.

III. Respecting physics, probing interpretability.

Achieving energy conservation, summarizing convective dynamics.
Road Map

I. Emulating classical superparameterization on a simple aquaplanet.
   Summary of results, challenges & surprises from Rasp, Pritchard & Gentine (2018).

II. New tests towards adding land and real geography.
   Successful one-way land coupling, fit quality w/ geography & seasons, new challenges.

III. Respecting physics, probing interpretability.
   Achieving energy conservation, summarizing convective dynamics.
I. Emulating classical superparameterization on a simple aquaplanet.
   Summary of results, challenges & surprises from Rasp, Pritchard & Gentine (2018).

II. New tests towards adding land and real geography.
   Successful one-way land coupling, fit quality w/ geography & seasons, new challenges.

III. Respecting physics, probing interpretability.
   Achieving energy conservation, summarizing convective dynamics.
Successful one-way coupling to land model despite fit imperfections.
Same nonlinear land model drift structure & unsteady carbon cycle adjustments as benchmark.

5-year CLM4 integrations for fictitious Amazon forest tiles exposed to surface inputs from NNCAM, across 112 points within 15S-15N.
## Relaxing the aquaplanet idealizations

<table>
<thead>
<tr>
<th>Model version:</th>
<th>SPCAM3.0</th>
<th>SPCAM5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamical core:</td>
<td>Spectral + semi-Lagrangian</td>
<td>Finite-volume, 2-deg</td>
</tr>
<tr>
<td>Physics columns:</td>
<td>~8k</td>
<td>~14k</td>
</tr>
<tr>
<td>No geography or land</td>
<td></td>
<td>Real geography &amp; land</td>
</tr>
<tr>
<td>Perpetual equinox</td>
<td></td>
<td>Full seasonality</td>
</tr>
<tr>
<td>Weak oceanic diurnal cycles</td>
<td></td>
<td>Realistic diurnal cycles</td>
</tr>
<tr>
<td>Zonal symmetry</td>
<td></td>
<td>Walker cells, asymmetric storm tracks, etc.</td>
</tr>
</tbody>
</table>
Successful **composite** diurnal rainfall cycle in new DNN fit.

Benchmark solution:

Neural network predictions:
But also many unrealistically “detectable” diurnal signals.

Benchmark solution:

Neural network predictions:

A new rainfall emulation challenge over subtropical arid land regions
For high frequency details, it is harder to fit with geography & seasons.

Zonal mean convective heating rate assessed via misfit of timestep-level (15-min) prediction quality

Aquaplanet benchmark
Rasp, Pritchard & Gentine (2018)

Preliminary result from real-geography
Mooers, Pritchard et al. (in prep)
But encouraging results for longer than diurnal timescales so far.

Tropical band: **Daily mean** skill, convective heating rate: 15S-15N.

Figure courtesy of Griffin Mooers
First-year UCI PhD student.
I. Emulating classical superparameterization on a simple aquaplanet.
  Summary of results, challenges & surprises from Rasp, Pritchard & Gentine (2018).

II. New tests towards adding land and real geography.
  Successful one-way land coupling, fit quality w/ geography & seasons, new challenges.

III. Respecting physics, probing interpretability.
  Achieving energy conservation, summarizing convective dynamics.
Road Map

I. Emulating classical superparameterization on a simple aquaplanet.
   Summary of results, challenges & surprises from Rasp, Pritchard & Gentine (2018).

II. New tests towards adding land and real geography.
    Successful one-way land coupling, fit quality w/ geography & seasons, new challenges.

III. Respecting physics, probing interpretability.
    Achieving energy conservation, summarizing convective dynamics.
I. Emulating classical superparameterization on a simple aquaplanet.
   Summary of results, challenges & surprises from Rasp, Pritchard & Gentine (2018).

II. New tests towards adding land and real geography.
   Successful one-way land coupling, fit quality w/ geography & seasons, new challenges.

III. Respecting physics, probing interpretability.

→ Achieving energy conservation, summarizing convective dynamics.

Road Map
The neural network begins to learn on its own to quasi-conserve column moist static energy without direction, but with error.

Quandary: data-driven NN parameterizations don’t strictly obey conservation laws.

Figure courtesy of Tom Beucler, UCI Postdoc.
How to physically constrain neural network parameterizations?

