Machine Learning the Warm Rain Process

Andrew Gettelman[1], David John Gagne[2], Chih-Chieh Chen[3], Matthew Christensen[4], Zachary Lebo[5], Hugh Morrison[1], and Gabrielle Gantos[6]

[1]National Center for Atmospheric Research (UCAR)
[2]NCAR
[3]NCAR CGD
[4]University of Oxford
[5]Unknown
[6]National Center for Atmospheric Research

November 22, 2022

Abstract

Clouds are one of the most critical yet uncertain aspects of weather and climate prediction. The complex nature of sub-grid scale cloud processes makes traceable simulation of clouds across scales difficult (or impossible). Often models and measurements are used to develop empirical relationships for large-scale models to be computationally efficient. Machine learning provides another potential tool to improve our empirical parameterizations of clouds. To explore these opportunities, we replace the warm rain formation process in a General Circulation Model (GCM) with a detailed treatment from a bin microphysical model that causes a 400\% slowdown in the GCM. We analyze the changes in climate that result from the use of the bin microphysical calculation and find improvements in the rain onset and frequency of light rain compared to detailed models and observations. We also find a resulting change in the cloud feedback response of the model to warming, which will significantly impact the climate sensitivity. We then emulate this process with an emulator consisting of multiple neural networks that predict whether specific tendencies will be nonzero and the magnitude of the nonzero tendencies. We describe the risks of over-fitting, extrapolation, and linearization of a non-linear problem by using perfect model experiments with and without the emulator and show we can recover the solutions with the emulators in almost all respects, and recover nearly all the speed to get simulations that perform as the detailed model, but with the computational cost of the control simulation.

Machine Learning the Warm Rain Process

A. Gettelman[1,2], D. J. Gagne[1], C.-C. Chen[1], M. W. Christensen[2], Z. J. Lebo[3], H. Morrison[1], and G. Gantos[1]

[1]National Center for Atmospheric Research, Boulder, CO, USA
[2]Department of Physics, Oxford University, Oxford, UK
[3]Department of Atmospheric Science, University of Wyoming, Laramie, WY, USA

Key Points:

- Advanced treatments of warm rain formation are added to a GCM at high computational cost
- Key warm rain metrics are improved
- Neural Networks can efficiently replicate the results of the advanced treatments with low computational cost

Corresponding author: Andrew Gettelman, `andrew@ucar.edu`

Abstract

Clouds are one of the most critical yet uncertain aspects of weather and climate prediction. The complex nature of sub-grid scale cloud processes makes traceable simulation of clouds across scales difficult (or impossible). Often models and measurements are used to develop empirical relationships for large-scale models to be computationally efficient. Machine learning provides another potential tool to improve our empirical parameterizations of clouds. To explore these opportunities, we replace the warm rain formation process in a General Circulation Model (GCM) with a detailed treatment from a bin microphysical model that causes a 400% slowdown in the GCM. We analyze the changes in climate that result from the use of the bin microphysical calculation and find improvements in the rain onset and frequency of light rain compared to detailed models and observations. We also find a resulting change in the cloud feedback response of the model to warming, which will significantly impact the climate sensitivity. We then emulate this process with an emulator consisting of multiple neural networks that predict whether specific tendencies will be nonzero and the magnitude of the nonzero tendencies. We describe the risks of over-fitting, extrapolation, and linearization of a non-linear problem by using perfect model experiments with and without the emulator and show we can recover the solutions with the emulators in almost all respects, and recover nearly all the speed to get simulations that perform as the detailed model, but with the computational cost of the control simulation.

Plain Language Summary

Cloud processes are perhaps the most critical and uncertain processes for weather and climate prediction. The complex nature of clouds and their variation at small spacial scales makes simulation of clouds difficult. There exist many observations and detailed simulations of clouds that are used to develop and evaluate larger-scale models. Many times these models and measurements are used to develop empirical relationships for large-scale models to be computationally efficient. Machine learning provides another potential tool to improve our empirical parameterizations of clouds. We replace the warm rain formation process in an earth system model with emulators that use detailed treatments from small-scale and idealized models. We target specific processes that are computationally intensive and difficult to approximate at large scales. The emulator consists of multiple neural networks that predict whether specific tendencies will be nonzero and the magnitude of the nonzero tendencies. We describe the opportunity (massive speed up of cloud process calculations) and the risks of over-fitting, extrapolation and linearization of a non-linear problem by using perfect model experiments with and without the emulator.

1 Introduction

Clouds are one of the most critical yet uncertain aspects of weather and climate prediction. The complex nature of sub-grid scale cloud processes makes traceable simulation of clouds across scales difficult (or impossible). There exist many observations and detailed simulations of clouds that are used to develop and evaluate larger-scale models. Many times these models and measurements are used to develop empirical relationships for large-scale models to be computationally efficient, because using more detailed treatments is computationally prohibitive. Machine learning provides another potential tool to improve such parameterizations, by using detailed models either off-line or on-line and then building emulators for them to reduce simulation time (Krasnopolsky et al., 2005). Here we present a comprehensive investigation of replacing the warm rain formation process in an earth system model with on-line emulators that use detailed treatments from small-scale and idealized models to represent key cloud microphysical processes.

The warm rain formation process is critical for weather and climate prediction and governs the location, intensity, and duration of rainfall events, critical for weather and the hydrologic cycle. Rain formation also affects cloud lifetime and Cloud Radiative Effects (CRE), making it critical for predicting climate (Twomey, 1977; Albrecht, 1989). The specific process of rain formation is altered by the microphysical properties of clouds, making warm rain formation (with no ice involved) dependent on the size distribution of cloud drops, and thus ultimately susceptible to changes in the distribution of aerosol particles that act as Cloud Condensation Nuclei (CCN).

Ice of course will complicate the precipitation process. Supercooled liquid drops can exist, and these will either precipitate in a similar manner to warm precipitation or subsequently may freeze once they are rain drops. Cloud droplets may also freeze and form ice crystals, which precipitate and collect liquid, freezing or riming as they fall. We will not concern ourselves in this work with processes involving (or potentially involving) ice. This of course is a critical issue for weather (Forbes & Ahlgrimm, 2014) and climate (Gettelman et al., 2019; Bodas-Salcedo et al., 2019) prediction, but is beyond the scope of this initial proof of concept.

The representation of rain formation in clouds involves the interaction of a population of hydrometeors. For warm clouds, the process is one of condensation, and then collision and coalescence, the latter usually calculated in models by solving the quasi-stochastic collection equation Pruppacher & Klett (1997). This treatment neglects correlations and fluctuations that impact collision/collection (Grabowski et al., 2019), and thus cannot capture stochastic impacts on rain formation like "lucky drops" that might be important for warm rain formation (Kostinski & Shaw, 2005; Wilkinson, 2016). We prefer the term "quasi-stochastic collection" rather than "stochastic collection" as the equation has sometimes been referred to previously). The quasi-stochastic collection process describes how each size particle interacts with other sizes. Quasi-stochastic collection can result in bimodal distributions of hydrometeors (or, at least, distributions that cannot be well represented by a single mode gamma function); these modes are usually termed 'cloud' (small) and 'rain' (large) drops. Inherently, the processes evolving cloud droplets and rain drops are different. For example, cloud droplets grow primarily by condensation whereas raindrops grow primarily by collision-coalescence. Moreover, sedimentation is in the Stokes' regime for cloud droplets but not so for rain. Thus, in nearly all bulk microphysics schemes the cloud and raindrop populations are modeled using separate distributions.

The quasi-stochastic collection process is computationally expensive to treat directly in large-scale global models for weather and climate prediction. It requires the pre-computation of a collection kernel for how different sizes of hydrometeors will interact due to differential fall speeds, and it requires tracking populations of drops discretized by size bins. The tracking and advection of at least order 60 different bin quantities for liquid and ice combined makes bin schemes computationally expensive. Moreover, there is a conceptual mismatch in using detailed and computationally costly representations of microphysics in large-scale models that cannot resolve cloud- or mesoscale motions, although this may become less of a concern with increasing resolution of global models with even convection-permitting (order km-scale) global simulations now becoming routine. Thus, traditionally, large-scale models with bulk microphysics have treated the quasi-stochastic collection process of warm rain formation in a heavily parameterized fashion (Khairoutdinov & Kogan, 2000; Seifert & Beheng, 2001). For conceptual simplicity, the process is often broken up into two processes. Autoconversion is the transition of cloud drops into rain as part of a cloud droplet distribution grows to large sizes. Methods for determining autoconversion and accretion in bulk schemes vary widely. For instance, the earliest approaches were based on heuristics, requiring a threshold cloud water to be exceeded for autoconversion to commence, with "continuous collection" assumed for accretion (e.g., Kessler, 1969). More recent approaches have fit autoconversion and accretion rates to output from bin microphysics in large eddy simulation (LES) models (e.g., Khairoutdinov & Kogan, 2000; Kogan, 2013), or have incorporated theoretical aspects (e.g., Y. Liu & Daum, 2004). Because they are the major loss mechanism for cloud water, different descriptions of these processes can result in very different model evolution and climates (e.g., Michibata & Takemura, 2015).

Because many existing formulations for autoconversion and accretion are simply empirical fits to data or other models, they are readily applicable to replacement with more sophisticated tools. Neural networks are multivariate function approximators that allow many more degrees of freedom than traditional polynomial or power-law methods, for example. They are usually trained on large data sets from observations or other models. Neural networks were first used to emulate radiation parameterizations in weather (Chevallier et al., 2000) and climate (Krasnopolsky et al., 2005) models and provided significant speedups with limited reductions in predictive accuracy. More recent work in this area has focused on emulating the effects of convection based on convection-resolving simulations (Brenowitz & Bretherton, 2018) or super-parameterization (Rasp et al., 2018; Gentine et al., 2018) with promising emulation results limited by issues with numerical model stability when running the emulator for an extended time or when running the emulator outside the training climate. Constraints on the neural network loss function, architecture, and inputs appear to assist in better performance (Beucler et al., 2020). Other machine learning frameworks, such as random forests, feature architectures that inherently conserve energy and limit predictions to within the bounds of the training data, resulting in more stable simulations (Yuval & O'Gorman, n.d.). Another path to greater numerical stability is to focus on emulating a smaller, but important, subset of the sub-grid physical processes, which we investigate in this paper.

In this work we replace the traditional empirically-fit autoconversion and accretion rates in a GCM (following the approach of Khairoutdinov & Kogan (2000)) by solving the quasi-stochastic collection equation using a detailed bin microphysical model. This approach is implemented directly on-line in the GCM microphysics. The resulting code is too computationally expensive for practical simulations (as will be shown below), so we use a neural network to then emulate the code. We pose two hypotheses:

1. Hypothesis 1: Simulating warm rain in a GCM by directly solving quasi-stochastic collection using a bin approach will greatly increase the computational cost, but will result in a qualitatively different warm rain formation processes and timing, as well as quantitatively different climate means and even emergent properties (e.g. aerosol-cloud interactions and cloud feedbacks).
2. Hypothesis 2. Machine Learning (ML) Neural Network emulators can speed up the process and reproduce the qualitative and quantitative changes seen during testing of Hypothesis 1.

The first hypothesis is independent of the neural network emulator, and tests whether there is sensitivity for climate to how the warm rain process is treated. The second question is a general test of the neural network emulator concept: can it work on-line in a 'standard' climate simulation?

The details of the model and methodology are discussed in section 2. Results for emulator performance relative to the bin code are presented in Section 3. Simulation results replacing the existing autoconversion and accretion formulations are described in section 4, including discussion of process rates, mean climate and emergent properties. Discussion and conclusions are in section 5.

