

RESEARCH LETTER

10.1029/2018GL078202

Key Points:

- We use a global atmospheric model with embedded cloud resolving model, as training data set for a machine learning algorithm of convection
- The machine learning algorithm can reproduce most of the key features of the embedded cloud resolving model
- The machine learning algorithm is much more computationally efficient than a super parameterization, but does not behave stochastically

Supporting Information:

- Supporting Information S1

Correspondence to:

P. Gentine,
pg2328@columbia.edu

Citation:

Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G., & Yacalis, G. (2018). Could machine learning break the convection parameterization deadlock? *Geophysical Research Letters*, 45, 5742–5751. <https://doi.org/10.1029/2018GL078202>

Received 6 APR 2018

Accepted 18 MAY 2018

Accepted article online 23 MAY 2018

Published online 12 JUN 2018

©2018. The Authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Could Machine Learning Break the Convection Parameterization Deadlock?

P. Gentine¹ , M. Pritchard² , S. Rasp³ , G. Reinaudi¹, and G. Yacalis² 

¹Earth and Environmental Engineering, Columbia University, New York, NY, USA, ²Earth System Science, University of California, Irvine, CA, USA, ³Faculty of Physics, LMU Munich, Munich, Germany

Abstract Representing unresolved moist convection in coarse-scale climate models remains one of the main bottlenecks of current climate simulations. Many of the biases present with parameterized convection are strongly reduced when convection is explicitly resolved (i.e., in cloud resolving models at high spatial resolution approximately a kilometer or so). We here present a novel approach to convective parameterization based on machine learning, using an aquaplanet with prescribed sea surface temperatures as a proof of concept. A deep neural network is trained with a superparameterized version of a climate model in which convection is resolved by thousands of embedded 2-D cloud resolving models. The machine learning representation of convection, which we call the Cloud Brain (CBRAIN), can skillfully predict many of the convective heating, moistening, and radiative features of superparameterization that are most important to climate simulation, although an unintended side effect is to reduce some of the superparameterization's inherent variance. Since as few as three months' high-frequency global training data prove sufficient to provide this skill, the approach presented here opens up a new possibility for a future class of convection parameterizations in climate models that are built "top-down," that is, by learning salient features of convection from unusually explicit simulations.

Plain Language Summary The representation of cloud radiative effects and the atmospheric heating and moistening due to moist convection remains a major challenge in current generation climate models, leading to a large spread in climate prediction. Here we show that neural networks trained on a high-resolution model in which moist convection is resolved can be an appealing technique to tackle and better represent moist convection in coarse resolution climate models.

1. Introduction

Convective parameterization remains one of the main roadblocks to weather and climate prediction (Bony et al., 2015; Medeiros et al., 2014; Sherwood et al., 2014; Stevens & Bony, 2013). In fact, most of the intermodel spread in equilibrium climate sensitivity can be traced back to the representation of clouds (Schneider et al., 2017). Convective schemes exhibit systematic biases in the vertical structure of heating and moistening, precipitation intensity, and cloud cover (Daleu et al., 2015, 2016). These errors, in turn, feed back onto larger-scale circulations, deteriorating general circulation model (GCM) simulations, and prediction skill (Bony et al., 2015). A challenge in current convective schemes is representing the transitions between different types of convection, such as the transition from shallow to deep convection (Couvreur et al., 2015; D'Andrea et al., 2014; Khouider et al., 2003, 2010; Khouider & Majda, 2006; Peters et al., 2013; Rochetin, Couvreur, et al., 2014; Rochetin, Grandpeix, et al., 2014), which is especially crucial to predicting both continental precipitation and modes of climate variability (Arnold et al., 2014). In addition, most convective parameterizations do not represent important processes, such as convective aggregation, which are essential to accurately predicting the response of clouds and precipitation to global warming, as well as modes of climate variability (Arnold & Randall, 2015; Bony et al., 2015; Bretherton & Khairoutdinov, 2015; Coppin & Bony, 2015; Jeevanjee & Roms, 2013; Muller & Bony, 2015; Wing & Emanuel, 2014).

Current generation climate models (and typical weather forecast models) with parameterized convection do not capture much of the degree of organization, nor do they represent mesoscale convective systems (MCS; Hohenegger & Stevens, 2016). MCS and the impact of shear are, however, crucial for correct representation of rainfall and radiative feedback (Cao & Zhang, 2017; Houze, 2004; Moncrieff, 2010; Moncrieff et al., 2012, 2017; Tan et al., 2015). Finally, another challenge is that climate sensitivity is strongly related to the interaction between deep and shallow convection (Bony et al., 2015), and the coupling between clouds, convection, and the large-scale circulation, which is currently poorly captured by parameterized convection (Bony et al., 2015; Daleu et al., 2016; Hohenegger & Stevens, 2016; Nie et al., 2016).

Many of the previously mentioned problems related to the representation of convection are partly alleviated when using convective-permitting resolutions, that is, at horizontal grid spacing of ~ 2 km or less. For instance, the transition between shallow and deep convection can be correctly captured at convective-permitting scale (Khairoutdinov et al., 2009; Khairoutdinov & Randall, 2006). Convective aggregation is observed at convective-permitting scale (Hohenegger & Stevens, 2016) so that cloud resolving models (CRMs) have been the tool of choice to understand convective aggregation (Arnold & Randall, 2015; Bony et al., 2015; Bretherton & Khairoutdinov, 2015; Coppin & Bony, 2015; Jeevanjee & Romps, 2013; Muller & Bony, 2015; Wing & Emanuel, 2014). CRMs (at convective-permitting scales < 2 km) also correctly reproduce MSCs and squall lines (Moncrieff & Liu, 2006; Taylor et al., 2009), as well as extreme precipitation events driven by larger scale anomalies. Convective-permitting simulations better represent modes of tropical climate variability (Arnold et al., 2014), shallow to deep convection (Guichard et al., 2004), and mesoscale propagation (Hohenegger et al., 2015).

Therefore, modeling at convective-permitting scales is transformative to the representation of convection. It is, however, impractical at present to use convective-resolving resolution at global scale for climate prediction given its computational requirements (Satoh et al., 2008). While global cloud-resolving models (GCRMs) can be run easily for months, multidecadal simulations are computationally challenging. To alleviate this problem, an interesting approach has been to use cloud “superparameterization (SP),” which computes the sub-grid vertical heating and moistening profiles within a GCM grid cell by sampling a curtain of an embedded 2-D CRM that uses convective-permitting resolution (Grabowski, 1999; Khairoutdinov et al., 2005). This has led to many successes such as the possibility to rectify the diurnal continental cycle, to improve the representation of the MJO, and to represent both some MCS propagation and some degree of aggregation, and reduce overly strong land-atmosphere coupling (Benedict & Randall, 2009; Grabowski, 2001; Holloway et al., 2012, 2015; Khairoutdinov et al., 2005; Kooperman et al., 2016a, 2016b; Pritchard & Somerville, 2009; Pritchard et al., 2011; Qin et al., 2018; Randall, 2013; Sun & Pritchard, 2016).