Tom Beucler’s idea:
Write physical constraints as function of input \( x \) and output \( y \).

\[
\begin{bmatrix} C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0
\]

Four constraints:
Conservation of column energy & mass
Consistency of longwave & shortwave radiative heating
How to physically constrain neural network parameterizations?

Option #1: Through the loss function:

\[
\text{Loss} = \left(1 - \alpha\right) (\text{Mean - squared error}) , \quad \alpha \in [0, 1]
\]
Option #2: Through the architecture:

Tom Beucler’s idea: Enforce n constraints within the neural net architecture.
Tom’s new architecture-constrained version of our neural network obeys physical constraints close to numerical precision.

Interpreting the black-box:
Neural network assisted dynamical analysis.

“Dynamical response matrices” like this that summarize moist convection have been made before but just for idealized tropical basic states.
They are usually hard-won.

Jacobian of the neural network fit to a superparameterized aquaworld.
(nondimensional convective growth rate in response to input T, q perturbations)
First glimpse of the basic state dependence of Kuang’s linear convective response matrix.

Diagnostics like this come along with neural network training “for free”

Deriving this with standard methods would be inconceivably labor-intensive.
Philosophical remarks & outlook.
Quandary: Even if it can be made stable and robust, what is sacrificed in relinquishing physics to a black-box?

Experimental process knobs?

Cheap skill

Interpretable parameter groups?

Tunability?
A case for the black box

Might DOF of system be too big for human brain to encapsulate in aesthetic, interpretable cartoons & parameters?

Schematic courtesy of Chris Bretherton, UW.
A case for millions of parameters

Chollet’s “geometric interpretation of deep learning”

Deep NNs do this by “incrementally decomposing a complicated geometric transformation into a long chain of elementary ones”
Deep learning has breakthrough potential. Already a surprisingly good emulator of deep moist turbulence. What else might be satisfyingly “emulatable”? In-cloud chemistry coupled to dynamics? Spectral bin microphysics? Better discretizations for PDE solvers? Species-level ecosystem dynamics?

Even short superparameterized simulations can be mined for their essence. To create efficient emulators with same emergent benefits. But issues of instability are yet to be resolved!

For compactly interpreting & intercomparing highly complex dynamical systems. Our community has only scratched the surface.

WHERE CAN WE GO FROM
THANKS

It is an exciting time for numerical climate dynamics!

mspritch@uci.edu
NEW GROUP MEMBERS

Crystal Cove State Park
(10 min drive)
NEW GROUP MEMBERS

Laguna Beach
(20 min drive)
CONTRASTS TO IMAGE
In what ways is the cloud parameterization emulation problem different?

What is interesting in the comparison?
DIMENSIONALITY & DATA

Moderate data amount

VS

Massive data amount

e.g. 10,000 labeled images

100,000,000 synthetic training samples
**DIMENSIONALITY & DATA**

- **High dimensional input per sample**
  - $100 \times 100$ pixels x 3 colors = 30,000 incoming values.

- **Moderate data amount**
  - e.g. 10,000 labeled images

- **Massive data amount**
  - 100,000,000 synthetic training samples

- **Low dimensional input...**
  - 3 state vars x 30 levels
  - ~ 100 incoming values (300x less)
Image processing is at totally different limits.

Computer scientists have no reason to know yet what is possible at ours!
CONVNETS & IMAGE
VISUALIZING WHAT CONVNETS

Once trained, hit the NN with a test image.
See what lights up.
VISUALIZING WHAT CONVNETS

“Fourth channel of the activation of the 1st layer on the test cat picture”

“Seventh channel of the 1st layer”
FROM SMALL BUILDING
TO ABSTRACT
(My favorite)
Sheer size of images demands some type of dimensionality reduction.

ConvNets accomplish this yielding hierarchies of increasingly abstract “representations” of “catness” learnt from big labeled imaged libraries.

Simple way to see what “features” of an image the NN exploits to achieve its skill.

Natural to wonder if there is an analogy to our question of interest..

What “features” of "environmental thermodynamics" give our NN emulator of SP its skill in predicting “convective adjustments”??
IF THIS SOUNDS FUN TO YOU..
POSTDOC & PROJECT SCIENTIST

The University of California, Irvine
NEW GROUP MEMBERS

Crystal Cove State Park
(10 min drive)
NEW GROUP MEMBERS

Laguna Beach
(20 min drive)
THANKS

It is an exciting time for numerical climate dynamics!

mspritch@uci.edu