2 Methods

Here we describe the model used (Section 2.1), details of the quasi-stochastic collection treatment (Section 2.2) and details of the subsequent machine learning emulator methods (Section 3).

2.1 CAM6

The Community Atmosphere Model version 6 (CAM6) is the atmospheric GCM component of the Community Earth System Model version 2 (Danabasoglu et al., 2020). CAM6 features a two-moment stratiform cloud microphysics scheme (Gettelman & Morrison, 2015; Gettelman et al., 2015, hereafter MG2) with prognostic cloud liquid, cloud ice, rain, and snow hydrometeor classes. MG2 permits ice supersaturation. CAM6 includes a physically based ice mixed phase dust ice nucleation scheme (Hoose et al., 2010) with modifications for a distribution of contact angles Wang et al. (2014), and accounts for preexisting ice in the cirrus ice nucleation of X. Liu & Penner (2005) as described by Shi et al. (2015).

MG2 is coupled to a unified moist turbulence scheme, Cloud Layers Unified by Binormals (CLUBB), developed by Golaz et al. (2002) and Larson et al. (2002) and implemented in CAM by Bogenschutz et al. (2013). CLUBB handles stratiform clouds, boundary layer moist turbulence, and shallow convective motions. CAM6 also has an ensemble plume mass flux deep convection scheme described by Zhang & McFarlane (1995) and Neale et al. (2008), which has very simple microphysics. The radiation scheme is The Rapid Radiative Transfer Model for GCMs (RRTMG) (Iacono et al., 2000).

Within the MG2 parameterization, the warm rain formation process is represented by expressions for autoconversion and accretion from (Khairoutdinov & Kogan, 2000, hereafter KK2000). KK2000 uses empirical power law fits to LES with bin-resolved microphysics to define:

$$\left(\frac{\partial q_r}{\partial t}\right)_{AUTO} = 13.5 q_c^{2.47} N_c^{-1.1} \tag{1}$$

$$\left(\frac{\partial q_r}{\partial t}\right)_{ACCRE} = 67 (q_c q_r)^{1.15} \tag{2}$$

where q_c and q_r are mass mixing ratios for condensate and rain, and N_c is the number mixing ratio of condensate. For CAM6, the autoconversion rate exponent on N_c (-1.1) and prefactor (13.5) in equation 1 have been adjusted from the original Khairoutdinov & Kogan (2000) scheme to better match observations (Gettelman et al., 2019).

2.2 Quasi-Stochastic Collection

We replace the KK2000 process rate equations with an estimate of the quasi-stochastic collection process from the Tel Aviv University (TAU) bin microphysical model. The TAU model uses a “bin” or “sectional” approach, where the drop size distribution is resolved into 35 size bins. It differs from most other microphysical codes in that it solves for two moments of the drop size distribution in each of these bins. This allows for an accurate transfer of mass between bins and alleviates anomalous drop growth (Tzivion et al., 1987). The original components were developed by Tzivion et al. (1987, 1989); Feingold et al. (1988), with later applications and development documented in Reisin et al. (1996); Stevens et al. (1996); Tzivion et al. (1999); Yin et al. (2000); Harrington et al. (2000); Lebo & Seinfeld (2011). Note that process rates are one aspect of bin microphysics, another critical aspect being that bin schemes prognose multiple microphysical variables and thus evolve hydrometeor size distributions with many degrees of freedom. On the other hand, MG2 represents the drop size distribution with only 4 degrees of freedom, corresponding to 2 bulk prognostic variables each for cloud and rain. Here we will employ the bin approach to obtain autoconversion and accertion rates, but will retain the bulk MG2 approach with 4 prognostic microphysical variables to evolve cloud and rain.

The method of application in CAM is as follows. First, we discretize the MG2 bulk size distributions for liquid and rain into number concentrations in individual bins. Liquid and rain are put in the same continuous distribution of 35 mass-doubling bins for the TAU code. Then we use this as input to the TAU code for quasi-stochastic collection, assigning the mass variables (which is needed since TAU is a two-moment bin scheme predicting number and mass in each bin) based on the mean bin mass. The quasi-stochastic collection code has 60 substeps in the 1800s GCM time step, effectively a 30s timestep to evolve the distributions. This was found to yield similar results to smaller timesteps (5s) but with more computational efficiency. The result is a revised set of 35 bins with number and concentrations in each bin. We then find a local minimum in the distribution of drop number across bins: this is always found in the case where there is rain and condensate present after the application of the collection kernel. The minimum is typically between 40 and 100 microns diameter. This minimum is used to divide the bins into liquid and rain. The total number and mass of liquid and rain is defined, and tendencies calculated as the final total mass and number resulting from the quasi-stochastic collection calculation minus the initial mass and number divided by the 1800s GCM time step. A limiter is applied to the tendencies to ensure that the final mass and number are non-zero. This estimated quasi-stochastic collection tendency is then directly applied instead of the KK200 accretion and autoconversion tendencies in the code. Nothing else is changed. MG2 couples the KK2000 process rates to the Sub-Grid Scale (SGS) distribution of cloud water, but this is not done with the TAU bin code. Considering the SGS distribution of cloud water in MG2 would be a linear scaling factor in front of the TAU process rates, and would not affect the higher order interactions, but might affect the overall loss of mass and number. We neglect this adjustment because we're looking at proof of concept for applying machine learning, not final model tuning, but SGS distributions of cloud and rain should probably be considered in the future.

The code is also set up to simulate the accretion and autoconversion rates from MG2 on the same state, and this is saved as a diagnostic. This allows a direct comparison of the original MG2 KK2000 tendency (autoconversion + accretion) with the stochastic collection tendency from the TAU code.

2.3 Simulations

CAM6 is run in a standard 0.9 x 1.25 degree (latitude and longitude) configuration with 32 levels in the vertical. Boundary conditions are climatological averages of Sea Surface Temperatures (SSTs), greenhouse gases and emissions of aerosols, and precursors appropriate for 1990-2010 (i.e., averaged around 2000). To build a training data set for the emulator, we output the instantaneous inputs and outputs from the quasi-stochastic collection code, which consists of the input state and tendencies of mass and number mixing ratios, along with air pressure and temperature. The advantage of this method is we can efficiently generate independent 4D samples (space and time) for training the emulator, of whatever size necessary. We provide output every 3 days + 3 hours, so that the local time precesses through the diurnal cycle over a month. This method is run for two years to generate approximately 250 different timesteps over different seasons and times of day at 192 lat x 288 lon x 32 levels or 1.8 million samples per time step (about 500 million training samples total).

Simulations for evaluation were run for one year with three hour instantaneous output frequency. We then ran further simulations for 9 years to estimate long-term climate impacts. To estimate anthropogenic Aerosol-Cloud Interactions (ACI), identical simulations were conducted with aerosol and precursor emissions only set back to 1850 'preindustrial' conditions. To estimate cloud feedbacks, simulations with the same forcing but with SSTs increased uniformly by +4 K were conducted following Cess (1987).

2.4 Emulation

The machine learning emulator system consists of three classifier neural networks to predict whether each tendency is non-zero and four regression neural networks to predict the magnitude of the tendency. The classifier networks for q_r and N_c predict either zero or nonzero, but the N_r classifier network predicts whether the tendency is negative, zero, or positive since self-collection results in a negative N_r tendency and autoconversion results in a positive N_r tendency. Each neural network consists of 4 fully connected hidden layers with 60 neurons in each layer, and Rectified Linear Unit (ReLU) activation functions. Each network is trained for 10 epochs (passes through the training data) with a batch size of 4096 examples. The Adam optimizer is used with a learning rate of 0.001. Binary cross-entropy is used as the loss function for the classifier neural networks, and mean squared error is used for the regression neural networks. Ridge, or L2 regularization of the hidden layer weights with a penalty weight λ of 10^{-4} helps constrain the magnitude of the neural network weights. The neural networks are trained with Tensorflow. The weights are saved to an intermediate netCDF file format that is then read into CESM using a custom-built Fortran neural network inference module.

As noted above, training data for the networks are based on simulations of CAM with the TAU bin code, using individual timestep samples in space and time as individual training events. There are about 250 time samples in 3 dimensions (latitude, longitude, pressure) for nearly 500 million individual events. With this size of training data, there was little sensitivity to the number of time samples.

3 Results-Emulator Performance

First we describe some basic metrics of emulator performance before we analyze results. The first metric is timing. We have analyzed timing statistics from simulations with the CAM6 control code, the TAU quasi-stochastic collection code, and simulations where the TAU results have been replaced with the ML emulator (TAU-ML). We use standard CESM timing metrics for the total atmosphere model cost of 9 year, 1 degree ($\sim$100km) horizontal resolution simulation. All simulations were performed using the same number of tasks and layout on the same supercomputer (taking about 1 wall clock hour per simulated year). Multiple control simulations estimate a standard deviation on the timing numbers between runs of the same code at about $\pm$5%. We find that using the TAU approach for autoconversion and accretion running in CAM results in a model run time over 4 times (+410%) longer than the control code. The emulator used to replace the TAU code in MG2 (TAU-ML) runs 8% slower than the control case. This is just beyond one standard deviation of the speed of individual years from the control and TAU-ML simulations, but may not be strictly significant.

Next we analyze if the emulator reproduces the TAU process rates it is designed to reproduce. The emulator produces a q_c and N_c tendency, with q_r and N_r being the negative of the q_c and N_c tendency respectively. Figure 1 illustrates a scatter plot of the rates between the emulated code and the underlying bin code. Focusing on the neural network versus the TAU bin code it is trained on (Figure 1, top row), the emulator does an excellent job of reproducing the training data. Most of the density is on the 1:1 line and Pearson correlation coefficients are 0.98-1.00. Note that there are some extremes and they are asymmetric, indicating that the emulator does not reproduce all the extremes exactly.

The slight asymmetries for extremes can be seen in Figure 2 as a difference in frequency for large tendency values. The PDF of the bin code is the purple line and the

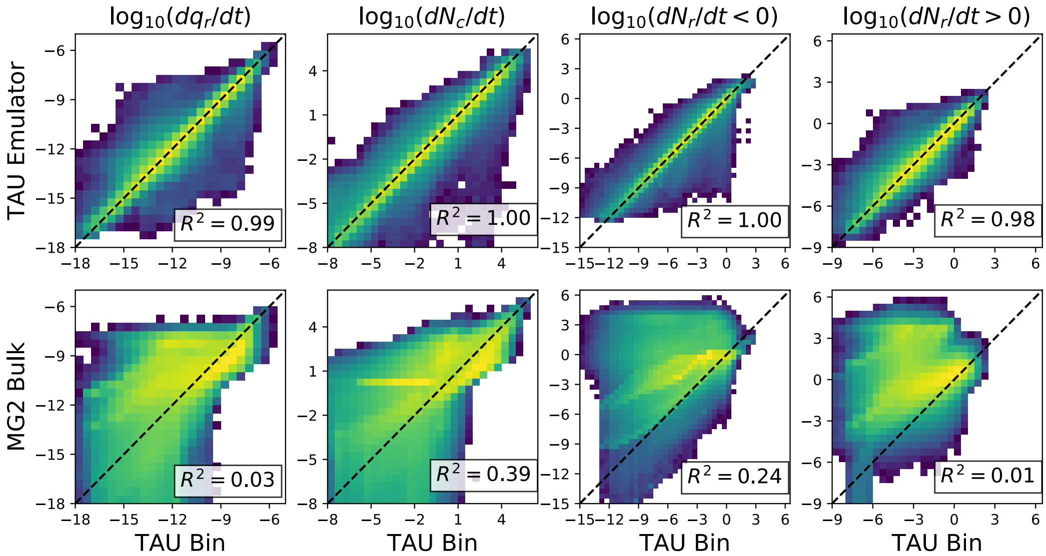

Figure 1. Frequency plots of the logarithm of TAU bin rates (horizontal) versus Neural Network Emulator (Top row) and MG2 Bulk Scheme (Bottom Row). Shown (left to right) are the rain mass tendency (dq_r/dt), condensate number tendency (dN_c/dt), negative rain number tendency ($dN_r/dt < 0$) and positive rain number tendency ($dN_r/dt > 0$). Correlation coefficient shown on the plots.

emulator PDF the blue bars. Here it is clear the emulator slightly narrows the distribution of process rates, with lower frequency of extreme high values. The distribution is narrower on the high end for dN_c/dt (Figure 2B), and positive dN_r/dt (Figure 2 D). There are some anomalies on the high end for number concentrations (Figure 1), indicating that rarely the emulator produces larger number tendencies. These are on the order of 1×10^4 s^{-1} for N_c. This represents about 9 cm^{-3} per half hour time step.