While promising, SP is not without its own idealizations that also limit its predictive ability and usefulness for climate simulation. For instance, restricting explicit convection to two dimensions makes it difficult to represent momentum transport (Arakawa, 2011; Jung & Arakawa, 2010; Tulich, 2015; Woelfle et al., 2018), and the limited CRM domain extent artificially constrains vertical mixing efficiency (Pritchard et al., 2014). Meanwhile, the typical use of 1–4 km CRM horizontal resolution and 250-m vertical resolution cannot resolve important boundary layer turbulence, lower tropospheric inversions, and associated entrainment that are critical to low cloud dynamics (Parishani et al., 2017).

In light of this ongoing deadlock, we propose to use an alternative approach to convective parameterization in which convection is represented using a machine-learning algorithm based on artificial neural networks (ANNs), trained on superparameterized simulations, called Cloud Brain (CBRAIN). ANNs can approximate any nonlinear deterministic function, a property called the universal approximation theorem (Schmidhuber, 2015). Clearly, parameterizing convection appears as an ideal problem for the use of machine learning algorithms and especially ANNs. Indeed, machine-learning algorithms have been used in many applications where a clear physically based algorithm could not be defined. Applications have included self-driving cars, society games (chess and go; Silver et al., 2016), speech recognition (Hinton et al., 2012), object recognition and detection, medical detection of cancers (Karabatak & Ince, 2009; Khan et al., 2001; Zhou et al., 2002), and genomics. There are also applications of ANNs to the geosciences, such as for rainfall prediction (Miao et al., 2015; Moazami et al., 2013; Tao et al., 2016), weather forecast, soil moisture (Kolassa et al., 2013, 2016; Kolassa, Gentile, et al., 2017; Kolassa, Reichle, & Draper, 2017), and surface turbulent flux retrievals (Alemohammad et al., 2017; Jimenez et al., 2009; Jung et al., 2011). Specifically, the development of deep learning and deep neural networks, that is, those with multiple hidden layers, has led to important developments in many different fields such as object detection or game strategy learning (Dahl et al., 2011; Hinton et al., 2012; LeCun et al., 2015; Silver et al., 2016; Tao et al., 2016). One of the advantages of ANNs is that once trained, they are computationally efficient, as most of the computational burden is dedicated to the training phase.

Our aim here is to use such ANN techniques to better parameterize convection in coarse-scale climate simulations by learning from cloud-permitting SP-simulations, while trying to minimize the computational cost compared to those cloud-permitting simulations, which are still computationally prohibitive.

2. Data

2.1. SuperParameterized Community Atmosphere Model

To evaluate this idea, we use a well-validated version of the SuperParameterized Community Atmosphere Model (SPCAM3) in a simplified aquaplanet configuration with zonally symmetric SSTs following a realistic meridional distribution (Andersen & Kuang, 2012). The global model uses a spectral dynamical core with approximately two-degree horizontal resolution (T42 triangular truncation) and 30 levels in the vertical. The CRM uses a simplified bulk one-moment microphysics scheme and a Smagorinsky 1.5-order subgrid scale turbulence closure as described by (Khairoutdinov et al., 2003) and shares the host GCM's vertical grid. For computational efficiency and convenience we use the "micro-CRM" (8-column) CRM domain discussed by Pritchard et al. (2014) for this proof of concept. Following a three-month spin-up period, we save global data at the host global model time step frequency (every 30 min) representing arterial inputs to (and outputs from) each of 8,192 cloud-resolving arrays embedded SPCAM. The simulation is run for two years, yielding around 140 million training samples per year. One year of data represents 375 Gb.

3. Neural Network Setup

We are using an ANN to predict SPCAM's total physics package tendencies, that is, the cumulative tendency produced by turbulence, convection, and radiation. Rather than purely isolating any of the above subtendencies from the CRM or GCM parameterizations, we chose a holistic approach in representing their sum—that is, the arterial total heating and moistening profiles that ultimately link a GCM's subgrid physics to its dynamical core. This has practical advantages in that the individual physical subprocesses—turbulence, convection, microphysics, and radiation—can interact in complex, nonlinear ways. Approximating the net effect of such interactions is one of the big strengths of ANNs. The ANN is not interacting with the dynamical core and uses the same inputs as SPCAM at each time step.

The ANN is written using the Python library Keras (<https://keras.io>), a high-level wrapper around TensorFlow (<http://www.tensorflow.org>). The code for the ANN training as well as for the validation and analysis below can be found at <https://github.com/raspstephan/CBRAIN-CAM>. Training took on the order of 12 hr on a graphical processing unit (Nvidia GTX 970). The first year of SP-CAM data was used for training, while the second year was used for independent validation.

The feedforward ANNs consist of interconnected layers, each of which has a certain number of nodes (Figure S1). The input and output variables are listed in Table 1. The first layer is the input layer, which in our case is a stacked vector containing the input variables including their vertical variation for a specific column. No latitude or longitude information is specifically passed to the neural network, meaning that we train a single neural network to be used for every column. The last layer is the output layer, which again is a stacked vector of the four output vertical profile variables. All layers in between are called hidden layers. Deep neural networks have more than one hidden layer. The values in the nodes of each layer are weighted sums of all node values in the previous layer plus a bias, passed through a nonlinear activation function. Here we used the Leaky Rectified Linear Unit (LeakyReLU) $a(x) = \max(0.3x, x)$, which resulted in better scores compared to other common activation functions such as tanh, sigmoid, or regular ReLU. The output layer is purely linear without an activation function.

Training an ANN means optimizing the weight matrices and bias vectors that define it, to minimize a loss function—in our case the mean squared error—between the ANN outputs and the truth for a given input. The loss is computed for a shuffled (in space and time) minibatch of the training data with a batch size of 1,024 samples. To reduce the loss, the gradient of the loss function with respect to all weights and biases is computed using a backpropagation algorithm, followed by a step down the gradient—that is, stochastic gradient descent. In particular, we use a version of stochastic gradient descent called Adam (Kingma & Ba, 2014). How much to step down the gradient is determined by the learning rate. We started with a learning rate of 10^{-3} , dividing it by 5 every 5 epochs (i.e., five passes through the entire training data set). In total, we trained for 30 epochs. Regularization techniques were not necessary because we did not see any signs of overfitting given the large number of training samples. Despite the random initialization of the ANN weights and biases, the final result proved robust between training realizations.

