Deep learning for climate modeling
Week 2: A data-driven approach to convective parametrization

September 11, 2018

Presenters: Pierre Gentine and Mike Pritchard
Notes takes by: Tom Beucler

Contents

1 What is a convective parametrization?

A model aimed at getting the cloud-scale mass flux profile and the corresponding turbulent tracer transport (e.g. turbulent vertical heat flux $\overline{w\prime T\prime}$, where $w\prime$ is vertical velocity turbulent anomaly and $T\prime$ is temperature turbulent anomaly). Typical convective parametrizations rely on a lot of assumptions:

- Specify mass flux at cloud base
- Specify entrainment and detrainment profiles
- Invoke a quasi-equilibrium assumption
- Based on a connection between boundary layer and convection.
- Can then add a triggering function preventing convection from triggering every time the static stability is negative (e.g. over land)

Pathologies of man-made convective parametrizations:

- [Bechtold et al., 2014] Problem of diurnal precipitation peak over land (e.g. peak too early, no late-night convection)
- [Couvreux et al., 2015] Bias in the convective heating profile.
- Other problems include convective memory and short-time scale behavior.
- Quasi-equilibrium closure imposes too short timescales over land.
- Entrainment depends on environmental conditions in practice, for instance would increase when relative humidity increases in reality.
- Furthermore, entrainment is a stochastic process (dependence on stochastic mixing rather than solely cloud base conditions, since the memory of cloud base conditions is quickly lost).
- Convective parametrizations tend to trigger convection too frequently and too little ("drizzle problem"), which means that precipitation extremes are ill-resolved (typically top 1% precipitation extremes).
- Cannot simulate mesoscale convective systems, which are mesoscale (~100km) precipitation clusters that bring most of the high-impact precipitation events.
- When the atmosphere precipitates, some of the precipitation re-evaporates, which cools the atmosphere and triggers gravity currents. These gravity currents can trigger convection remotely and therefore organize convection.

Although progress can be made by introducing a lag in the response of the boundary layer or parametrizing cold pools, the limits of convective parametrizations limit our ability to predict the changes in cloud radiative effects and precipitation with climate, even in an aquaplanet configuration [e.g. Stevens, Bony 2013]. These limitations impede our ability to predict regional climate change.

2 Super-parametrization of convection

2.1 Cloud-permitting models

Cloud-permitting models (~1-10km horizontal resolution) can resolve convective processes, and therefore correct many of the biases intrinsic to convective parametrizations. Examples include:

1. The diurnal cycle of precipitation over the Amazon (convective parametrizations typically underestimate the precipitation intensity in this case)
2. Captures the magnitude of surface precipitation extremes in mid-latitudes (e.g. Oklahoma)

2.2 What is super parametrization?

It is possible to embed two-dimensional cloud-permitting models in global climate models to better resolve the convection-equilibrated thermodynamic tendencies. This considerably reduces the biases in moist large-scale circulations as well as precipitation patterns (macro-statistics of convection). However, certain aspects of convection, such as momentum transport, cannot be represented because the embedded cloud-permitting model is not three-dimensional. Because the macro-statistics are better represented, mesoscale radiation patterns, cold pools, etc. can be studied in the context of climate models.

2.3 Limits of super-parametrization

Because the cloud-permitting model is doubly periodic, it cannot properly simulate mesoscale disturbances such as mesoscale convective complexes, but they can resolve the winds necessary to advect them. Furthermore, they are very expensive (more than 100 times the cost of a regular global climate model).

3 A data-driven approach: Machine learning to represent convection

3.1 Setup

The idea is to use the wealth of data available from super-parametrized global climate models to train a network that takes as inputs:

- The temperature profile
- The specific humidity profile
- The surface pressure
- The surface enthalpy fluxes
- The top-of-the-atmosphere shortwave radiative flux

and predicts the physical tendencies due to convection as outputs:

- The temperature tendency profile due to both convection and radiation
- The moisture tendency profile
- The precipitation
- The top-of-the-atmosphere radiative flux

The cost function during the training phase is the mean square error of the output vector.

3.2 Results using the super-parametrized community atmospheric model

The mesoscale precipitation and circulation features are well reproduced, as well as the heating and moistening profiles. Once embedded in the climate model to parametrize convection, the neural network decreases the exaggerated Kelvin wave signal and has a good Madden-Julian Oscillation signal.

3.3 Predictive abilities

Although the initial aquaplanet training data set did not have any longitudinal asymmetry, the neural network was able to predict a Walker circulation once zonal asymmetries were introduced. Furthermore, the neural net parametrization also significantly reduces the double Inter-Tropical Convergence Zone (ITCZ) bias.

There are limitations if the neural net is out of its training dataset: For instance, if trained on past climates, it cannot reproduce the future (+1K,+2K,+3K,+4K) climates (e.g. double ITCZ, etc.). If trained on the past (+0K) and future (+4K) climate, then it can properly interpolate in between (+1K,+2K,+3K).

In conclusion, it is possible to represent sub-grid scale processes given only coarse-grained values, although some of the stochastic variability is missing (e.g. problematic for weather predictions).

4 Story of the making/Discussion

4.1 Neural network characteristics

- The workflow evolved from the Tensorflow to the Keras library for the deep learning scripts
- Only requires 3-6 months of data to train the neural network, which opens the possibility of using large eddy simulation data to train even better neural network parametrizations
- 256×8 layer neural network
- Using a convolution network did not improve the predictive abilities
- Non-dimensionalizing the data did not help either; the neural network seems to learn the non-dimensionalization by itself
- The cost function was simply the RMSE from the output vector
- Most of the results were insensitive to the output normalization decisions
- The learning rate was the most important hyper-parameter

4.2 Neural-net global climate model performances

- There is less and less skill near the surface as quantified by mean square error, but the neural-driven climate models produce reasonable signals when coupled one way with a land model.
- Conservation properties almost satisfied: the neural network seems to be able to "learn" physical laws to some extent
- Runs 20 times faster than the super-parametrized model

4.3 Issues

- Stability problems, especially when trying to extend to the 32 column version of SPCAM

4.4 Future directions

- Using three-dimensional data to train the network