Next we look at the difference between the emulator (TAU-ML) and the TAU bin code. Since we use different simulations, we evaluate it based on monthly means at each grid location on the planet. Figure 3 illustrates the ratio of dq_c/dt between the emulator and the bin code. As expected, emulated tendencies on average are within ±20% of the bin code. White regions have values $< 1 \times 10^{-9}$ kg kg^{-1} s^{-1} (a few parts per million per minute). Lowest correspondence is in the deep tropics. This may be due to the prominence of deep convection there. But in most of the regions with significant tendencies, the process rate ratios are within ±20% (green region).

Next we examine individual process rates from the warm cloud microphysics in different regions and compare the TAU and TAU-ML codes (Figure 4). Here we use instantaneous output from the simulations. TAU (red lines) is the bin or emulator tendency for autoconversion and accretion. Over the S. Ocean (65°S-50°S, 0–360°E), in mostly supercooled liquid clouds (Gettelman et al 2020: SOCRATES paper), the emulator (TAU-ML, Figure 4B), has nearly 50% less loss of condensate to precipitation near 800 hPa than the TAU bin code it is representing at the top of shallow cloud layers. Over the Subtropical Atlantic near Barbados (10°S-25°N, 290–320°E), the TAU-ML emulator (Figure 4E) has slightly more loss at the top of the shallow cloud layer, and less below, for very little change in the vertical column. Note that none of the other process rates are significantly impacted by the shift between the TAU code and the emulator (TAU-ML).

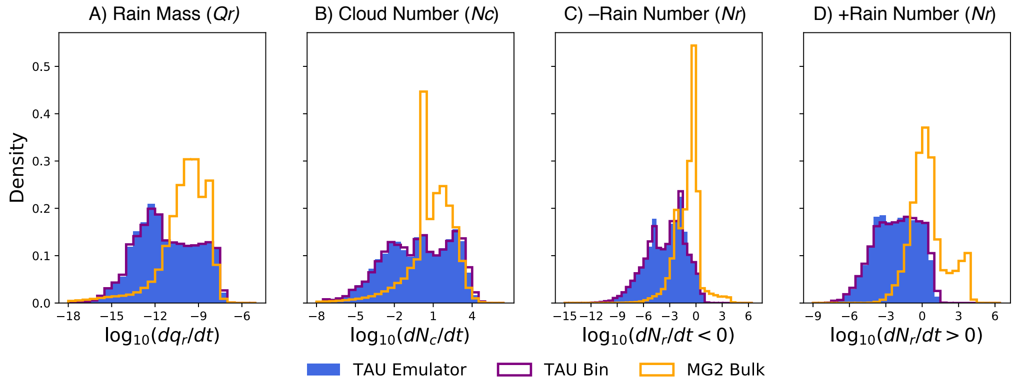

Figure 2. Probability distributions of the logarithm of process rates. A) Rain mass tendency (dq_r/dt), B) Cloud condensate number tendency (dN_c/dt), C) Negative rain number tendency ($dN_r/dt < 0$) and D) Positive rain number tendency ($dN_r/dt > 0$). The TAU bin distribution is shown in purple, the emulator (TAU-ML) solid blue and the MG2 bulk autoconversion+accretion tendencies in orange.

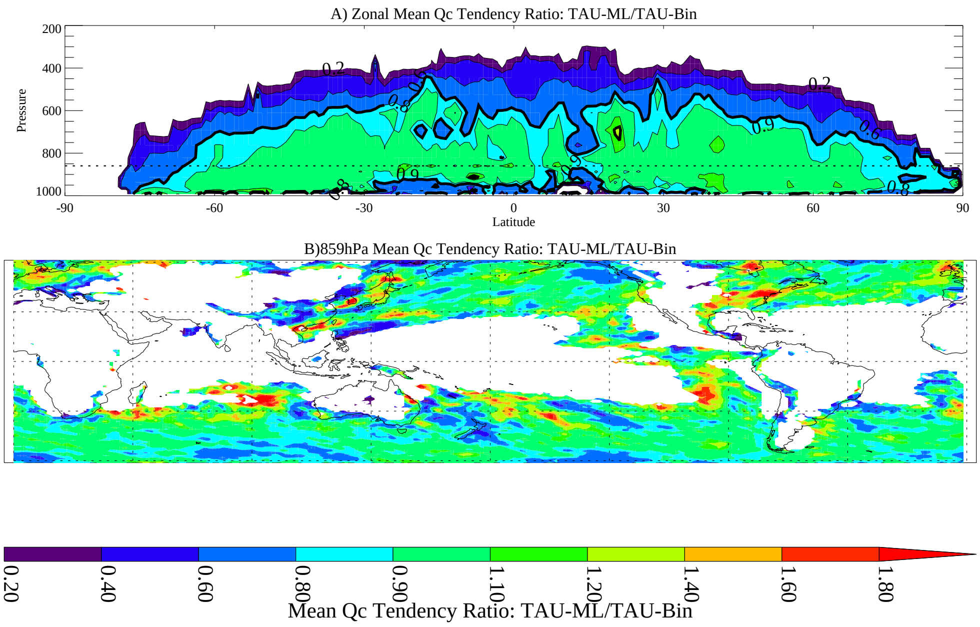

Figure 3. Ratio of q_c tendency (dq_c/dt) between the TAU-ML code and TAU bin code it is designed to reproduce. A) Zonal mean latitude-height, B) horizontal map from 60S-60N at the 859hPa level near the top of the planetary boundary layer. Blank regions have tendencies less that 1×10^{-9} kg kg^{-1} s^{-1}.

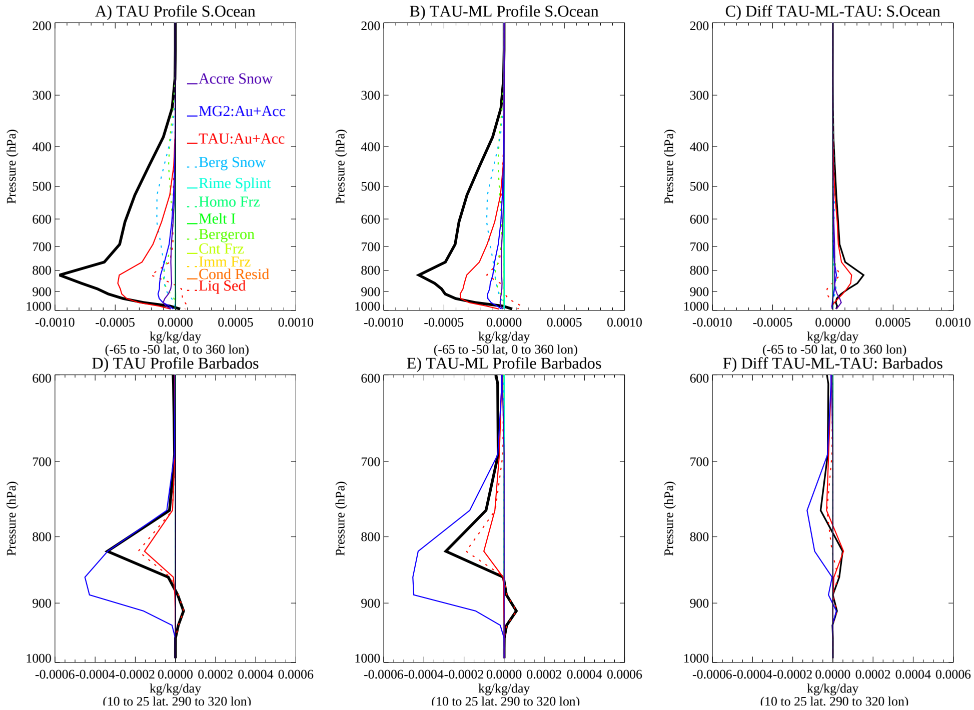

Figure 4. Mean process rates in the S. Ocean (65°S-50°S, 0–360°E, Top: A,B,C) and the ocean region around Barbados (10°S-25°N, 290–320°E, Bottom: D, E, F) regions as defined in the text. Process rates from the TAU bin code (A,C Left), Emulated code (TAU-ML B,D center) and the difference (TAU-ML minus TAU C, F right). Total of all rates is the thick black line.

When the emulator is trained on present day climate, it will occasionally produce tendencies that would result in a negative mass or number to either cloud or rain when applied to the model state. We have built the model to check for this, and correct ('fix') the mass if necessary. This is done by ensuring that the final masses and number concentrations of both cloud and rain are positive or zero, and reducing the tendencies accordingly. The fixer is not necessary in present day simulations (the climate where the emulator is trained), but if we attempt to run the emulator with a perturbed simulation (e.g., SST+4K), then the model will crash without the fixer. Figure 5 illustrates the frequency of occurrence of the mass fixer. The highest frequency occurs in the S. Hemisphere subtropics close to the surface, with another lower peak frequency in the deep tropics at 800 hPa (Figure 5A). Figure 5 shows annual means, but seasonal means are similar: the peak remains in the S. in all seasons. Figure 5B indicates that the peak fixer invocation occurs nearly 1/3 of the time in regions of trade cumulus clouds in the subtropics. These are regions with small grid box average liquid water content due to small cloud fractions, but potentially high in-cloud water contents. We return to this later when looking at derived properties of the system. It is hypothesized that we could eliminate the need for the fixer by including perturbed climate samples in the training data.