Table 1*List of Input and Output Variables Used for the Neural Network*

Input variables	Vertical levels	Output variables	Vertical levels
Temperature at beginning of time step	30	Convective and turbulent temperature tendency	30
Humidity at beginning of time step	30	Convective and turbulent humidity tendency	30
Surface pressure	1	Longwave heating tendency	30
Sensible heat flux	1	Shortwave heating tendency	30
Latent heat flux	1		
Temperature tendency from dynamics	30		
Humidity tendency from dynamics	30		
Incoming solar radiation	1		
Size of stacked array	124		120

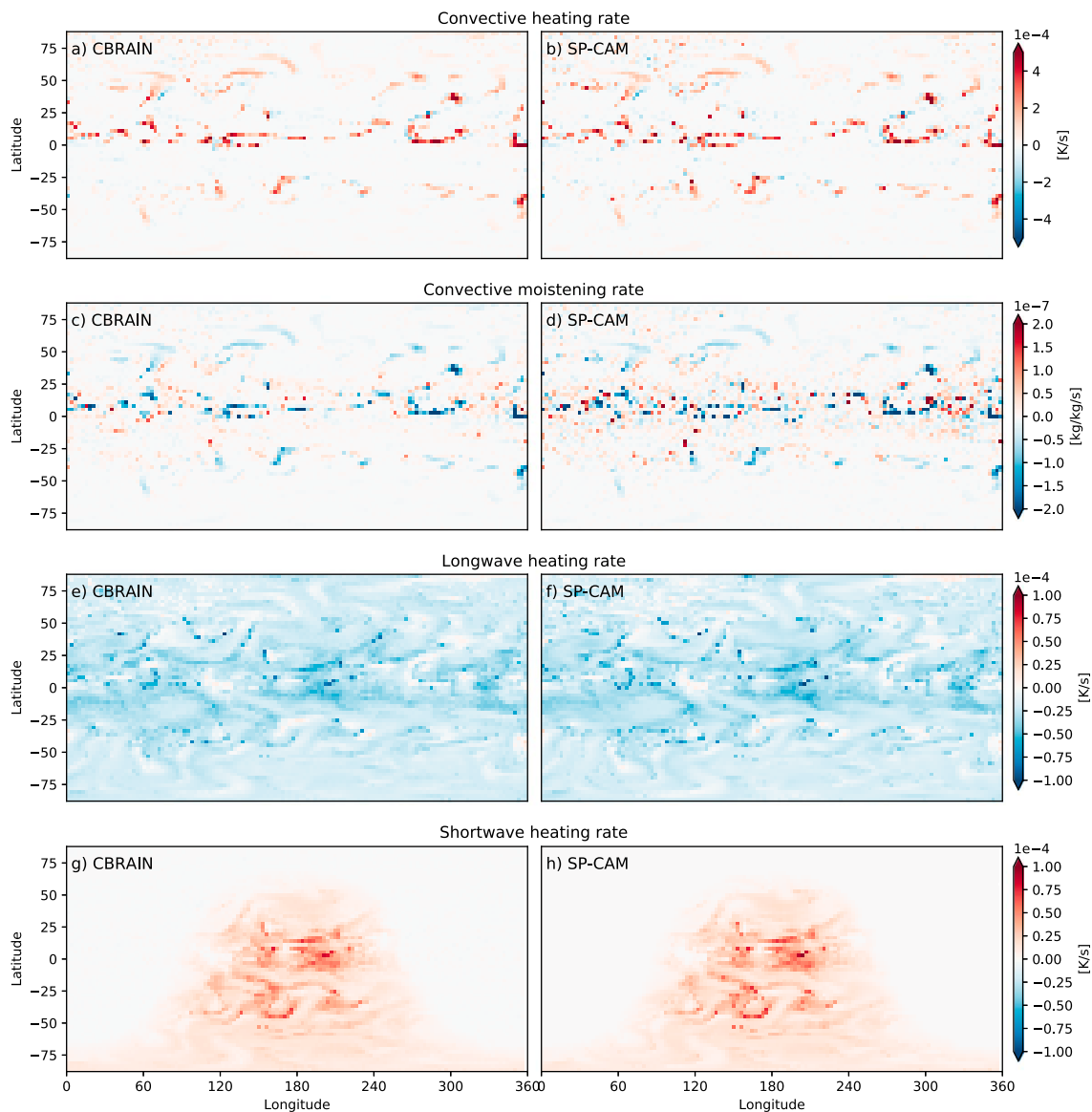


Figure 1. Latitude-longitude snapshot of neural network predictions and the corresponding SP-CAM truth at model level 20 (roughly 700 hPa) for one time step in the validation set.