4 Results-Model Output

Having established that the emulator works to reproduce the results of the bin code, we now discuss the differences between the CAM base code (CAM6) and the code that directly solves the quasi-stochastic collection equation (TAU). We analyze first process rates and key metrics of warm rain formation, then the mean state climate, and finally

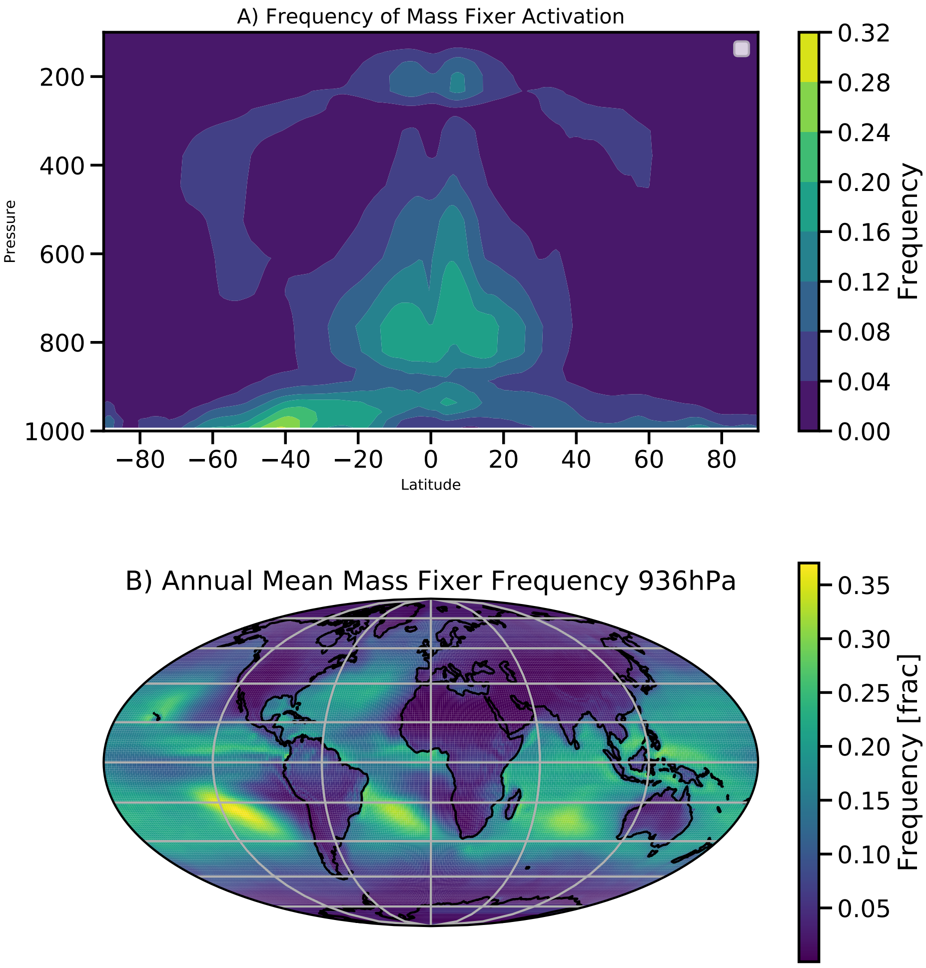

Figure 5. A) Zonal mean annual frequency of occurrence of the mass fixer in the TAU-ML code. B) Horizontal map of annual mean frequency of occurrence of the mass fixer in the TAU-ML code at 936 hPa.

emergent properties such as aerosol forcing and cloud feedbacks. We also assess whether there are further differences between TAU and TAU-ML simulations.

4.1 Process Rates

Figure 4 illustrates the liquid process rates in the TAU (A,D) and TAU ML (emulator) simulations (B,E) , and their difference (C,F), as discussed above. In each of the two simulations, the MG2 control case autoconversion and accretion rates from Khairoutdinov & Kogan (2000) are also run on the same state to generate tendencies. These are not applied to the model evolution (prognostic tendencies) but are saved diagnostically. They represent, however, how the Khairoutdinov & Kogan (2000) autoconversion and accretion would have responded to the same model state, so are valuable for comparison. Over the S. Ocean (Figure 4A and B), the TAU code (and its emulator, TAU-ML) produces a larger loss of water than the KK2000 scheme (MG2, blue). The vertical structure is similar. However, over the subtropics around Barbados (Figure 4 D and E), the Bin code and its emulator both produce much lower rates of condensate loss than MG2, and only in the upper regions of the clouds, with a peak at ∼850 hPa, rather than closer to 900 hPa found in KK2000 (Figure 4 D and E, difference between red and blue lines). This has implications for the overall climate simulations, as we will see below.

Figure 6 is similar to Figure 3, except it compares the TAU bin process rates for autoconversion and accretion to those produced by MG2 in the same simulation. As in Figure 4, the process rates are calculated in the same simulation on the same state, where 'MG2' is KK2000 and 'TAU_Bin' replaces this with the direct calculation of quasi-stochastic collection. The figure uses monthly means, but instantaneous fields yield the same results. Consistent with Figure 4, in the extra-tropics (such as the S. Ocean), the TAU Bin process rates are more negative (larger loss) than MG2, while in the sub-tropics, the rates are typically less negative (smaller loss), except in regions with high water content over the N. E. and S. E. Pacific, and S. E. Atlantic. The results are consistent with Figure 2, bottom row, where the MG2 control case has slightly more frequent high process rates ($1{\times}10^{-9}$ kg kg^{-1} s^{-1}), and fewer moderate rates ($1{\times}10^{-12}$ kg kg^{-1} s^{-1}) than the bin scheme. Note that the TAU bin code has significantly less change in rain number than the MG2 KK2000 autoconversion and accretion for either the negative (Figure 2, bottom row 2nd from left) or the positive (Figure 2, bottom row, right) cases.

We have also analyzed the onset of precipitation by looking at the average rain rate as a function of drop effective radius and Liquid Water Path (LWP). Rosenfeld et al. (2012), Figure 1, illustrate for a LES that significant rain rates are rarely seen for an effective radius (Re) <15 microns, a result they note is also found in observations. This is only for one set of cases, but Figure 7A illustrates that with the KK2000 scheme in MG2 there are significant rain rates for high LWP but small effective radius. The TAU Bin (Figure 7B and emulated TAU Bin (Figure 7C) simulations do not show this behavior: they have much lower rain rates for high LWP but Re<15 microns, in better agreement with observations and LES.

We have also examined the diurnal cycle of precipitation in the different simulations. The TAU bin and emulator code show no significant changes from the base code in the diurnal cycle of precipitation. There are small changes over ocean, but they are not significant. Over land there are no changes, likely because over land the diurnal cycle is mostly dominated by the deep convection parameterization which is not directly affected by the TAU bin code.

We do see changes in the intensity of precipitation. Figure 8 illustrates the frequency of occurrence of different rain rates in the simulations, based on 3-hourly precipitation averages. The shaded region represents one standard deviation of precipitation values over a month (typically 240 time samples) from each simulation in each intensity bin. Shown are the control case with KK2000 (blue), the TAU bin code (orange), the emu-

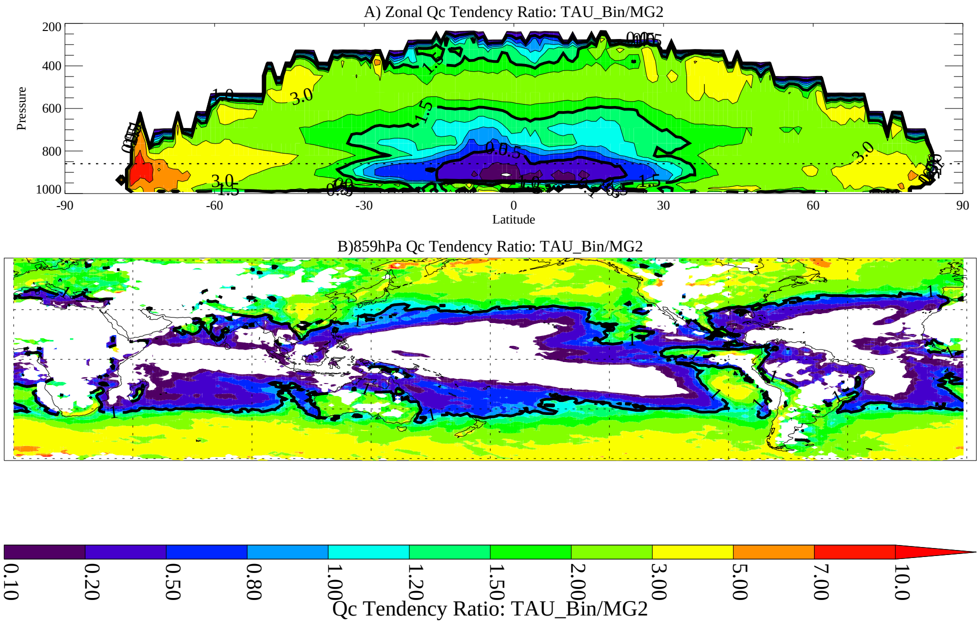

Figure 6. Ratio of Qc tendency (dQ_c/dt) between the TAU bin code and MG2 code it is designed to reproduce. A) Zonal mean latitude-height, B) horizontal map from 60°S–60°N at the 859hPa level.

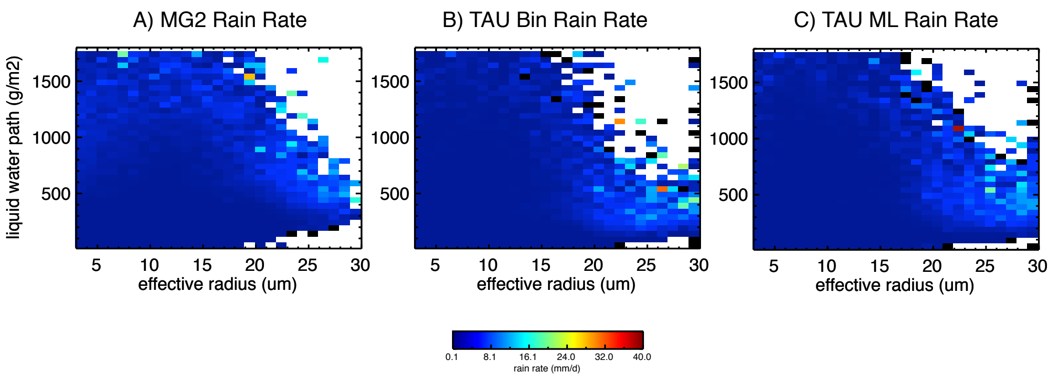

Figure 7. Contoured frequency distributions of rain rate plotted against effective radius (microns) and liquid water path (g/m^2) for MG2 control (A), TAU bin (B), and TAU-ML emulator simulations (C).

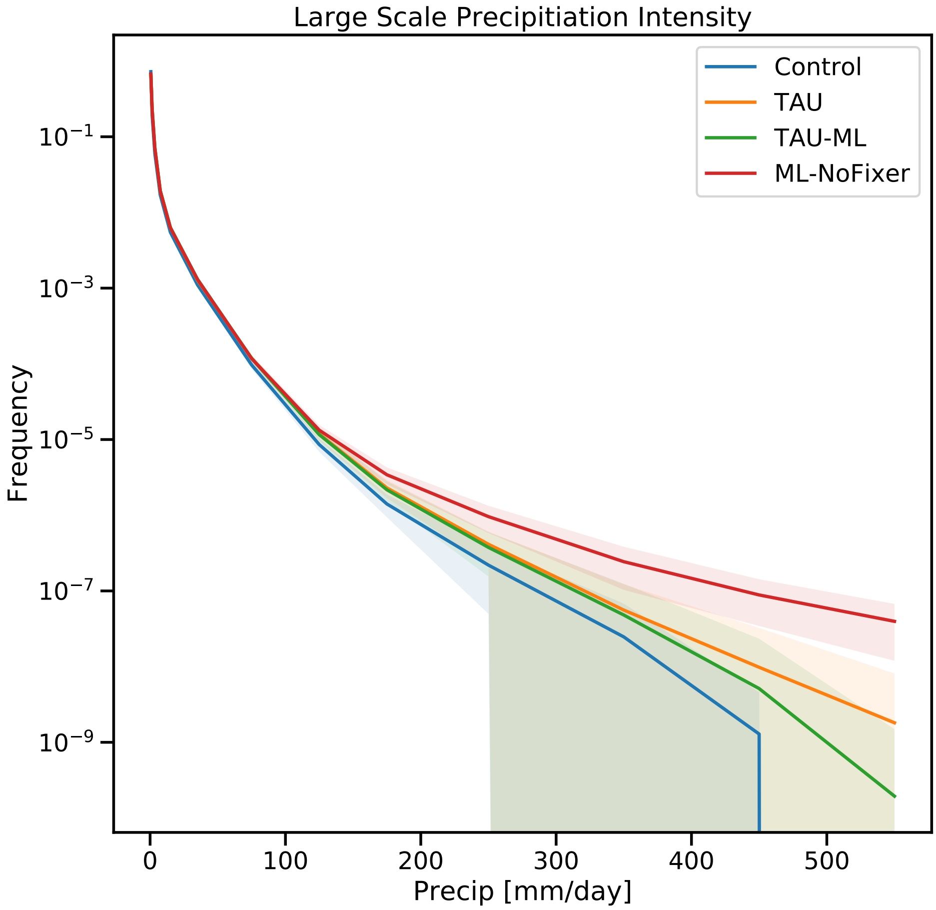

Figure 8. Frequency of occurrence of different rain rates (mm/day) in the simulations, based on 3 hourly precipitation averages. The shaded region represents one standard deviation of the monthly frequency (typically 240 time samples) from each simulation. Shown are the control case with KK2000 (blue), the TAU bin code (orange), the TAU-ML emulated code (green), and the emulator without the mass fixer (ML-NoFixer, red).

lated code (green), and the emulator without the mass fixer (red). The histogram of counts has been normalized into frequency (integral of 1). In general the bin code produces higher rain rates at large values (>200 mm/day) than the control case with KK2000. The emulator code produces a lower frequency of occurrence of these values, but only when the mass fixer is applied. Without the fixer there are significant anomalies in the ML code, and the emulator produces a small frequency of very high precipitation values.