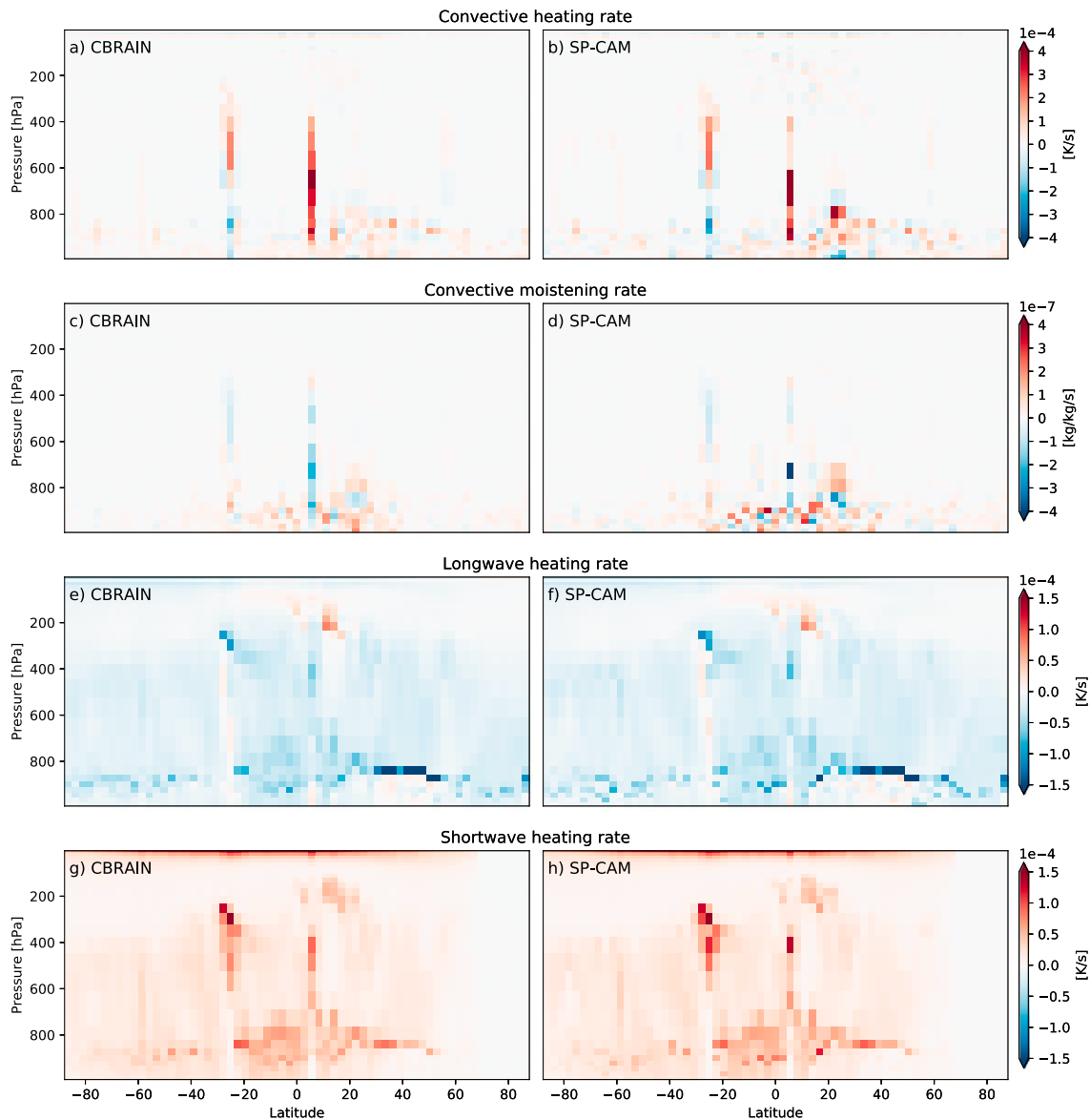


Figure 2. Pressure-latitude snapshot at 180° longitude corresponding to Figure 3.

For an ANN to train efficiently, all input values should be on the same order of magnitude. For this purpose, for each input variable, we subtracted the mean and divided by the standard deviation, independently for each vertical level; not normalizing did not modify any results but extended the duration of the training process. To make the outputs comparable, we converted the output variables (i.e., convective and radiative heating as well as convective moistening rates) to common energy units.

4. Results

4.1. Sensitivity to ANN Architecture and Amount of Training Data

We start by testing how the amount of ANN parameters and their configuration impacts the performance. Table S1 summarizes 12 separate ANN architectures tested. As a first metric of skill we assess a mean squared error statistic computed across all four output variables, all space, and all time during the second simulated year. That is, given knowledge of the inputs to each CRM, we measure the error across 143 million separate

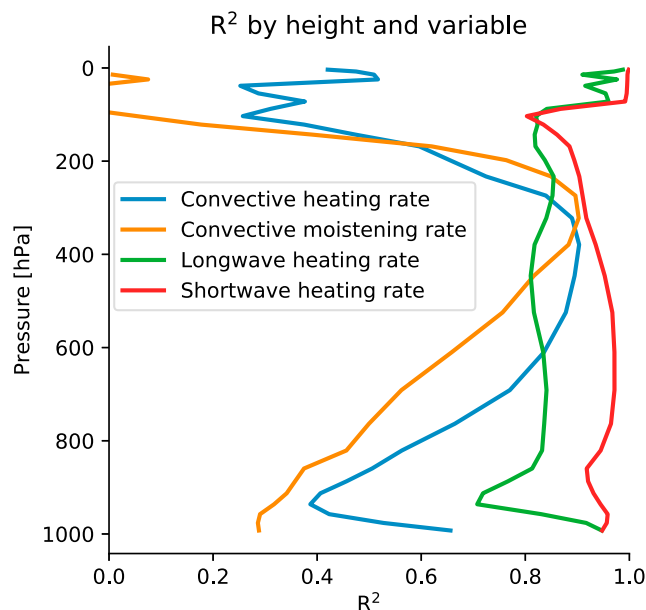


Figure 3. R^2 computed for each model pressure level and variable as described in the text.

ANN predictions of the CRM heating and moistening output profiles received by SPCAM's dynamical core, during a one-year time period that was not included in the training data set.

Figure S2a shows strong sensitivities to network architecture that underscore the importance of the ANN design—more parameters generally produce better scores and deeper networks give better results, because they also allow for more nonlinear interactions. For all subsequent analyses we thus only use our best performing network—a large, deep network with eight hidden layers of 512 nodes each.

A key question for the generalizability of our approach is how much training data is needed. For this we incrementally increase the length of continuous simulation data for training up to one year (Figure S2b). As expected, more training data do lead to better scores on the validation set. But, interestingly, three months appear to be sufficient to yield most of the information (Figure S2b). This suggests promising potential to generalize our approach beyond an SPCAM demonstration test bed to other simulation strategies that do even more justice to the true physics of moist convection. Indeed, three-month simulations are practical even for GCRMs or high-resolution, 3-D variants of SP.

4.2. Evaluation of NN Predictions

Latitude-longitude and pressure-latitude snapshots (Figures 1 and 2) provide a good qualitative starting point for evaluating the NN predictions (supplement videos). Overall, the NN predictions agree remarkably well with the SP-CAM truth in terms of horizontal and vertical structure. Lower tropospheric convective (turbulent and latent) heating and moistening associated with the intertropical convergence zone and extratropical cyclones occur at approximately the correct geographic locations (Figures 1a and 1d). The radiative heating rates show very good agreement, which is particularly impressive given the fact that there is no cloud condensate information in the input; that is, cloud-radiative feedback is all internal to the ANN. For instance, ANN skillfully predicts the geographic location of shortwave absorption by water vapor and regional cloud anomalies (Figures 1g and 1h) as well as the vertical location of longwave cooling maxima at the tops of subtropical boundary layer clouds and deep tropical clouds (Figures 2e and 2f). However, one issue for the convective heating and particularly moistening rates is that the NN predictions are smoother and do not exhibit as much of the variability as SP-CAM (internal stochastic variability). Indeed, the ANN is by definition deterministic and thus cannot reproduce any stochasticity.

To assess the quality of the predictions in more detail, we analyze R^2 averaged over both time and horizontal dimensions to yield statistics for each level and predicted variable (Figure 3). R^2 is defined as one minus the ratio of the sum of squared error to the true variance. The radiative heating rates are well represented throughout the column, particularly for shortwave heating. The convective tendencies interestingly show a distinct profile with less predictive skill in the boundary layer and the stratosphere. In the stratosphere, this lower skill is simply due to the near absence of convection at upper levels and likely not a concern. In the boundary layer, the reasons for reduced skill are discussed more below.

First, for a closer analysis of the skill in the troposphere, we also look at spatial statistics. Pressure-latitude maps of R^2 and the standard deviation (Figure 4) reveal patches of especially high skill in the midlevels at the equator and midlatitudes, which correspond to the locations of the Intertropical Convergence Zone and the midlatitude storm tracks. Since these are the locations of latent heating most fundamental to forcing the free tropospheric general circulation, this is reassuring regarding the potential of CBRAIN to reproduce important heating and moistening tendencies in future tests that could allow it to feedback with a dynamical core.

The skill in the boundary layer is significantly lower, again. One possibility is that this reflects the difficulty in representing mesoscale effects and subcloud layer organization as well as its memory (D'Andrea et al., 2014; Mapes & Neale, 2011). SPCAM does include some degree of convective aggregation (Arnold et al., 2015) and

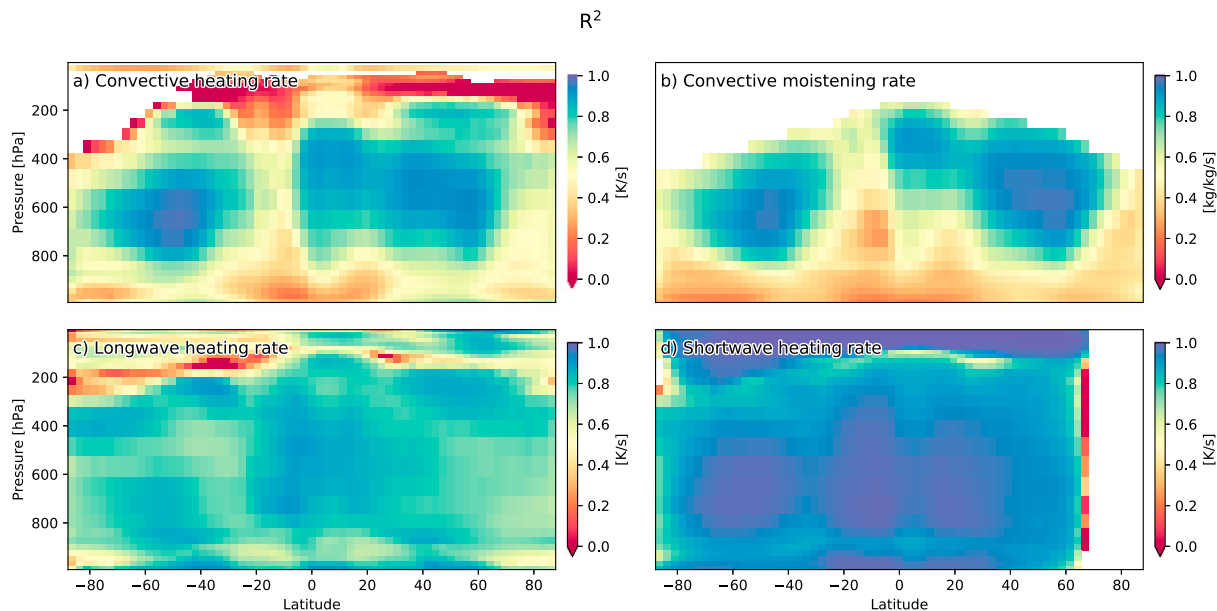


Figure 4. Pressure-latitude maps of (a and b) R^2 and (c and d) true standard deviation averaged over time and longitude. Regions where the variance was less than 0.05% of the global variance were masked out.

also carries memory of CRM organization from one-time step to the next through the embedded CRM (Pritchard et al., 2011). Our ANN does not include memory, as our objective was to mimic most current practice in convective parameterization, which is local in space and time. Another source of lower R^2 is related to the higher internal variability in SPCAM simulations compared to the ANN prediction, evident in Figures 1 and 2. This may be less of an issue in configurations that use larger, or 3-D CRMs; the small-extent 2-D CRMs used here are known to throttle deep updrafts and lead to unrealistically intense extremes (Pritchard et al., 2014). SPCAM does have some internal stochasticity (Subramanian & Palmer, 2017), which, by definition, a deterministic ANN cannot reproduce. The boundary layer and shallow convection tendencies, particularly for the moistening rate, are much noisier and thus appear much more stochastic than at higher levels. In these lower levels, the predictions here have significantly less variability in terms of its mean squared error loss function, which encourages the ANN to predict just an average value in cases where it is not certain.

5. Discussion and Conclusion

We have demonstrated that machine learning, and neural networks in particular, can skillfully represent many of the effects of unresolved clouds and convection, including their vertical transport of heat and moisture and the interaction of radiation with clouds and water vapor. The concept was proven in an idealized test bed using SPCAM over an aquaplanet. The implication of the success in this context is that an approach like CBRAIN could glean the advantages of GCRMs or high-resolution, 3-D SPs not yet practical for multidecadal climate simulations.

There are, however, important steps required for full implementation of CBRAIN in a GCM. First, neural networks do not intrinsically preserve energy and moisture. This can be fine for implementation in a weather forecast model but energy and moisture conservation are required for climate prediction. Second, neural networks are inherently deterministic. It was shown here that the resulting CBRAIN representation of heating and moistening tendencies was too smooth compared to the original SPCAM field used for training, which is more variable especially in the lower levels of the atmosphere (below 700 hPa). An important next test is to examine how CBRAIN feeds back with the GCM's resolved scale dynamics and surface fluxes. A final challenge is related to the fact that inherently a machine-learning algorithm is trained on existing data. For climate prediction, the algorithm should be able to generalize to situations that have potentially

not been seen such as changes in trace gas profile and concentrations or aerosols, and should be able to represent convection over continents.

Notwithstanding the above challenges, we believe that our preliminary results motivate the case that machine learning represents a powerful alternative to GCRMs or embedded-2-D CRM parameterizations. It is computationally efficient, even for relatively large networks. For instance, without specific optimization, a preliminary test showed that CBRAIN was 10 times faster than the micro-CRM form of SP used in our study and produces tendencies of unresolved physics comparable to SP. It would thus be several orders of magnitude faster than an SP equipped with large, 3-D, high-resolution domains, or a GCRM. CBRAIN is also naturally fitted for data assimilation since computation of the adjoint is straightforward and analytical, making it a natural candidate for operational weather forecasting. CBRAIN could represent a useful alternative to current parameterizations, which have followed a “bottom-up” deterministic strategy that still exhibits too many biases for satisfying prediction of the future hydrological cycle. A “top-down” strategy that instead learns the realistic complexity of simulated convection, as captured in short multimonth simulations at convection permitting resolution, is an attractive alternative. As global temperature sensitivity to CO₂ is strongly linked to convective representation, this might also improve our estimates of future temperature.

Acknowledgments

The neural network and analysis code can be found at <https://github.com/raspstephan/CBRAIN-CAM>. The exact version of the code base used for this study is tagged `grl_submission`. The raw SP-CAM output is very large (several TB) and available upon request to Mike Pritchard. M. P. acknowledges funding from the DOE SciDac and Early Career Programs (DE-SC0012152 and DE-SC00-12548) as well as the NSF (AGS-1734164). Stephan Rasp was funded by the German Research Foundation (DFG) Transregional Collaborative Research Center SFB/TRR 165 “Waves to Weather”. Computational resources for our SPCAM3 simulations were provided through the NSF Extreme Science and Engineering Discovery Environment (XSEDE) under allocation TG-ATM120034.