Finally, we examine the frequency of occurrence of surface precipitation in Figure 9, and compare this to data from CloudSat. CloudSat-retrieved precipitation is averaged over 1 degree regions at monthly intervals based on the 2C-Rain (Lebsock & L'Ecuyer, 2011)and 2C-Snow Profile (Wood et al., 2014) datasets. CloudSat is upscaled by aggregating profiles along the orbit at a native resolution of 1.75 km to 111 km to match the model resolution. If at least one of the profiles has a precipitation rate greater than 0.01

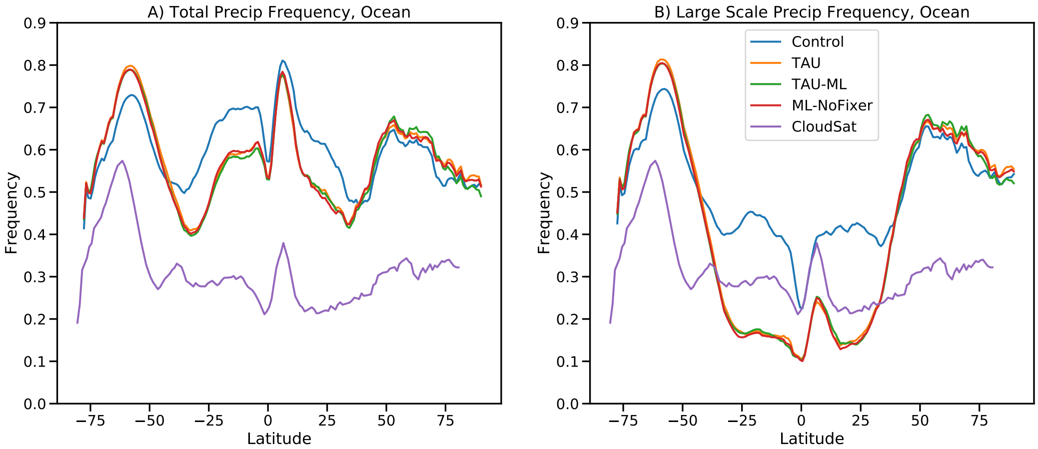

Figure 9. Annual mean frequency of occurrence of (A) Total (Large Scale + Deep Convection) and (B) Large-scale (stratiform) precipitation greater than 0.01 mm/hr. Shown are the control case with KK2000 (blue), the TAU bin code (orange), the TAU-ML emulated code (green), the emulator without the mass fixer (ML-NoFixer, red) and CloudSat observations (purple). CloudSat precipitation is obtained from 2C-Rain-Profile and 2C-Snow-Profile products as described in the text and frequency calculations follow Stephens et al. (2010).

mm/hr the whole upscaled bin is considered precipitating. This is then aggregated to 1x1 degree regions. Both day and nighttime retrievals are combined. The method is identical to Stephens et al. (2010). Model 3-hourly average precipitation is binned to the same resolution, and a threshold of 0.01mm/hr (0.24 mm/day) determines precipitating locations to match the CloudSat threshold used. CAM has too frequent total precipitation over the ocean when compared to observations from CloudSat (Figure 9A). This is common with many other models, and with earlier versions of CAM (Stephens et al., 2010). In the tropics and sub-tropics, most of the precipitation frequency is convective precipitation, as the average frequency is 0.6 for total precipitation, but only 0.2 for large scale precipitation (which includes shallow convection in CAM6). The large-scale precipitation from the MG microphysics scheme is shown in Figure 9B frequency is low, particularly in the sub-tropics and tropics (30°S–30°N). The TAU code and the TAU-ML emulator significantly reduce the frequency of large scale precipitation in the tropics and subtropics, reducing it from nearly 0.4 to 0.2, and reducing the bias in total precipitation frequency, which is largely due to the deep convective scheme which dominates total precipitation frequency between 30°S and 30°N. The mass fixer (ML-NoFixer) does not change these results. There is still too frequent large scale precipitation in the storm tracks. We have examined a higher precipitation threshold of 0.05 mm/hr and found that there is little difference in total or large scale precipitation frequency between the TAU and Control simulations, which both have more similar frequency distributions to CloudSat (see Figure S1). Thus the TAU code differences are only for drizzle, but significantly reduce the 'dreary' state of the model (Stephens et al., 2010) for light rain.

4.2 Mean Climate

Next we turn to analysis of the mean climate in the simulations, and compare the TAU-Bin code and the emulator with each other and the control model with KK2000. Figure 10 provides an overview of the mean state climate, focusing on clouds and radiation in the simulations. We have added, where available, observations from the NASA

Clouds in the Earth Radiant Energy System (CERES) Energy Balance Adjusted Flux (EBAF) product (Loeb et al., 2018), version 4.1. We focus on our two hypotheses in the questions above. First we explore whether the emulator (TAU-ML) produces the same mean climate as the TAU Bin code it is trying to emulate, and second how the TAU Bin and/or emulator climate differs from the control climate with KK2000.

The emulator is trained on instantaneous output from 2 years of data of the TAU Bin code. To evaluate the emulator, we run a further 7 years of the TAU Bin code, and compare this to 9 years with the emulator. Results indicate that the emulator has l0-15% more LWP than the TAU Bin code at most latitudes outside the tropics (Figure 10A). This is associated with slightly larger drop sizes (Figure 10D), but the same number concentration (Figure 10E) and cloud fraction (Figure 10C). Thus the change is solely in the mass of liquid, not its number. The increased LWP then results in 10% higher cloud optical depth (Figure 10F) in the storm tracks, and a corresponding small difference in SW Cloud Radiative Effect (CRE, Figure 10G). There is no change in Ice Water Path (Figure 10B) or in the LW CRE (Figure 10H). This is consistent with Figure 3 which illustrates that most of the atmosphere has a ratio of TAU-ML emulator to TAU Bin code q_c tendencies of slightly less than 1, resulting in less loss of water and more remaining cloud liquid. It is also seen in the vertical structure of process rates over the S. Ocean in Figure 4C, showing less q_c tendency.

Figure 10 also compares the emulator code with a simulation of the emulator run without the mass fixer for present day climate. These simulations are not bit-for-bit, but Figure 10 indicates that their mean climates are very similar: adding the mass fixer does not change climate. There do not appear to be appreciable differences in the S. Hemisphere subtropical regions where the mass fixer is most active. Only a few differences appear in cloud optical depth (Figure 10F) at high latitudes, likely from cases with low liquid water.

Next we compare the TAU Bin code (and TAU-ML emulator) results to the control climate with KK2000 autoconversion and accretion. The code with KK2000 (Control) has much higher LWP in the storm track latitudes of 30–60°N and S (Figure 10A) with the same cloud fraction in the storm tracks (Figure 10C). In the subtropical S. Hemisphere (30–15°S), there is similar or less LWP and lower cloud fraction in the control case than the TAU cases, and similarly lower cloud fraction in the control case in the N. Hemisphere subtropics. Effective radius is smaller with the TAU Bin or TAU-ML code (Figure 10D), while number concentration is not substantially different between any of the cases (Figure 10E). We have added for reference (where available) observations from the NASA CERES EBAF product (Loeb et al., 2018), which indicate that the reduction in LWP may be too large (though this is uncertain from satellite observations), while the changes to sub-tropical cloud fraction may be an improvement. The impact of these microphysical changes is a reduction in cloud optical depth in the storm tracks, and increases in cloud optical depth in the sub-tropics with the Bin code. This is an improvement in the subtropics, and creates similar biases of opposite sign to the control code in the extratropics (Figure 10D). The overall radiative impact is mostly on SW CRE (Figure 10G), with improvements at high latitudes, and a degradation with too much cloud forcing in the subtropics compared to observations. It is interesting that increased optical depth (but still lower than CERES) and increased cloud fraction (similar to CERES) at high latitudes yield CRE that is too strong (more negative) in the TAU BIN and TAU-ML emulator simulations. This likely results from higher cloud frequency (not just fraction) in these regions seen in Figure 9 Note that we do not necessarily expect a 'better climate' with the TAU code than the control model, since we have made major changes to the microphysics scheme and not made an effort to work to compensating biases in other parts of the cloud microphysics, macrophysics, or the turbulence scheme.