References

- Alemohammad, S. H., Fang, B., Konings, A. G., Aires, F., Green, J. K., Kolassa, J., et al. (2017). Water, Energy, and Carbon with Artificial Neural Networks (WECANN): a statistically based estimate of global surface turbulent fluxes and gross primary productivity using solar-induced fluorescence. *Biogeosciences*, 14(18), 4101.
- Andersen, J. A., & Kuang, Z. (2012). Moist static energy budget of MJO-like disturbances in the atmosphere of a zonally symmetric aquaplanet. *Journal of Climate*, 25(8), 2782–2804. <https://doi.org/10.1175/JCLI-D-11-00168.1>
- Arakawa, A., Jung, J. H., & Wu, C. M. (2011). Toward unification of the multiscale modeling of the atmosphere. *Atmospheric Chemistry and Physics*, 11(8), 3731–3742.
- Arnold, N. P., Branson, M., Burt, M. A., Abbot, D. S., Kuang, Z., Randall, D. A., & Tziperman, E. (2014). Effects of explicit atmospheric convection at high CO₂. *Proceedings of the National Academy of Sciences of the United States of America*, 111(30), 10,943–10,948. <https://doi.org/10.1073/pnas.1407175111>
- Arnold, N. P., Branson, M., Kuang, Z., & Randall, D. A. (2015). MJO intensification with warming in the superparameterized CESM. *Journal of Climate*, 28(7), 2706–2724. <https://doi.org/10.1175/JCLI-D-14-00494.1>
- Arnold, N. P., & Randall, D. A. (2015). Global-scale convective aggregation: Implications for the Madden-Julian Oscillation. *Journal of Advances in Modeling Earth Systems*, 7(4), 1499–1518. <https://doi.org/10.1002/2015MS000498>
- Benedict, J. J., & Randall, D. A. (2009). Structure of the Madden-Julian oscillation in the superparameterized CAM. *Journal of the Atmospheric Sciences*, 66(11), 3277–3296.
- Bony, S., Stevens, B., Frierson, D. M. W., Jakob, C., Kageyama, M., Pincus, R., et al. (2015). Clouds, circulation and climate sensitivity. *Nature Geoscience*, 8(4), 261–268. <https://doi.org/10.1038/ngeo2398>
- Bretherton, C. S., & Khairoutdinov, M. F. (2015). Convective self-aggregation feedbacks in near-global cloud-resolving simulations of an aquaplanet. *Journal of Advances in Modeling Earth Systems*, 7(4), 1765–1787. <https://doi.org/10.1002/2015MS000499>
- Cao, G., & Zhang, G. J. (2017). Role of vertical structure of convective heating in MJO simulation in NCAR CAM 5.3. *Journal of Climate*, 30(18), 7423–7439. <https://doi.org/10.1175/JCLI-D-16-0913.1>
- Coppin, D., & Bony, S. (2015). Physical mechanisms controlling the initiation of convective self-aggregation in a general circulation model. *Journal of Advances in Modeling Earth Systems*, 7(4), 2060–2078. <https://doi.org/10.1002/2015MS000571>
- Couvreux, F., Roehrig, R., Rio, C., Lefebvre, M. P., Caian, M., Komori, T., et al. (2015). Representation of daytime moist convection over the semi-arid tropics by parametrizations used in climate and meteorological models. *Quarterly Journal of the Royal Meteorological Society*, 141(691), 2220–2236. <https://doi.org/10.1002/qj.2517>
- Dahl, G. E., Yu, D., Deng, L., & Acero, A. (2011). Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition. *IEEE Transactions on Audio, Speech and Language Processing*, 20(1), 30–42. <https://doi.org/10.1109/TASL.2011.2134090>
- Daleu, C. L., Plant, R. S., Woolnough, S. J., Sessions, S., Herman, M. J., Sobel, A., et al. (2015). Intercomparison of methods of coupling between convection and large-scale circulation: 1. Comparison over uniform surface conditions. *Journal of Advances in Modeling Earth Systems*, 7(4), 1576–1601. <https://doi.org/10.1002/2015MS000468>
- Daleu, C. L., Plant, R. S., Woolnough, S. J., Sessions, S., Herman, M. J., Sobel, A., et al. (2016). Intercomparison of methods of coupling between convection and large-scale circulation: 2. Comparison over nonuniform surface conditions. *Journal of Advances in Modeling Earth Systems*, 8(1), 387–405. <https://doi.org/10.1002/2015MS000570>
- D’Andrea, F., Gentine, P., Betts, A. K., & Lintner, B. R. (2014). Triggering deep convection with a probabilistic plume model. *Journal of the Atmospheric*, 71(11), 3881–3901. <https://doi.org/10.1175/JAS-D-13-0340.1>
- Grabowski, W. W. (1999). A parameterization of cloud microphysics for long-term cloud-resolving modeling of tropical convection. *Atmospheric research*, 52(1–2), 17–41.
- Grabowski, W. W. (2001). Coupling cloud processes with the large-scale dynamics using the cloud-resolving convection parameterization (CRCP). *Journal of the Atmospheric Sciences*, 58(9), 978–997.
- Guichard, F., Petch, J. C., Redelsperger, J. L., Bechtold, P., Chaboureaud, J. P., Cheinet, S., et al. (2004). Modelling the diurnal cycle of deep precipitating convection over land with cloud-resolving models and single-column models. *Quarterly Journal of the Royal Meteorological Society*, 130(604), 3139–3172. <https://doi.org/10.1256/qj.03.145>
- Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A. R., Jaitly, N., et al. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6), 82–97. <https://doi.org/10.1109/MSP.2012.2205597>
- Hohenegger, C., Schlemmer, L., & Silvers, L. (2015). Coupling of convection and circulation at various resolutions. *Tellus Series A-Dynamic Meteorology And Oceanography*, 67(0). <https://doi.org/10.3402/tellusa.v67.26678>