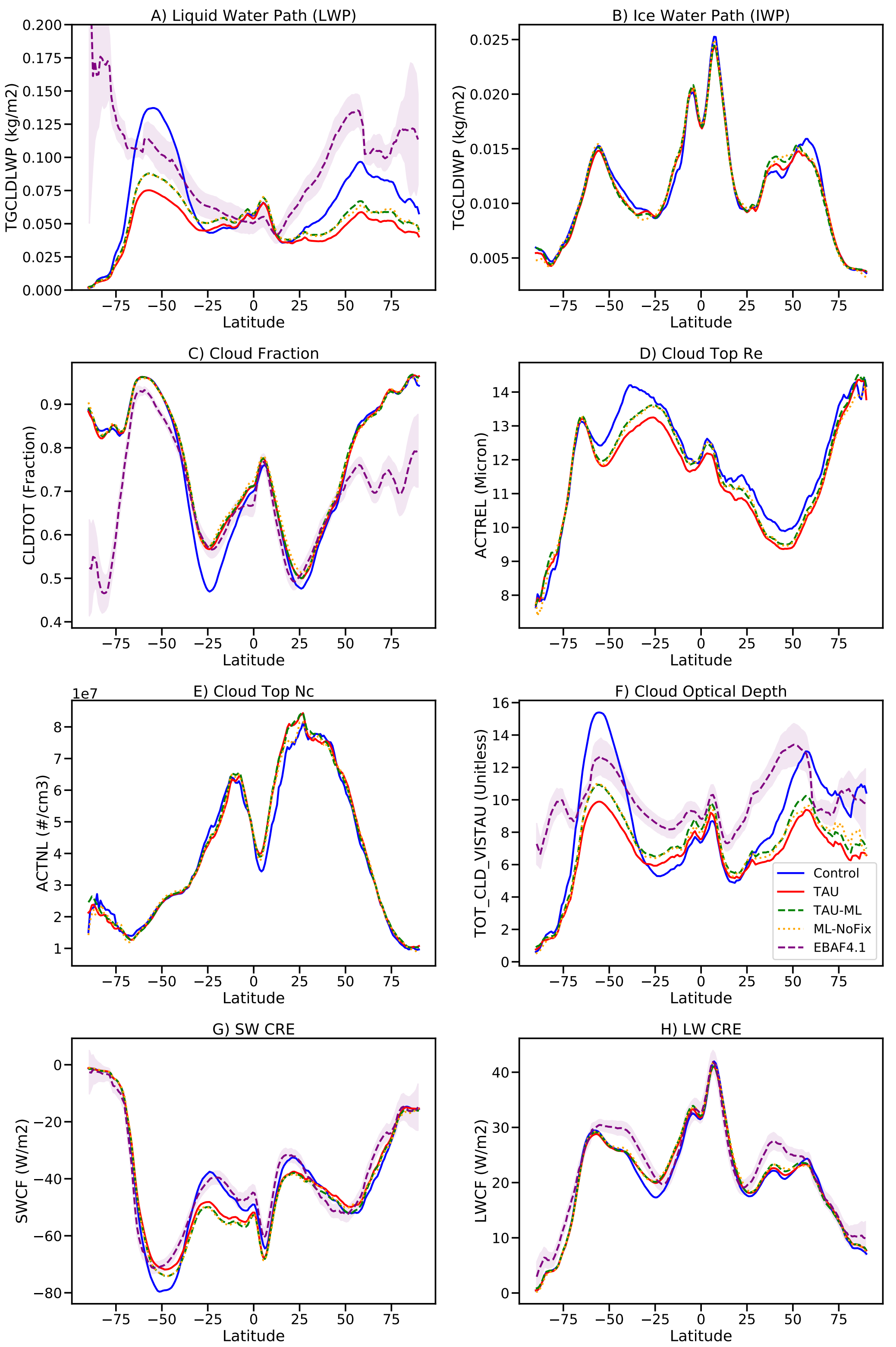

Figure 10. Zonal mean climatologies from Control (blue solid), TAU bin (red solid), TAU-ML (green dash), the emulator without the mass fixer (ML-NoFix, orange) and CERES EBAF4.1 satellite observations (purple dash). A) Liquid Water Path (LWP), B) Ice Water Path (IWP), C) Cloud fraction, D) Cloud Top Effective Radius (Re), E) Cloud Top Droplet number concentration (Nc), F) Cloud optical Depth, G) Shortwave Cloud Radiative Effect (SW CRE), and H) Longwave Cloud Radiative Effect (LW CRE). Shading shows ± 1 standard deviation of monthly anomalies for CERES data.

4.3 Emergent Properties

The warm rain formation process is critical for the mean state of clouds. It may be also critical for the response of clouds to perturbations. A very important global response is the response of cloud to changes in aerosols that nucleate cloud drops, called Aerosol Cloud Interactions (ACI). ACI result when changes in aerosols affect Cloud Condensation Nuclei (CCN) and hence cloud drop number. This results in significant radiative perturbations and adjustments of cloud microphysical processes; see Bellouin et al. (2020) for a review. ACI are the largest uncertainty in historical and present anthropogenic climate forcing (Boucher et al., 2013). Second, we look at cloud feedbacks, the response of cloud radiative effects to surface temperature changes (Gettelman & Sherwood, 2016; Stephens, 2005), which are the largest uncertainty in understanding the sensitivity of climate to forcing (Boucher et al., 2013).

Figure 11 illustrates the magnitude of ACI from the different simulations. The unit of peta-watts (1 PW=10^{15} W) per degree of latitude in Figure 11 is essentially area weighted around a latitude circle (dividing by m^2 deg^{-1} yields Wm^{-2}). ACI are calculated by running another simulation with forcing identical to the baseline cases discussed above, but with aerosol emissions set to 1850 values. All other forcings remain at 2000 levels. The differences of the two simulation climates are the ACI. The total ACI defined as the change in cloud radiative effect for the Control (CAM6) KK2000, TAU bin, and emulated (TAU-ML) simulations are -1.85, -1.95 and -2.0 Wm^{-2}, respectively. Differences result from slightly higher ACI in the TAU and TAU-ML code in the S. Hemisphere (Figure 11A). The change in LWP (Figure 11B) is slightly lower in the N. Hemisphere in the TAU Bin and emulator code, while changes in drop number concentration are nearly identical (Figure 11C). On the whole, these differences are not significantly different from each other, given the variance of annual radiation differences by latitude.

Finally we examine cloud feedbacks. Feedbacks are estimated as the radiative kernel adjusted change in cloud radiative effect as defined by Soden et al. (2008), with kernels from Shell et al. (2008) used as in Gettelman & Sherwood (2016). For this analysis we use a uniform +4 K perturbation in the sea surface temperature from the control case, a standard metric used in many studies following Cess (1987). Figure 12 illustrates the results for (A) SW, (B) LW, and (C) net cloud feedback. Feedbacks are comparable, except over the S. Ocean where the control CAM6 case with KK2000 has much higher positive SW feedbacks, resulting in a significantly higher net cloud feedback. The globally integrated zonal mean net cloud feedback is 0.81 $Wm^{-2}K^{-1}$ for CAM6, 0.77 $Wm^{-2}K^{-1}$ for TAU Bin, and 0.70 $Wm^{-2}K^{-1}$ for TAU ML.

The bin code and emulator have significantly lower SW cloud feedbacks over the S. Ocean. This region was identified by Gettelman et al. (2019) as a region where cloud feedbacks increased significantly in CAM6 due to better representation of cloud phase as supercooled liquid clouds. This removed a negative cloud phase feedback (Tan et al., 2016) and increased cloud feedback and climate sensitivity. The feedback in the emulator code is reduced because the ice fraction extends lower in the atmosphere over the S. Ocean. The lower liquid water path (Figure 10A) with the same ice water path (Figure 10B) increases the ice fraction and results in reduced positive cloud feedback. Note that the overall radiative effect in the S. Ocean is improved in the TAU code simulations (Figure 10G), indicating this change might reduce biases and improve realism. Comparisons between CAM6 and in-situ aircraft observations over the S. Ocean by Gettelman et al. (2020) indicates that CAM6 has too little ice, indicating that the increased ice fraction may be more realistic.

5 Conclusions

We return to our hypotheses and note two significant conclusions.

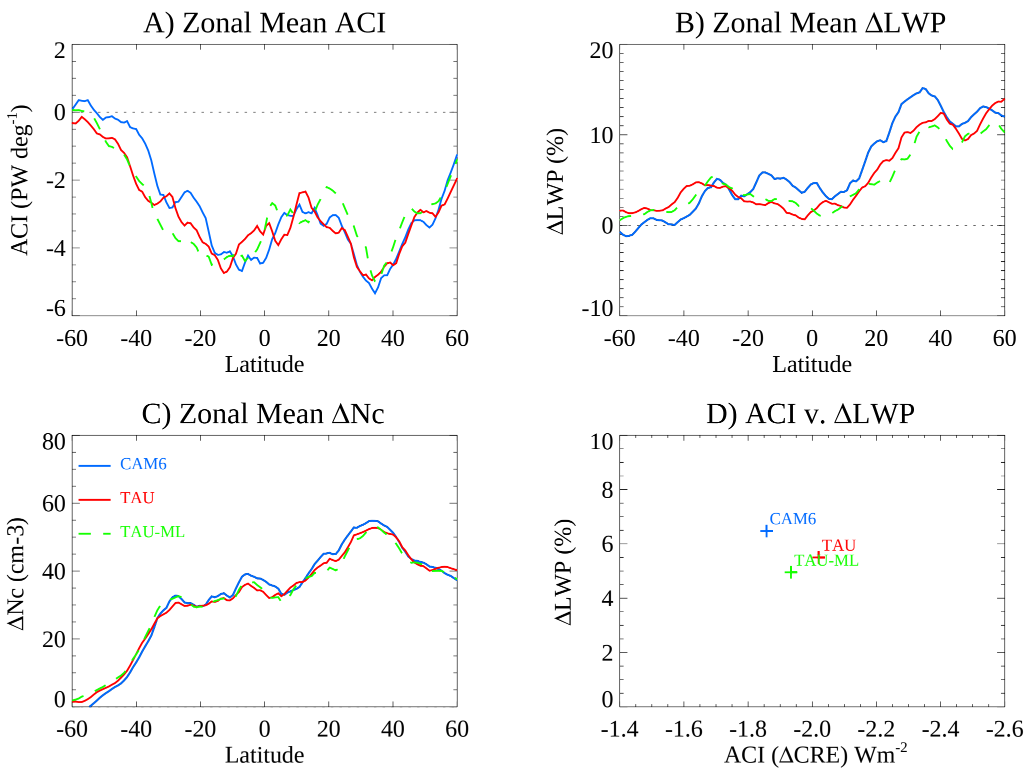

Figure 11. A) Zonal mean Aerosol Cloud Interactions (ACI) in PW deg^{-1}, B) Zonal mean percent change in LWP, C) Zonal mean change in droplet number concentration (Nc), and D) Global-average ACI (Wm^{-2} v. percent change in LWP. Simulations shown: CAM6 control (blue solid), TAU bin code (red solid), and TAU-ML emulator (green dash).

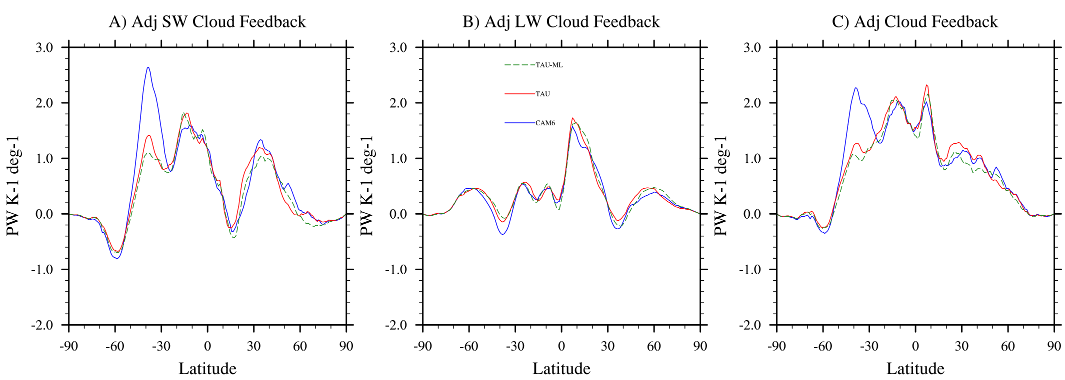

Figure 12. Zonal mean kernel adjusted cloud feedbacks. A) SW, B) LW, and C) Net. Simulations: CAM6 (blue solid), TAU bin (red solid), and TAU ML emulator (Green Dash).

First, the TAU-ML emulator code is able to reproduce almost all the metrics of the TAU bin formulation and recover almost all the computational cost penalty. There are some differences in mean climate, mostly in liquid water path. The differences between TAU-ML and TAU result from the emulator underestimating the mean process rates (Figure 4 and Figure 3), so the LWP is slightly higher in the emulator code. A mass fixer was found to be needed to get the code to run stably for perturbed climates. The fixer is active in sub-tropical regions with low mean liquid water path. This does not appear to impact any of the climate results or emergent properties like cloud feedbacks or aerosol cloud interactions. However, the emulator does have excessive precipitation intensity if the mass fixer is not applied. We note that mass fixers are currently applied to other process rates in the code as well, since the microphysical process rates are process split (they operate on the same state and are then combined). It is further hypothesized (but left for future work) that training on a combined data set including perturbed climate simulations might reduce or eliminate the need for the fixer. Thus the emulator may be a useful option for simulating process rates, with the noted caveats above.

Second, the use of the TAU bin quasi-stochastic collection process results in a different climate simulation from the base code with KK2000. The overall climate shows more significant biases than the well tuned and adjusted code with KK2000, but there are several interesting and important features. First, the onset of precipitation is improved significantly, and mostly occurs only when mean drop size is large, in agreement with bin microphysics in LES models and observations (Rosenfeld et al., 2012). Second, the frequency of occurrence of light large-scale precipitation (drizzle) in the sub-tropics and tropics is significantly lower with the TAU code or TAU-ML emulator of the code than the control case. This is a substantial improvement and reduces a longstanding and common model bias. It is likely coupled to the precipitation onset being limited to large mean drop size. Third, the bin code modifies the intensity of precipitation for extreme but low-frequency events. The emulator is able to reproduce all of these differences with the control simulation. Finally, the mean state in the sub-tropics degrades (with respect to radiative effects) but it improves over the S. Ocean. This is due mostly to improvements with lower LWP, resulting in a different balance of ice and liquid in S. Ocean cloud systems.

There are two important conclusions related to emergent properties. First, the change in S. Ocean clouds results in a significant drop in cloud feedback strength, which would significantly impact the climate sensitivity. CAM6 has been seen to have little ice when compared in detail to in-situ observations of S. Ocean supercooled liquid clouds (Gettelman et al., 2020), so this may be an improvement. Further analysis here would be very useful and critical for understanding cloud feedbacks.

Second, the cloud radiative effects of aerosols (ACI) are virtually unchanged in the simulations, despite very different warm rain process rates and a very different representation of the interaction of drop number with rain formation. Recent work has focused on the importance of rain formation and autoconversion and accretion as governing ACI (Bellouin et al., 2020; Gettelman, 2015). However, this work shows little sensitivity of ACI to the formulation of autoconversion and accretion. CAM6 still sees a large LWP response (cloud adjustment) to increased aerosols, despite a very different dependence of the warm rain process on drop number. This argues that the cloud adjustments to aerosols might be damped by other buffering processes (Stevens & Feingold, 2009), such as interactions with turbulence or entrainment. Recently, Karset et al. (2020) linked drop number to turbulent entrainment in CAM5 and did not find significant sensitivity of ACI to parameterized turbulent entrainment. Further investigation of the interaction of turbulence and cloud macrophysics with aerosols is warranted, as cloud adjustments to aerosols and ACI in CAM6 do not seem sensitive to the microphysical representation of the warm rain process. This result should be tested in other modeling systems.

Finally, it is important to note there is uncertainty in the bin microphysics calculations. Different bin schemes will do different things, depending on the specified collection efficiencies, numerics, bin resolution, etc. This needs to be kept in mind when thinking of bin schemes as a "benchmark".

In summary, machine learning emulators do appear to provide useful speedups with accurate representations of complex microphysical processes that can be used to provide insight into important uncertainties in climate models. Methods do require bespoke development of emulators, and also require more inputs that would seem to be required to accurately predict and simulate model evolution.

Acknowledgments

The National Center for Atmospheric Research is sponsored by the United States National Science Foundation. Simulation output is available on the Earth System Grid. DOIs for summary simulation output will be put here in proof stage.

References

Albrecht, B. A. (1989). Aerosols, cloud microphysics and fractional cloudiness. *Science*, *245*, 1227–1230.

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., ... Stevens, B. (2020). Bounding Global Aerosol Radiative Forcing of Climate Change. *Reviews of Geophysics*, *58*(1), e2019RG000660. Retrieved 2020-06-08, from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019RG000660` (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019RG000660) doi: 10.1029/2019RG000660

Beucler, T., Pritchard, M., Gentine, P., & Rasp, S. (2020, February). Towards Physically-consistent, Data-driven Models of Convection. *arXiv:2002.08525 [physics]*. Retrieved 2020-02-28, from `http://arxiv.org/abs/2002.08525` (arXiv: 2002.08525)

Bodas-Salcedo, A., Mulcahy, J. P., Andrews, T., Williams, K. D., Ringer, M. A., Field, P. R., & Elsaesser, G. S. (2019). Strong Dependence of Atmospheric Feedbacks on Mixed-Phase Microphysics and Aerosol-Cloud Interactions in HadGEM3. *Journal of Advances in Modeling Earth Systems*, *11*(6), 1735–1758. Retrieved 2019-09-25, from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001688` doi: 10.1029/2019MS001688

Bogenschutz, P. A., Gettelman, A., Morrison, H., Larson, V. E., Craig, C., & Schanen, D. P. (2013). Higher-order turbulence closure and its impact on Climate Simulation in the Community Atmosphere Model. *Journal of Climate*, *26*(23), 9655–9676. doi: 10.1175/JCLI-D-13-00075.1

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., ... Zhang, X. Y. (2013). Clouds and Aerosols. In T. F. Stocker et al. (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge Universtiy Press.

Brenowitz, N. D., & Bretherton, C. S. (2018). Prognostic validation of a neural network unified physics parameterization. *Geophysical Research Letters*, *45*(12), 6289-6298. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL078510` doi: 10.1029/2018GL078510

Cess, R. D. (1987). Exploratory studies of cloud radiative forcing with a general circulation model. *Tellus*, *39A*, 460–473.

Chevallier, F., Morcrette, J.-J., Chéruy, F., & Scott, N. A. (2000). Use of a neural-network-based long-wave radiative-transfer scheme in the ecmwf atmospheric model. *Quarterly Journal of the Royal Meteorological Society*, *126*(563), 761-776. Retrieved from https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49712656318 doi: 10.1002/qj.49712656318

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., ... Strand, W. G. (2020). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001916. Retrieved 2020-02-10, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001916 doi: 10.1029/2019MS001916

Feingold, G., Tzivion, S., & Leviv, Z. (1988, November). Evolution of Raindrop Spectra. Part I: Solution to the Stochastic Collection/Breakup Equation Using the Method of Moments. *J. Atmos. Sci.*, *45*(22), 3387–3399. Retrieved 2020-07-03, from https://journals.ametsoc.org/jas/article/45/22/3387/22330/Evolution-of-Raindrop-Spectra-Part-I-Solution-to (Publisher: American Meteorological Society) doi: 10.1175/1520-0469(1988)045⟨3387:EORSPI⟩2.0.CO;2

Forbes, R. M., & Ahlgrimm, M. (2014, September). On the Representation of High-Latitude Boundary Layer Mixed-Phase Cloud in the ECMWF Global Model. *Monthly Weather Review*, *142*(9), 3425–3445. Retrieved 2015-05-04, from http://journals.ametsoc.org/doi/abs/10.1175/MWR-D-13-00325.1 doi: 10.1175/MWR-D-13-00325.1

Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G., & Yacalis, G. (2018). Could Machine Learning Break the Convection Parameterization Deadlock? *Geophysical Research Letters*, *45*(11), 5742–5751. Retrieved 2019-11-28, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL078202 doi: 10.1029/2018GL078202

Gettelman, A. (2015, November). Putting the clouds back in aerosol–cloud interactions. *Atmos. Chem. Phys.*, *15*(21), 12397–12411. Retrieved 2016-01-07, from http://www.atmos-chem-phys.net/15/12397/2015/ doi: 10.5194/acp-15-12397-2015

Gettelman, A., Bardeen, C. G., McCluskey, C. S., & Jarvinen, E. (2020). Simulating Observations of Southern Ocean Clouds and Implications for Climate. *J. Geophys. Res.*. doi: 10.1029/2020JD032619

Gettelman, A., Hannay, C., Bacmeister, J. T., Neale, R. B., Pendergrass, A. G., Danabasoglu, G., ... Mills, M. J. (2019). High Climate Sensitivity in the Community Earth System Model Version 2 (CESM2). *Geophysical Research Letters*, *46*(14), 8329–8337. Retrieved 2020-05-03, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL083978 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019GL083978) doi: 10.1029/2019GL083978

Gettelman, A., & Morrison, H. (2015). Advanced Two-Moment Bulk Microphysics for Global Models. Part I: Off-Line Tests and Comparison with Other Schemes. *J. Climate*, *28*(3), 1268–1287. Retrieved 2015-07-16, from http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-14-00102.1 doi: 10.1175/JCLI-D-14-00102.1

Gettelman, A., Morrison, H., Santos, S., Bogenschutz, P., & Caldwell, P. M. (2015). Advanced Two-Moment Bulk Microphysics for Global Models. Part II: Global Model Solutions and Aerosol–Cloud Interactions. *J. Climate*, *28*(3), 1288–1307. Retrieved 2015-07-16, from http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-14-00103.1 doi: 10.1175/JCLI-D-14-00103.1

Gettelman, A., & Sherwood, S. C. (2016, October). Processes Responsible for Cloud Feedback. *Curr Clim Change Rep*, 1–11. Retrieved 2016-10-03, from http://link.springer.com/article/10.1007/s40641-016-0052-8 doi: 10.1007/s40641-016-0052-8

Golaz, J.-C., Larson, V. E., & Cotton, W. R. (2002). A PDF-Based Model for

Boundary Layer Clouds. Part II: Model Results. *J. Atmos. Sci.*, *59*, 3552–3571.

Grabowski, W. W., Morrison, H., Shima, S.-I., Abade, G. C., Dziekan, P., & Pawlowska, H. (2019). Modeling of Cloud Microphysics: Can We Do Better? *Bull. Amer. Meteor. Soc.*, *100*(4), 655–672. Retrieved 2020-05-04, from https://journals.ametsoc.org/doi/full/10.1175/BAMS-D-18-0005.1 (Publisher: American Meteorological Society) doi: 10.1175/BAMS-D-18-0005.1

Harrington, J. Y., Feingold, G., & Cotton, W. R. (2000, March). Radiative Impacts on the Growth of a Population of Drops within Simulated Summertime Arctic Stratus. *J. Atmos. Sci.*, *57*(5), 766–785. Retrieved 2020-07-03, from https://journals.ametsoc.org/jas/article/57/5/766/101301/Radiative-Impacts-on-the-Growth-of-a-Population-of (Publisher: American Meteorological Society) doi: 10.1175/1520-0469(2000)057⟨0766:RIOTGO⟩2.0.CO;2

Hoose, C., Kristjánsson, J. E., Chen, J.-P., & Hazra, A. (2010, March). A Classical-Theory-Based Parameterization of Heterogeneous Ice Nucleation by Mineral Dust, Soot, and Biological Particles in a Global Climate Model. *J. Atmos. Sci.*, *67*(8), 2483–2503. Retrieved 2015-10-28, from http://journals.ametsoc.org/doi/abs/10.1175/2010JAS3425.1 doi: 10.1175/2010JAS3425.1

Iacono, M. J., Mlawer, E. J., Clough, S. A., & Morcrette, J.-J. (2000). Impact of an improved longwave radiation model, RRTM, on the energy budget and thermodynamic properties of the NCAR community climate model, CCM3. *jgr*, *105*(D11), 14,873–14,890.

Karset, I. H. H., Gettelman, A., Storelvmo, T., Alterskjær, K., & Berntsen, T. K. (2020). Exploring Impacts of Size-Dependent Evaporation and Entrainment in a Global Model. *Journal of Geophysical Research: Atmospheres*, *125*(4), e2019JD031817. Retrieved 2020-06-08, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JD031817 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019JD031817) doi: 10.1029/2019JD031817

Kessler, E. (1969). On the Distribution and Continuity of Water Substance in Atmospheric Circulations. In E. Kessler (Ed.), *On the Distribution and Continuity of Water Substance in Atmospheric Circulations* (pp. 1–84). Boston, MA: American Meteorological Society. Retrieved 2018-11-25, from https://doi.org/10.1007/978-1-935704-36-2_1 doi: 10.1007/978-1-935704-36-2_1

Khairoutdinov, M. F., & Kogan, Y. (2000). A new cloud physics parameterization in a large-eddy simulation model of marine stratocumulus. *Monthly Weather Review*, *128*, 229–243.