- Hohenegger, C., & Stevens, B. (2016). Coupled radiative convective equilibrium simulations with explicit and parameterized convection. *Journal of Advances in Modeling Earth Systems*, 8(3), 1468–1482. <https://doi.org/10.1002/2016MS000666>
- Holloway, C. E., Woolnough, S. J., & Lister, G. M. S. (2012). Precipitation distributions for explicit versus parametrized convection in a large-domain high-resolution tropical case study. *Quarterly Journal of the Royal Meteorological Society*, 138(668), 1692–1708.
- Holloway, C. E., Woolnough, S. J., & Lister, G. M. S. (2015). The effects of explicit versus parameterized convection on the MJO in a large-domain high-resolution tropical case study. Part II: Processes leading to differences in MJO development. *Journal of the Atmospheric Sciences*, 72(7), 2719–2743. <https://doi.org/10.1175/JAS-D-14-0308.1>
- Houze, R. A. (2004). Mesoscale convective systems. *Reviews of Geophysics*, 42, RG4003. <https://doi.org/10.1029/2004RG000150>
- Jeevanjee, N., & Romps, D. M. (2013). Convective self-aggregation, cold pools, and domain size. *Geophysical Research Letters*, 40, 994–998. <https://doi.org/10.1002/grl.50204>
- Jimenez, C., Prigent, C., & Aires, F. (2009). Toward an estimation of global land surface heat fluxes from multisatellite observations. *Journal of Geophysical Research*, 114, D06305. <https://doi.org/10.1029/2008JD011392>
- Jung, J. H., & Arakawa, A. (2010). Development of a Quasi-3D Multiscale Modeling Framework: Motivation, Basic Algorithm and Preliminary results. *Journal of Advances in Modeling Earth Systems*, 2, 11. <https://doi.org/10.3894/JAMES.2010.2.11>
- Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., et al. (2011). Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research*, 116, G00J07. <https://doi.org/10.1029/2010JG001566>
- Karabatak, M., & Ince, M. C. (2009). An expert system for detection of breast cancer based on association rules and neural network. *Expert Systems with Applications*, 36(2), 3465–3469. <https://doi.org/10.1016/j.eswa.2008.02.064>
- Khairoutdinov, M., & Randall, D. (2006). High-resolution simulation of shallow-to-deep convection transition over land. *Journal of the Atmospheric Sciences*, 63(12), 3421–3436. <https://doi.org/10.1175/JAS3810.1>
- Khairoutdinov, M., Randall, D., & DeMott, C. (2005). Simulations of the atmospheric general circulation using a cloud-resolving model as a superparameterization of physical processes. *Journal of the Atmospheric Sciences*, 62(7), 2136–2154. <https://doi.org/10.1175/JAS3453.1>
- Khairoutdinov, M. F., Krueger, S. K., Moeng, C.-H., Bogenschutz, P. A., & Randall, D. A. (2009). Large-eddy simulation of maritime deep tropical convection. *Journal of Advances in Modeling Earth Systems*, 2, 15. <https://doi.org/10.3894/JAMES.2009.1.15.S1>
- Khairoutdinov, M. F., & Randall, D. A. (2003). Cloud resolving modeling of the ARM summer 1997 IOP: Model formulation, results, uncertainties, and sensitivities. *Journal of the Atmospheric Sciences*, 60(4), 607–625.
- Khan, J., Wei, J. S., Ringner, M., Saal, L. H., & Ladanyi, M. (2001). Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nature*, 411(6837), 522. <https://doi.org/10.1038/35079160>
- Khouider, B., Biello, J., & Majda, A. J. (2010). A stochastic multicloud model for tropical convection. *Communications in Mathematical Sciences*, 8(1), 187–216.
- Khouider, B., & Majda, A. (2006). A simple multicloud parameterization for convectively coupled tropical waves. Part I: Linear analysis. *Journal of the Atmospheric Sciences*, 63(4), 1308–1323. <https://doi.org/10.1175/JAS3677.1>
- Khouider, B., Majda, A., & Katsoulakis, M. (2003). Coarse-grained stochastic models for tropical convection and climate. *Proceedings of the National Academy of Sciences of the United States of America*, 100(21), 11,941–11,946. <https://doi.org/10.1073/pnas.1634951100>
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization, *arXiv*, cs. LG.
- Kolassa, J., Aires, F., Polcher, J., Prigent, C., Jimenez, C., & Pereira, J. M. (2013). Soil moisture retrieval from multi-instrument observations: Information content analysis and retrieval methodology. *Journal of Geophysical Research: Atmospheres*, 118, 4847–4859. <https://doi.org/10.1029/2012JD018150>
- Kolassa, J., Gentine, P., Prigent, C., & Aires, F. (2016). Soil moisture retrieval from AMSR-E and ASCAT microwave observation synergy. Part 1: Satellite data analysis. *Remote Sensing of Environment*, 173(C), 1–14. <https://doi.org/10.1016/j.rse.2015.11.011>
- Kolassa, J., Gentine, P., Prigent, C., Aires, F., & Alemohammad, S. H. (2017). Soil moisture retrieval from AMSR-E and ASCAT microwave observation synergy. Part 2: Product evaluation. *Remote Sensing of Environment*, 195, 202–217. <https://doi.org/10.1016/j.rse.2017.04.020>
- Kolassa, J., Reichle, R. H., & Draper, C. S. (2017). Merging active and passive microwave observations in soil moisture data assimilation. *Remote Sensing of Environment*, 191(C), 117–130. <https://doi.org/10.1016/j.rse.2017.01.015>
- Kooperman, G. J., Pritchard, M. S., Burt, M. A., Branson, M. D., & Randall, D. A. (2016a). Impacts of cloud superparameterization on projected daily rainfall intensity climate changes in multiple versions of the Community Earth System Model. *Journal of Advances in Modeling Earth Systems*, 8(4), 1727–1750. <https://doi.org/10.1002/2016MS000715>
- Kooperman, G. J., Pritchard, M. S., Burt, M. A., Branson, M. D., & Randall, D. A. (2016b). Robust effects of cloud superparameterization on simulated daily rainfall intensity statistics across multiple versions of the Community Earth System Model. *Journal of Advances in Modeling Earth Systems*, 8, 140–165. <https://doi.org/10.1002/2015MS000574>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Mapes, B., & Neale, R. (2011). Parameterizing convective organization to escape the entrainment dilemma. *Journal of Advances in Modeling Earth Systems*, 3(2), M06004. <https://doi.org/10.1029/2011MS000042>
- Medeiros, B., Stevens, B., & Bony, S. (2014). Using aquaplanets to understand the robust responses of comprehensive climate models to forcing. *Climate Dynamics*, 44(7–8), 1957–1977. <https://doi.org/10.1007/s00382-014-2138-0>
- Miao, C., Ashouri, H., Hsu, K.-L., Sorooshian, S., & Duan, Q. (2015). Evaluation of the PERSIANN-CDR daily rainfall estimates in capturing the behavior of extreme precipitation events over China. *Journal of Hydrometeorology*, 16(3), 1387–1396. <https://doi.org/10.1175/JHM-D-14-0174.1>
- Moazami, S., Golian, S., Kavianpour, M. R., & Hong, Y. (2013). Comparison of PERSIANN and V7 TRMM Multi- satellite Precipitation Analysis (TMPA) products with rain gauge data over Iran. *International Journal of Remote Sensing*, 34(22), 8156–8171. <https://doi.org/10.1080/01431161.2013.833360>
- Moncrieff, M. W. (2010). The multiscale organization of moist convection and the intersection of weather and climate. In D.-Z. Sun & F. Bryan (Eds.), *Climate Dynamics: Why Does Climate Vary? Geophysical Monograph Series* (pp. 3–26). Washington, DC: American Geophysical Union. <https://doi.org/10.1029/2008GM000838>
- Moncrieff, M. W., & Liu, C. (2006). Representing convective organization in prediction models by a hybrid strategy. *Journal of the Atmospheric Sciences*, 63(12), 3404–3420. <https://doi.org/10.1175/JAS3812.1>
- Moncrieff, M. W., Liu, C., & Bogenschutz, P. (2017). Simulation, modeling and dynamically based parameterization of organized tropical convection for global climate models. *Journal of the Atmospheric Sciences*, 74(5), 1363–1380. <https://doi.org/10.1175/JAS-D-16-0166.1>
- Moncrieff, M. W., Waliser, D. E., Miller, M. J., Shapiro, M. E., Asrar, G., & Caughey, J. (2012). Multiscale convective organization and the YOTC Virtual Global Field Campaign. *Bulletin of the American Meteorological Society*, 93(8), 1171–1187. <https://doi.org/10.1175/BAMS-D-11-00233.1>