Kogan, Y. (2013). A Cumulus Cloud Microphysics Parameterization for Cloud-Resolving Models. *J. Atmos. Sci.*, *70*(5), 1423–1436. Retrieved 2015-07-16, from http://journals.ametsoc.org/doi/abs/10.1175/JAS-D-12-0183.1 doi: 10.1175/JAS-D-12-0183.1

Kostinski, A. B., & Shaw, R. A. (2005, February). Fluctuations and Luck in Droplet Growth by Coalescence. *Bull. Amer. Meteor. Soc.*, *86*(2), 235–244. Retrieved 2020-07-28, from https://journals.ametsoc.org/bams/article/86/2/235/58436/Fluctuations-and-Luck-in-Droplet-Growth-by (Publisher: American Meteorological Society) doi: 10.1175/BAMS-86-2-235

Krasnopolsky, V. M., Fox-Rabinovitz, M. S., & Chalikov, D. V. (2005, 05). New Approach to Calculation of Atmospheric Model Physics: Accurate and Fast Neural Network Emulation of Longwave Radiation in a Climate Model. *Monthly Weather Review*, *133*(5), 1370-1383. Retrieved from https://doi.org/10.1175/MWR2923.1 doi: 10.1175/MWR2923.1

Larson, V. E., Golaz, J.-C., & Cotton, W. R. (2002, December). Small-Scale and Mesoscale Variability in Cloudy Boundary Layers: Joint Probability Density Functions. *J. Atmos. Sci.*, *59*(24), 3519–3539. Retrieved 2016-11-07, from http://journals.ametsoc.org/doi/full/10.1175/1520-0469%282002%

29059%3C3519%3ASSAMVI%3E2.0.CO%3B2 doi: 10.1175/1520-0469(2002)059⟨3519:SSAMVI⟩2.0.CO;2

Lebo, Z. J., & Seinfeld, J. H. (2011, June). Theoretical basis for convective invigoration due to increased aerosol concentration. *Atmospheric Chemistry and Physics*, *11*(11), 5407–5429. Retrieved 2015-04-19, from http://www.atmos-chem-phys.net/11/5407/2011/ doi: 10.5194/acp-11-5407-2011

Lebsock, M. D., & L'Ecuyer, T. S. (2011). The retrieval of warm rain from CloudSat. *Journal of Geophysical Research: Atmospheres*, *116*(D20). Retrieved 2020-07-24, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JD016076 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2011JD016076) doi: 10.1029/2011JD016076

Liu, X., & Penner, J. E. (2005). Ice Nucleation Parameterization for Global Models. *Meteor. Z.*, *14*(499-514).

Liu, Y., & Daum, P. H. (2004, July). Parameterization of the Autoconversion Process.Part I: Analytical Formulation of the Kessler-Type Parameterizations. *J. Atmos. Sci.*, *61*(13), 1539–1548. Retrieved from http://journals.ametsoc.org/doi/full/10.1175/1520-0469(2004)061%3C1539:POTAPI%3E2.0.CO;2 doi: 10.1175/1520-0469(2004)061⟨1539:POTAPI⟩2.0.CO;2

Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., … Kato, S. (2018). Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 Data Product. *J. Climate*, *31*(2), 895–918. Retrieved 2019-02-27, from http://journals.ametsoc.org/doi/full/10.1175/JCLI-D-17-0208.1 doi: 10.1175/JCLI-D-17-0208.1

Michibata, T., & Takemura, T. (2015, September). Evaluation of autoconversion schemes in a single model framework with satellite observations. *J. Geophys. Res. Atmos.*, *120*(18), 2015JD023818. Retrieved 2015-10-26, from http://onlinelibrary.wiley.com/doi/10.1002/2015JD023818/abstract doi: 10.1002/2015JD023818

Neale, R. B., Richter, J. H., & Jochum, M. (2008). The Impact of Convection on ENSO: From a Delayed Oscillator to a Series of Events. *J. Climate*, *21*, 5904–+. doi: 10.1175/2008JCLI2244.1

Pruppacher, H. R., & Klett, J. D. (1997). *Microphysics of Clouds and Precipitation* (2nd ed.). Kluwer Academic.

Rasp, S., Pritchard, M. S., & Gentine, P. (2018, September). Deep learning to represent subgrid processes in climate models. *PNAS*, *115*(39), 9684–9689. Retrieved 2019-11-28, from https://www.pnas.org/content/115/39/9684 doi: 10.1073/pnas.1810286115

Reisin, T., Levin, Z., & Tzivion, S. (1996, February). Rain Production in Convective Clouds As Simulated in an Axisymmetric Model with Detailed Microphysics. Part I: Description of the Model. *J. Atmos. Sci.*, *53*(3), 497–519. Retrieved 2020-07-03, from https://journals.ametsoc.org/jas/article/53/3/497/24062/Rain-Production-in-Convective-Clouds-As-Simulated (Publisher: American Meteorological Society) doi: 10.1175/1520-0469(1996)053⟨0497:RPICCA⟩2.0.CO;2

Rosenfeld, D., Wang, H., & Rasch, P. J. (2012, July). The roles of cloud drop effective radius and LWP in determining rain properties in marine stratocumulus. *Geophysical Research Letters*, *39*(13), n/a–n/a. Retrieved 2015-04-19, from http://doi.wiley.com/10.1029/2012GL052028 doi: 10.1029/2012GL052028

Seifert, A., & Beheng, K. D. (2001). A double-moment parameterization for simulating autoconversion, accretion and selfcollection. *Atmos. Res.*, *59-60*, 265–281.

Shell, K. M., Kiehl, J. T., & Shields, C. A. (2008). Using the Radiative Kernel Technique to Calculate Climate Feedbacks in NCAR's Community Atmosphere Model. *J. Climate*, *21*, 2269–2282. doi: 10.1175/2007JCLI2044.1

Shi, X., Liu, X., & Zhang, K. (2015, February). Effects of pre-existing ice crystals on cirrus clouds and comparison between different ice nucleation parameterizations with the Community Atmosphere Model (CAM5). *Atmospheric Chemistry and Physics*, *15*(3), 1503–1520. Retrieved 2015-04-19, from http://www.atmos-chem-phys.net/15/1503/2015/ doi: 10.5194/acp-15-1503-2015

Soden, B. J., Held, I. M., Colman, R., Shell, K. M., Kiehl, J. T., & Shields, C. A. (2008). Quantifying Climate Feedbacks Using Radiative Kernels. *J. Climate*, *21*(14), 3504–3520. doi: 10.1175/2007JCLI2110.1

Stephens, G. L. (2005). Cloud Feedbacks in the Climate System: A Critical Review. *joc*, *18*(2), 237–273.

Stephens, G. L., L'Ecuyer, T., Forbes, R., Gettelman, A., Golaz, J.-C., Bodas-Salcedo, A., . . . Haynes, J. (2010). Dreary state of precipitation in global models. *jgr*, *115*(D24211). doi: 10.1029/2010JD014532

Stevens, B., & Feingold, G. (2009). Untangling aerosol effects on clouds and precipitation in a buffered system. *Nature*, *461*(1), 607–613.

Stevens, B., Feingold, G., Cotton, W. R., & Walko, R. L. (1996, April). Elements of the Microphysical Structure of Numerically Simulated Nonprecipitating Stratocumulus. *J. Atmos. Sci.*, *53*(7), 980–1006. Retrieved 2020-07-03, from https://journals.ametsoc.org/jas/article/53/7/980/24252/Elements-of-the-Microphysical-Structure-of (Publisher: American Meteorological Society) doi: 10.1175/1520-0469(1996)053⟨0980:EOTMSO⟩2.0.CO;2

Tan, I., Storelvmo, T., & Zelinka, M. D. (2016, April). Observational constraints on mixed-phase clouds imply higher climate sensitivity. *Science*, *352*(6282), 224–227. Retrieved 2016-04-08, from http://science.sciencemag.org/content/352/6282/224 doi: 10.1126/science.aad5300

Twomey, S. (1977). The influence of pollution on the shortwave albedo of clouds. *J. Atmos. Sci.*, *34*(7), 1149–1152.

Tzivion, S., Feingold, G., & Levin, Z. (1987, November). An Efficient Numerical Solution to the Stochastic Collection Equation. *J. Atmos. Sci.*, *44*(21), 3139–3149. Retrieved 2020-07-03, from https://journals.ametsoc.org/jas/article/44/21/3139/21838/An-Efficient-Numerical-Solution-to-the-Stochastic (Publisher: American Meteorological Society) doi: 10.1175/1520-0469(1987)044⟨3139:AENSTT⟩2.0.CO;2

Tzivion, S., Feingold, G., & Levin, Z. (1989, November). The Evolution of Raindrop Spectra. Part II: Collisional Collection/Breakup and Evaporation in a Rainshaft. *J. Atmos. Sci.*, *46*(21), 3312–3328. Retrieved 2020-07-03, from https://journals.ametsoc.org/jas/article/46/21/3312/22566/The-Evolution-of-Raindrop-Spectra-Part-II (Publisher: American Meteorological Society) doi: 10.1175/1520-0469(1989)046⟨3312:TEORSP⟩2.0.CO;2

Tzivion, S., Reisin, T. G., & Levin, Z. (1999, January). A Numerical Solution of the Kinetic Collection Equation Using High Spectral Grid Resolution: A Proposed Reference. *Journal of Computational Physics*, *148*(2), 527–544. Retrieved 2020-07-03, from http://www.sciencedirect.com/science/article/pii/S0021999198961289 doi: 10.1006/jcph.1998.6128

Wang, Y., Liu, X., Hoose, C., & Wang, B. (2014, October). Different contact angle distributions for heterogeneous ice nucleation in the Community Atmospheric Model version 5. *Atmos. Chem. Phys.*, *14*(19), 10411–10430. Retrieved 2017-05-01, from http://www.atmos-chem-phys.net/14/10411/2014/ doi: 10.5194/acp-14-10411-2014

Wilkinson, M. (2016, January). Large Deviation Analysis of Rapid Onset of Rain Showers. *Phys. Rev. Lett.*, *116*(1), 018501. Retrieved 2020-07-28, from https://link.aps.org/doi/10.1103/PhysRevLett.116.018501 (Publisher: American Physical Society) doi: 10.1103/PhysRevLett.116.018501

Wood, N. B., L'Ecuyer, T. S., Heymsfield, A. J., Stephens, G. L., Hudak, D. R.,

& Rodriguez, P. (2014). Estimating snow microphysical properties using collocated multisensor observations. *Journal of Geophysical Research: Atmospheres*, *119*(14), 8941–8961. Retrieved 2020-07-24, from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JD021303 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2013JD021303) doi: 10.1002/2013JD021303

Yin, Y., Levin, Z., Reisin, T. G., & Tzivion, S. (2000, March). The effects of giant cloud condensation nuclei on the development of precipitation in convective clouds — a numerical study. *Atmospheric Research*, *53*(1), 91–116. Retrieved 2020-07-03, from http://www.sciencedirect.com/science/article/pii/S0169809599000460 doi: 10.1016/S0169-8095(99)00046-0

Yuval, J., & O'Gorman, P. A. (n.d.). Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions. *Nature Communications*, *11*, 3295. Retrieved from https://doi.org/10.1038/s41467-020-17142-3

Zhang, G. J., & McFarlane, N. A. (1995). Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian Climate Center general circulation model. *Atmos. Ocean*, *33*, 407–446.