- Muller, C., & Bony, S. (2015). What favors convective aggregation and why? *Geophysical Research Letters*, 42, 5626–5634. <https://doi.org/10.1002/2015GL064260>
- Nie, J., Shaevitz, D. A., & Sobel, A. H. (2016). Forcings and feedbacks on convection in the 2010 Pakistan flood: Modeling extreme precipitation with interactive large-scale ascent. *Journal of Advances in Modeling Earth Systems*, 8(3), 1055–1072. <https://doi.org/10.1002/2016MS000663>
- Parishani, H., Pritchard, M. S., Bretherton, C. S., Wyant, M. C., & Khairoutdinov, M. (2017). Toward low cloud-permitting cloud superparameterization with explicit boundary layer turbulence. *Journal of Advances in Modeling Earth Systems*, 9(3), 1542–1571. <https://doi.org/10.1002/2017MS000968>
- Peters, K., Jakob, C., Davies, L., Khouider, B., & Majda, A. J. (2013). Stochastic behavior of tropical convection in observations and a multcloud model. *Journal of the Atmospheric Sciences*, 70(11), 3556–3575. <https://doi.org/10.1175/JAS-D-13-031.1>
- Pritchard, M. S., Bretherton, C. S., & Demott, C. A. (2014). Restricting 32–128 km horizontal scales hardly affects the MJO in the Superparameterized Community Atmosphere Model v3.0 but the number of cloud-resolving grid columns constrains vertical mixing. *Journal of Advances in Modeling Earth Systems*, 6(3), 723–739. <https://doi.org/10.1002/2014MS000340>
- Pritchard, M. S., Moncrieff, M. W., & Somerville, R. C. J. (2011). Orographic propagating precipitation systems over the United States in a global climate model with embedded explicit convection. *Journal of the Atmospheric Sciences*, 68(8), 1821–1840. <https://doi.org/10.1175/2011JAS3699.1>
- Pritchard, M. S., & Somerville, R. C. J. (2009). Assessing the diurnal cycle of precipitation in a multi-scale climate model. *Journal of Advances in Modeling Earth Systems*, 2, 12. <https://doi.org/10.3894/JAMES.2009.1.12>
- Qin, H., Pritchard, M. S., Kooperman, G. J., & Parishani, H. (2018). Global Effects of Superparameterization on Hydrothermal Land-Atmosphere Coupling on Multiple Timescales. *Journal of Advances in Modeling Earth Systems*, 10(2), 530–549.
- Randall, D. A. (2013). Beyond deadlock. *Geophysical Research Letters*, 40, 5970–5976. <https://doi.org/10.1002/2013GL057998>
- Rochetin, N., Couvreux, F., Grandpeix, J.-Y., & Rio, C. (2014). Deep convection triggering by boundary layer thermals. Part I: LES analysis and stochastic triggering formulation. *Journal of the Atmospheric Sciences*, 71(2), 496–514. <https://doi.org/10.1175/JAS-D-12-0336.1>
- Rochetin, N., Grandpeix, J.-Y., Rio, C., & Couvreux, F. (2014). Deep convection triggering by boundary layer thermals. Part II: Stochastic triggering parameterization for the LMDZ GCM. *Journal of the Atmospheric Sciences*, 71(2), 515–538. <https://doi.org/10.1175/JAS-D-12-0337.1>
- Satoh, M., Matsuno, T., Tomita, H., Miura, H., Nasuno, T., & Iga, S. (2008). Nonhydrostatic icosahedral atmospheric model (NICAM) for global cloud resolving simulations. *Journal of Computational Physics*, 227(7), 3486–3514. <https://doi.org/10.1016/j.jcp.2007.02.006>
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>
- Schneider, T., Teixeira, J., Bretherton, C. S., Brient, F., Pressel, K. G., Schär, C., & Siebesma, A. P. (2017). Climate goals and computing the future of clouds. *Nature Climate Change*, 7(1), 3–5. <https://doi.org/10.1038/nclimate3190>
- Sherwood, S. C., Bony, S., & Dufresne, J. L. (2014). Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature*, 505(7481), 37–42. <https://doi.org/10.1038/nature12829>
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587), 484–489. <https://doi.org/10.1038/nature16961>
- Stevens, B., & Bony, S. (2013). What are climate models missing? *Science*, 340(6136), 1053–1054. <https://doi.org/10.1126/science.1237554>
- Subramanian, A. C., & Palmer, T. N. (2017). Ensemble superparameterization versus stochastic parameterization: A comparison of model uncertainty representation in tropical weather prediction. *Journal of Advances in Modeling Earth Systems*, 9(2), 1231–1250. <https://doi.org/10.1002/2016MS000857>
- Sun, J., & Pritchard, M. S. (2016). Effects of explicit convection on global land-atmosphere coupling in the superparameterized CAM. *Journal of Advances in Modeling Earth Systems*, 8(3), 1248–1269. <https://doi.org/10.1002/2016MS000689>
- Tan, J., Jakob, C., Rossow, W. B., & Tselioudis, G. (2015). Increases in tropical rainfall driven by changes in frequency of organized deep convection. *Nature*, 519(7544), 451–454. <https://doi.org/10.1038/nature14339>
- Tao, Y., Gao, X., Hsu, K., Sorooshian, S., & Ihler, A. (2016). A deep neural network modeling framework to reduce bias in satellite precipitation products. *Journal of Hydrometeorology*, 17(3), 931–945. <https://doi.org/10.1175/JHM-D-15-0075.1>
- Taylor, C. M., Harris, P. P., & Parker, D. J. (2009). Impact of soil moisture on the development of a Sahelian mesoscale convective system: A case-study from the AMMA special observing period. *Quarterly Journal of the Royal Meteorological Society*, 136(S1), 456–470. <https://doi.org/10.1002/qj.465>
- Tulich, S. N. (2015). A strategy for representing the effects of convective momentum transport in multiscale models: Evaluation using a new superparameterized version of the Weather Research and Forecast model (SP-WRF). *Journal of Advances in Modeling Earth Systems*, 7(2), 938–962. <https://doi.org/10.1002/2014MS000417>
- Wing, A. A., & Emanuel, K. A. (2014). Physical mechanisms controlling self-aggregation of convection in idealized numerical modeling simulations. *Journal of Advances in Modeling Earth Systems*, 6(1), 59–74. <https://doi.org/10.1002/2013MS000269>
- Woelfle, M. D., Yu, S., Bretherton, C. S., & Pritchard, M. S. (2018). Sensitivity of Coupled Tropical Pacific Model Biases to Convective Parameterization in CESM1. *Journal of Advances in Modeling Earth Systems*, 10, 126–144. <https://doi.org/10.1002/2017MS001176>
- Zhou, Z. H., Jiang, Y., Yang, Y. B., & Chen, S. F. (2002). Lung cancer cell identification based on artificial neural network ensembles. *Artificial Intelligence in Medicine*, 24(1), 25–36. [https://doi.org/10.1016/S0933-3657\(01\)00094-X](https://doi.org/10.1016/S0933-3657(01)00094-X)